Statistics and Applications {ISSN 2454-7395 (online)}
Volume 22, No. 1, 2024 (New Series), pp 263–277
http://www.ssca.org.in/journal

Asymptotic Confidence Interval Approach to Estimate the Portion of Area Under Multi-Class ROC Curve

Arunima S. Kannan and R. Vishnu Vardhan
Department of Statistics, Pondicherry University, Puducherry

Received: 26 October 2022; Revised: 03 September 2023; Accepted: 06 September 2023

Abstract

The area under the curve (AUC) gives an overall summary measure of the performance of the Receiver Operating Characteristic (ROC) curve. AUC summarizes the entire area under the curve. Sometimes clinical studies need to focus on the area with low FPR and high TPR rates. To find the area of portion of an ROC curve, partial AUC (pAUC) came into use. The seminal works on estimating the pAUC was in the framework of Bi-normal ROC curve. However, in a real-life scenario, we may come across non-normality, and the data may be of multi-class. In such cases the existing methodology of binormal ROC curve will not be of use and this creates the need to bring out a new methodology for estimating the pAUC under non-normal data. In this paper we made an attempt to address the above point and derived the expressions for the pAUC of multi-class ROC curve. Further estimating the partial AUC has been carried out by means of asymptotic confidence intervals of the false positive rates. Adding to this the constraint on TPR has been considered to elicit the focused area of the ROC curve and termed it as two-way pAUC (TpAUC). Good amount of simulations and two real datasets have been considered for necessary illustrations.

Key words: AUC; Exponential; Multi-class; pAUC; ROC; TpAUC.

AMS Subject Classifications: 62P10

1. Introduction

The Receiver Operating Characteristic (ROC) curve is the popular classification tool to evaluate the performance of a diagnostic test/marker. ROC curve was first used for signal detection during World War II (Peterson *et al.*, 1954; Tanner and Swets, 1954). ROC curve analysis was introduced into diagnostic medicine by Lusted (1971), and its applications can be seen in several clinical domains that rely extensively on screening and diagnostic procedures, laboratory testing, epidemiology, radiology, etc.(Obuchowski, 2003).

ROC curve is generated using the coordinate pairs, namely the false positive rates (FPRs) and true positive rates (TPRs), which are usually referred as intrinsic measures. Out of these thresholds, one has to choose a threshold that can give better accuracy with reasonable values of false positives and true positives. AUC is the summary measure of the ROC

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

curve, which is used to determine the performance/accuracy of a diagnostic test/marker. AUC has a theoretical value lies between 0 and 1. An AUC of 0.5 indicates that the classification is random, the accuracy of a diagnostic test or procedure will increase as the value of AUC gets closer to 1.

Let us assume that H denote the population 1 with distribution function F and D be the population 2 with distribution function G; then the ROC form is given as

$$y(t) = G(F_0^{-1}(t)), 0 \leq t \leq 1 \tag{1}$$

where, $G(x) = \int_x^\infty g(x)dx$ and $F_0(y) = \int_y^\infty f(y)dy$; *g(x)* and *f(y)* is the density functions of H and D populations respectively. The probability of detecting/ identifying a subject with condition is called as the $TPR = P(S > t/D)$, and the probability of classifying a subject is $FPR = P(S > t/H)$. Here S denotes the data value or score observed from a subject and t is the threshold. Each data point in the ROC serves as a threshold point, with which one can calculate the TPRs and FPRs. Yet there is a need to determine the optimal threshold among the set of all possible thresholds, which is done by using Youden's J index. The optimal threshold is determined by taking the maximum value of Youden's J index from a vector of values obtained from (2). Now, the FPR and TPR values corresponding to this optimal threshold have to be considered as the optimal FPR and TPR.

$$J = maximum\,(TPR(t) - FPR(t)) \tag{2}$$

which is the maximum distance between the curve and the chance line. With this optimal threshold, the subjects will be classified with the atmost accuracy and can also be used to assign the status of unspecified subjects. When J = 1, the test is perfect, meaning there are no false positives or negatives.
The AUC of an ROC curve is defined as

$$AUC = \int_0^1 ROC(t)\,dt$$

AUC is defined as the average TPR value for all possible TNR (1-FPR) values, which will consider the entire area under the curve (Obuchowski, 2003; Zhou *et al.*, 2009; McClish, 1989; Obuchowski and Bullen, 2018). Analyzing the entire ROC curve involves both strict and lax thresholds; hence, considering a portion of ROC curve will be more meaningful in some instances, and such portion is named as partial AUC (pAUC) (see Figure 2). From Figure 1, we can see that the lax threshold provides high TPRs and high FPRs, which are not a region of interest for clinical studies. Since most clinical studies involve living subjects, the FPR must be reasonably low. However, the strict threshold provides low FPR and TPR values, which also does not indicate better classification. Hence, this generates the need to speak about the portion of the ROC curve above the strict thresholds and below the lax thresholds, which is shown in Figure 2 for some arbitrary values of c_1 and c_2 from the set of FPRs. pAUC is now employed in numerous medical applications and has gained popularity, particularly in screening research (Ricamato and Tortorella, 2011).

The pAUC consider the area of the ROC space where data have been observed or that correspond to clinically significant TPR or TNR values. For example, only the lower tail of the ROC curve is of interest for cancer screening because the FPR must be minimal to

be acceptable (Zhang *et al.*, 2018). Baker and Pinsky (2001) also pointed out that low FPR needs to be maintained in cancer screening studies, which is important because it will avoid costly biopsies. In such cases, analyzing a restricted portion will be more meaningful than analyzing the entire area of the ROC curve. Seminal work on pAUC was done by McClish (1989), where a method of analyzing the portion of the ROC curve was proposed and also gave a transformation to obtain the standardized pAUC value. Thompson and Zucchini (1989) introduced the method to estimate the partial area under the binormal ROC curve over any specified region of interest. Later, Jiang *et al.* (1996) adopted the methodology of McClish's work and extended it to describe the partial area index for highly sensitive diagnostic test. Hillis and Metz (2012) derived analytic expressions to estimate the pAUC under the assumption of a latent binormal model.

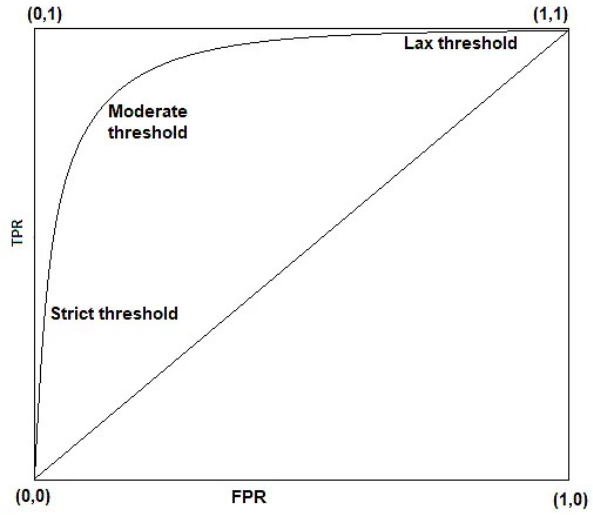


Figure 1: ROC curve depicting strict, moderate and lax thresholds

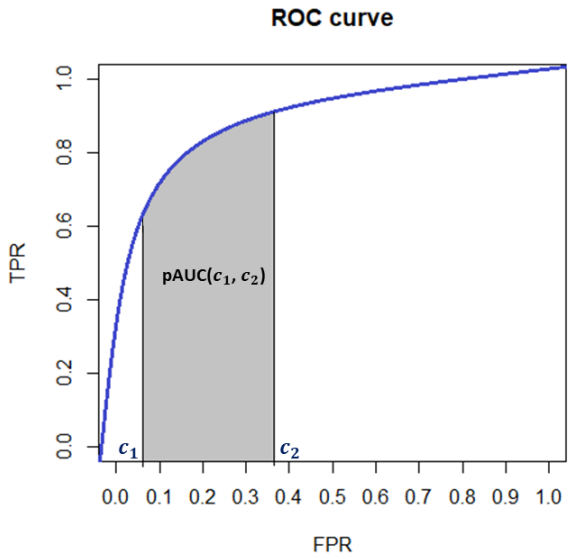


Figure 2: A typical plot of pAUC between a fixed range of FPR

In practice, diagnostic tests with high FPR lead to enormous economic expenses

because a significant fraction of healthy individuals would use up the limited supply of medical treatments. Furthermore, when diagnosing a fatal disease, failure to correctly identify severely ill patients (poor TPR) will result in severe ethical ramifications. Therefore, in such cases, it is necessary to simultaneously maintain FPR and TPR at low and high levels, respectively. So here, we have introduced a method to estimate the partial area by considering the constraints on both FPR and TPR simultaneously, which is termed as Two-Way pAUC (TpAUC). A diagrammatic representation of TpAUC is given in Figure 3 which considered the area of the ROC curve with FPR range (c_1,c_2) and minimum TPR of d_0, it can be denoted as TpAUC(c_1,c_2,d_0).

In literature, the works mainly focus on estimating the partial area of the binormal ROC curve, which can only be used when the data consists of two classes and follows normality. However, in a real-life scenario, we may come across non-normality, and the data may be of multi-class. In such cases the existing methodology of binormal ROC curve will not be of use and this creates the need to bring out a new methodology for estimating the pAUC under non-normal data. Estimation of AUC under the multi-class classification where data tend to follow normal distribution was addressed by Gönen (2013); Cheam and McNicholas (2016); Siva and Vishnu (2022). Recently Arunima and Vishnu (2022) proposed gamma mixture ROC curve to classify the multi-class data where the population follows gamma distributions, in which the gamma variate is transformed into normal by using the Wilson-Hilferty transformation. In this paper we have considered one of the well-known life time distribution; the exponential distribution and the partial area estimation of multi-class exponential ROC is discussed in detail. The study is supported with simulated and real datasets. Before we detail out the proposed methodology, a gentle introduction on Multi-class Exponential ROC Curve is given. Thereafter, along with the proposed methodology, the numerical illustrations are discussed in subsequent sections.

1.1. Multi-class exponential ROC curve

Let us assume that population 1, H $\sim exp(\theta_0)$ and population 2 has two sub populations namely D_1 and D_2 such that, $D_1 \sim exp(\theta_1)$ and $D_2 \sim exp(\theta_2)$. Then the expressions for intrinsic measures of mixture Exponential ROC (mEROC) are defined Arunima and Vishnu (2023) as below.
FPR of the mEROC (mFPR) is given as

$$mFPR = \pi_1 FPR_1 + \pi_2 FPR_2$$

where

$$FPR_1 = x(t_1) = P(S > t_1 \mid H) = e^{-\theta_0 t_1} \tag{3}$$

$$FPR_2 = x(t_2) = P(S > t_2 \mid D_1) = e^{-\theta_1 t_2} \tag{4}$$

where π_is are mixing proportions/weights; t_1 and t_2 are the respective threshold values for the classification of (H, D_1) and (D_1, D_2) respectively. From (3) and (4) we can write t_1 and t_2 as

$$t_1 = -\frac{log(x(t_1))}{\theta_0} \quad ; \quad t_2 = -\frac{log(x(t_2))}{\theta_1} \tag{5}$$

TPR of mEROC (mTPR) is given as

$$mTPR = \pi_1 TPR_1 + \pi_2 TPR_2$$

where

$$TPR_1 = y(t_1) = P(S > t_1 \mid D_1) = e^{-\theta_1 t_1} \tag{6}$$
$$TPR_2 = y(t_2) = P(S > t_2 \mid D_2) = e^{-\theta_2 t_2} \tag{7}$$

substituting (5) in (6) and (7) we will get the mEROC curve which be written as,

$$mEROC = \pi_1 x(t_1)^{\beta_1} + \pi_2 x(t_2)^{\beta_2}$$

where $\beta_1 = \frac{\theta_1}{\theta_0}$ and $\beta_2 = \frac{\theta_2}{\theta_1}$.
accuracy can be expressed notationally as

$$mAUC = \int_0^1 ROC(t)\, dt = \pi_1 \frac{\theta_0}{\theta_0 + \theta_1} + \pi_2 \frac{\theta_1}{\theta_1 + \theta_2}$$

2. Proposed methodology - partial area of mEROC curve

Let c_1 and c_2 denote any two arbitrary FPR values, then pAUC for mEROC can be defined as

$$mA_{(c_1,c_2)} = \pi_1 A_{1_{(c_1,c_2)}} + \pi_2 A_{2_{(c_1,c_2)}}$$

where $A_{1(c_1,c_2)}$ and $A_{2(c_1,c_2)}$ are the partial areas of H & D_1 and D_1 & D_2 respectively. which is defined as

$$\begin{aligned} A_{1_{(c_1,c_2)}} &= \int_{c_1}^{c_2} TPR_1(t)\, FPR_1'(t)\ dt = \int_{c_1}^{c_2} e^{-\theta_1 t}\, e^{-\theta_0 t}\, (-\theta_0)\, dt \\ &= \frac{\theta_0}{\theta_1 + \theta_0} \left[e^{-(\theta_0+\theta_1)c_2} - e^{-(\theta_0+\theta_1)c_1} \right] \end{aligned}$$

and

$$\begin{aligned} A_{2_{(c_1,c_2)}} &= \int_{c_1}^{c_2} TPR_2(t)\, FPR_2'(t)\ dt = \int_{c_1}^{c_2} e^{-\theta_2 t}\, e^{-\theta_1 t}\, (-\theta_1)\, dt \\ &= \frac{\theta_1}{\theta_2 + \theta_1} \left[e^{-(\theta_1+\theta_2)c_2} - e^{-(\theta_1+\theta_2)c_1} \right] \end{aligned}$$

The area (A+B) in Figure 3 indicates the pAUC between the FPR range c_1 and c_2. To this, one more additional constraint is added in the form of d_0. The area generated between c_1, c_2 and d_0 is termed as the two-way pAUC and denoted by $mTpAUC_{(c_1,c_2,d_0)}$. Here c_2 is the upper limit of FPR; d_0 is the lower limit of TPR, c_1 is the corresponding FPR value at d_0. The area encapsulated between the triplet combination (c_1, c_2, d_0) is indicated as B in Figure 3 and it can be obtained as

$$\begin{aligned} mTpAUC_{(c_1,c_2,d_0)} &= Area\ (A + B) - Area\ A \\ &= mA_{(c_1,c_2)} - [c_2 - c_1]d_0 \end{aligned}$$

In the above expression the constants c_1, c_2 and d_0 are to be estimated. To do so we need to employ two ways; one is arbitrarily choosing the values and the other is to estimate them using the method of asymptotic confidence interval. Below we describe the way of determining the c_1, c_2 and d_0.

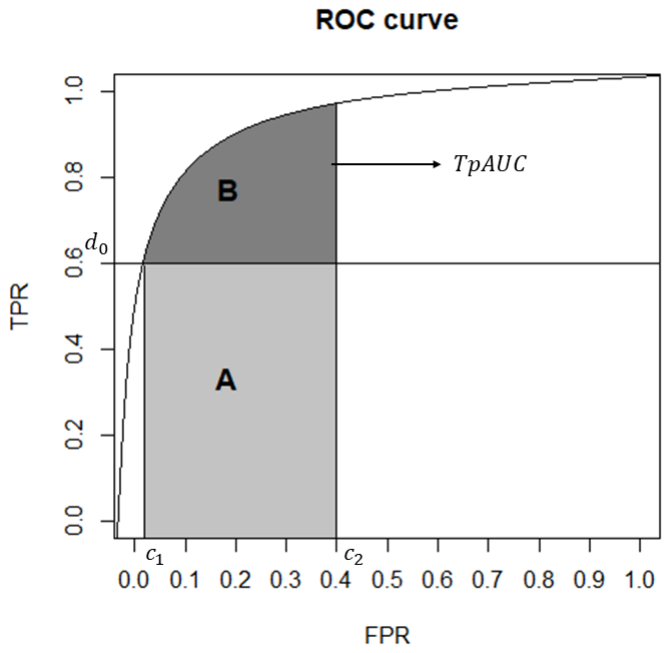


Figure 3: Depicting the area of A and B

Method I

In this method the c_2 and d_0 will be chosen arbitrarily and c_1 can be obtained from d_0. Instead of choosing arbitrarily one can impute using the knowledge from previous studies. In general clinicians prefer to have reasonably law FPR and moderate/high TPR.

Method II

Here we introduce the asymptotic confidence interval approach to define c_2 and d_0. For which the asymptotic confidence intervals are defined and respective variances for mFPR and mTPR are derived. The upper limit of mFPR will be taken as c_2 and lower limit of mTPR will be the d_0; the corresponding mFPR value is taken as c_1.

Asymptotic confidence intervals for mTPR and mFPR

The $100(1-\alpha)\%$ asymptotic confidence interval for mTPR is

$$mTPR \pm Z_{1-(\frac{\alpha}{2})}\sqrt{Var(mTPR)}$$

where $Z_{1-\frac{\alpha}{2}}$ is the $1-\frac{\alpha}{2}$ standard normal percentile and by delta method (Miller Jr, 1981), we can obtain variance of $mTPR$.

$$Var(mTPR) = \pi_1 Var(TPR_1) + \pi_2 Var(TPR_2)$$

where

$$Var(TPR_1) = \left(\frac{\partial TPR_1}{\partial \theta_0}\right)^2 Var(\theta_0) + \left(\frac{\partial TPR_1}{\partial \theta_1}\right)^2 Var(\theta_1)$$
$$= \left\{ \frac{e^{-\theta_1 \frac{(ln(\theta_0)-ln(\theta_1))}{\theta_0-\theta_1}} \left[\theta_1 \left(\theta_0 - \theta_1 - \theta_0(ln(\theta_0) + ln(\theta_1))\right)\right]}{\theta_0(\theta_0-\theta_1)^2} \right\}^2 \frac{{\theta_0}^2}{n_0}$$
$$+ \left\{ \frac{e^{-\theta_1 \frac{(ln(\theta_0)-ln(\theta_1))}{\theta_0-\theta_1}} \left[\theta_1 - \theta_0 ln(\theta_0) - \theta_0 + \theta_0 ln(\theta_0)\right]}{(\theta_0-\theta_1)^2} \right\}^2 \frac{{\theta_1}^2}{n_1}$$

$$Var(TPR_2) = \left(\frac{\partial TPR_2}{\partial \theta_1}\right)^2 Var(\theta_1) + \left(\frac{\partial TPR_2}{\partial \theta_2}\right)^2 Var(\theta_2)$$
$$= \left\{ \frac{e^{-\theta_2 \frac{(ln(\theta_1)-ln(\theta_2))}{\theta_1-\theta_2}} \left[\theta_2 \left(\theta_1 - \theta_2 - \theta_1(ln(\theta_1) + ln(\theta_2))\right)\right]}{\theta_1(\theta_1-\theta_2)^2} \right\}^2 \frac{{\theta_1}^2}{n_0}$$
$$+ \left\{ \frac{e^{-\theta_2 \frac{(ln(\theta_1)-ln(\theta_2))}{\theta_1-\theta_2}} \left[\theta_2 - \theta_1 ln(\theta_1) - \theta_1 + \theta_1 ln(\theta_1)\right]}{(\theta_1-\theta_2)^2} \right\}^2 \frac{{\theta_2}^2}{n_1}$$

For the method II we choose d_0 as lower limit of mTPR

$$d_0 = mTPR - Z_{1-(\frac{\alpha}{2})}\sqrt{Var(mTPR)} \tag{8}$$

then c_1 will be the corresponding FPR of the d_0.

Similarly, $100(1-\alpha)\%$ asymptotic confidence interval for $mFPR$ is,

$$mFPR \pm Z_{1-(\frac{\alpha}{2})}\sqrt{Var(mFPR)}$$

$$Var(mFPR) = \pi_1 Var(FPR_1) + \pi_2 Var(FPR_2)$$

where

$$Var(FPR_1) = \left(\frac{\partial FPR_1}{\partial \theta_0}\right)^2 Var(\theta_0) + \left(\frac{\partial FPR_1}{\partial \theta_1}\right)^2 Var(\theta_1)$$
$$= \left\{ \frac{e^{-\theta_0 \frac{(ln(\theta_0)-ln(\theta_1))}{\theta_0-\theta_1}} \left[\theta_0 - \theta_1 ln(\theta_0) - \theta_1 + \theta_1 ln(\theta_1)\right]}{(\theta_0-\theta_1)^2} \right\}^2 \frac{{\theta_0}^2}{n_0}$$
$$+ \left\{ \frac{e^{-\theta_0 \frac{(ln(\theta_0)-ln(\theta_1))}{\theta_0-\theta_1}} \left[\theta_0 \left(\theta_1 - \theta_0 + \theta_1 ln(\theta_0) + \theta_1 ln(\theta_1)\right)\right]}{\theta_1(\theta_0-\theta_1)^2} \right\}^2 \frac{{\theta_1}^2}{n_1}$$

and

$$Var(FPR_2) = \left(\frac{\partial FPR_2}{\partial \theta_1}\right)^2 Var(\theta_1) + \left(\frac{\partial FPR_2}{\partial \theta_2}\right)^2 Var(\theta_2)$$
$$= \left\{\frac{e^{-\theta_1 \frac{(ln(\theta_1)-ln(\theta_2))}{\theta_1-\theta_2}}\left[\theta_1 - \theta_2 ln(\theta_1) - \theta_2 + \theta_2 ln(\theta_2)\right]}{(\theta_1-\theta_2)^2}\right\}^2 \frac{{\theta_1}^2}{n_1}$$
$$+ \left\{\frac{e^{-\theta_1 \frac{(ln(\theta_1)-ln(\theta_2))}{\theta_1-\theta_2}}\left[\theta_1\left(\theta_2 - \theta_1 + \theta_2 ln(\theta_1) + \theta_2 ln(\theta_2)\right)\right]}{\theta_2(\theta_1-\theta_2)^2}\right\}^2 \frac{{\theta_2}^2}{n_2}$$

then the value of c_2 will be

$$c_2 = mFPR + Z_{1-\frac{\alpha}{2}}\sqrt{Var(mFPR)} \tag{9}$$

3. Numerical illustrations

For illustrating the proposed work both simulated and real datasets are considered and the results are tabulated accordingly.

3.1. Simulated datasets

Exponential random samples of size $n = (25, 50, 100, 200)$ are generated with different parameter combinations. The moderate and better classification scenarios will be demonstrated using the parameter combinations given in Table 1. Methods I and II are used to estimate the partial area on the samples that were generated.

For illustration purpose, we have chosen different parameter combinations of θ_0 and θ_1 and calculated the partial area by taking $c_2 = 0.5$ and $d_0 = 0.65$ (one may choose any values for c_2 and d_0 according to prior knowledge on the study). The results for each parameter combination with respective sample sizes are reported in Tables 1 and 2 and respective ROC curves are given in Figure 4. From Table 1, consider n=100 in set A, the overall $\widehat{mAUC}$ is observed to be 86.71%, with true positives about 76.61% and false positives of 13.40%, which indicates comparatively a better accuracy. In similar lines, in set B for n=100, the overall $\widehat{mAUC}$ is observed to be 68.27%, with true positives about 51.19% and false positives of 27.84%, which indicates a moderate accuracy. Table 2 gives the results pertaining to partial area for method I, and we can observe that, let say for n=100 for set A the values for d_0 and c_2 are chosen as 0.65 and 0.5 respectively, $\hat{c}_1$ is the corresponding mFPR at d_0 which is about 0.0816 provides mpAUC ($\widehat{mA}$) and $\widehat{mTpAUC}$ about 0.3786 and 0.1066 respectively. And for n=100 in set B, the corresponding mFPR at d_0 which is about 0.3911 provides mpAUC ($\widehat{mA}$) and $\widehat{mTpAUC}$ about 0.1496 and 0.0788 respectively. Here we can see that within a fixed range of mFPR and mTPR the $\widehat{mAUC}$, $\widehat{mA}_{(c_1,c_2)}$ and $\widehat{mTpAUC}_{(c_1,c_2,d_0)}$ are proportional to each other. For better understanding the $\widehat{mTpAUC}$ for n=100 of sets A and B using method II are depicted in Figures 5.

Table 1: ROC curve estimates of simulated dataset

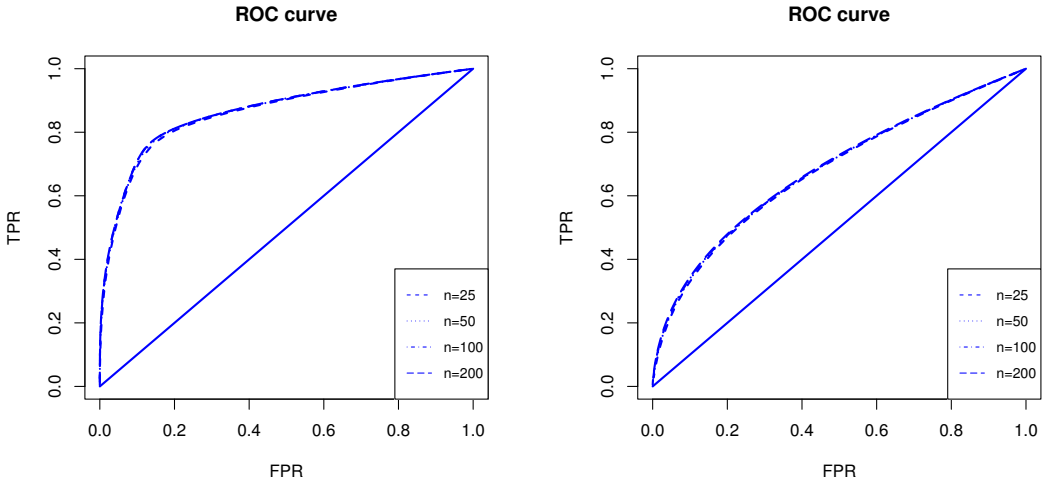
n	$\hat{\pi}_1$	$\hat{\pi}_2$	$\hat{\theta}_0$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{t}_1$	$\hat{t}_2$	$\hat{J}$	$\widehat{mFPR}$	$\widehat{mTPR}$	$\widehat{mAUC}$
Set A: $\hat{\theta}_0$=0.99; $\hat{\theta}_1$=0.3; $\hat{\theta}_2$=0.01											
25	0.5015	0.4985	1.0351	0.3136	0.0105	1.7019	11.5858	0.6277	0.1349	0.7625	0.8650
50	0.5024	0.4976	1.0151	0.3066	0.0102	1.7132	11.6439	0.6304	0.1335	0.7639	0.8662
100	0.4983	0.5017	0.9976	0.3035	0.0101	1.7264	11.6779	0.6325	0.1340	0.7664	0.8671
200	0.4986	0.5014	0.9908	0.3003	0.0101	1.7346	11.7392	0.6330	0.1342	0.7672	0.8674
Set B: $\hat{\theta}_0$=0.99; $\hat{\theta}_1$=0.6 $\hat{\theta}_2$=0.2											
25	0.5266	0.4734	1.0329	0.6287	0.2095	1.2612	2.6946	0.2485	0.2743	0.5228	0.6794
50	0.5143	0.4857	1.0122	0.6099	0.2033	1.2754	2.7382	0.2420	0.2792	0.5212	0.6831
100	0.5220	0.4780	0.9960	0.6066	0.2011	1.2823	2.7404	0.2415	0.2784	0.5199	0.6827
200	0.5109	0.4891	0.9940	0.6023	0.1997	1.2836	2.7512	0.2398	0.2819	0.5218	0.6851

Table 2: Partial area estimates of simulated datasets using method I

d_0	$\hat{c}_1$	c_2	$\widehat{mA}_{(c_1,c_2)}$	$\widehat{mTpAUC}_{(c_1,c_2,d_0)}$
Set A: $\hat{\theta}_0$=0.99; $\hat{\theta}_1$=0.3; $\hat{\theta}_2$=0.01				
0.65	0.08364	0.5	0.40231	0.13168
0.65	0.07992	0.5	0.39148	0.11842
0.65	0.08156	0.5	0.37855	0.10657
0.65	0.07934	0.5	0.37575	0.10233
Set B: $\hat{\theta}_0$=0.99; $\hat{\theta}_1$=0.6 $\hat{\theta}_2$=0.2				
0.65	0.39608	0.5	0.16230	0.09475
0.65	0.39208	0.5	0.15241	0.08226
0.65	0.39106	0.5	0.14953	0.07872
0.65	0.39220	0.5	0.14706	0.07699

Table 3: Partial area estimates of simulated datasets using method II

$Var(\widehat{mFPR})$	$Var(\widehat{mTPR})$	$\hat{d}_0$	$\hat{c}_1$	$\hat{c}_2$	$\widehat{mA}_{(c_1,c_2)}$	$\widehat{mTpAUC}_{(c_1,c_2,d_0)}$
Set A: $\hat{\theta}_0$=0.99; $\hat{\theta}_1$=0.3; $\hat{\theta}_2$=0.01						
0.00179	0.00206	0.68016	0.09353	0.21719	0.09823	0.01412
0.00096	0.00094	0.70578	0.10191	0.19527	0.07232	0.00642
0.00048	0.00046	0.72367	0.10988	0.17725	0.05133	0.00258
0.00025	0.00025	0.73621	0.11354	0.16539	0.03902	0.00084
Set B: $\hat{\theta}_0$=0.99; $\hat{\theta}_1$=0.6 $\hat{\theta}_2$=0.2						
0.00314	0.00302	0.41063	0.15767	0.35576	0.24532	0.16397
0.00118	0.00152	0.45334	0.17771	0.30952	0.15247	0.09272
0.00075	0.00116	0.46747	0.18847	0.29418	0.11795	0.06854
0.00046	0.00070	0.47924	0.20080	0.28167	0.08965	0.05089



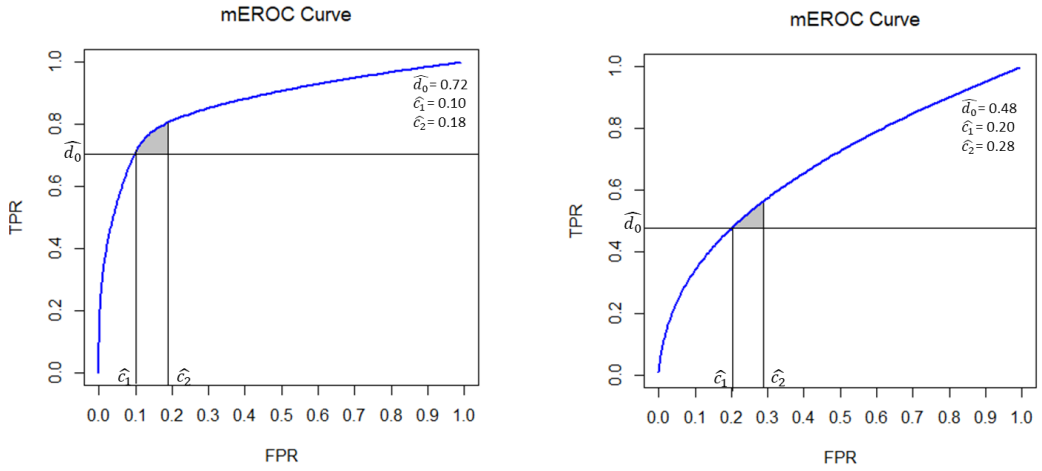
(a) Set A **(b) Set B**

Figure 4: ROC curve for simulated datasets

Coming to method II, from Table 3, we can observe that for n=100 in set A, by using the equations (8) and (9) the obtained value are $\hat{d}_0 = 0.7237$, $\hat{c}_2 = 0.1773$ and the respective $\widehat{mFPR}$ at $\hat{d}_0$ is $\hat{c}_1 = 0.1099$, and provides $\widehat{mA}_{(c_1,c_2)}$ and $\widehat{mTpAUC}_{(c_1,c_2,d_0)}$ of about 0.05133 and 0.00258 respectively. Similarly for n=100 in set B, the obtained value of $\hat{d}_0$, $\hat{c}_1$ and $\hat{c}_2$ are 0.46747, 0.18847 and 0.29418 which provides $\widehat{mA}_{(c_1,c_2)}$ and $\widehat{mTpAUC}_{(c_1,c_2,d_0)}$ of about 0.1179 and 0.06854 respectively.

3.2. Real datasets

To demonstrate the proposed methodology two real datasets are considered and their results are tabulated accordingly with respective ROC curves in Figure 7.



(a) Set A **(b) Set B**

Figure 5: TpAUC for set A and B for n=100

Data 1: Irradiated mice data

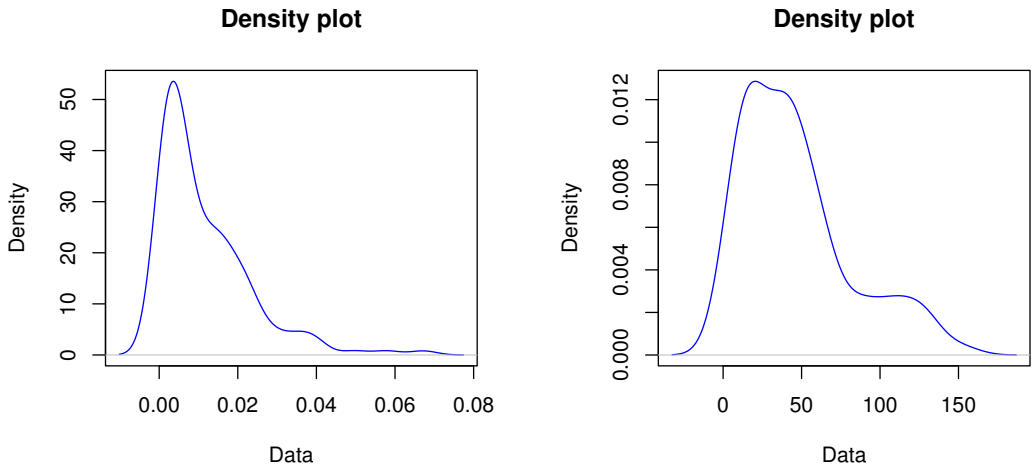
Irradiated mice data from Elandt-Johnson and Johnson (1980) is considered; the variable of interest is the time at the death of 99 mice. The p-value of Kolmogorov-Smirnov test for exponential distribution is 0.43 (test statistic (D)=0.113) which indicates that the data follows exponential distribution.

The density plot of the irradiated mice data is given in the Figure 6(a), and is very clear that there exists multi-modality which indicates the presence of sub-populations, i.e., the data is of multi-class. By using EM algorithm we identified that there are three classes and the estimated the parameters of the respective populations are $\hat{\theta}_0 = 0.3954$, $\hat{\theta}_1 = 0.0279$ and $\hat{\theta}_2 = 0.0201$. The proposed methodology is used to classify the data and the results are tabulated in the Tables 4 and 5.

It is observed that with thresholds $\hat{t}_1 = 12.8556$ and $\hat{t}_2 = 417.2446$, the overall $\widehat{mAUC}$ is about 0.7267, this indicates that the mEROC curve has accuracy about 72%, with false positives of 24% and true positives of 69%. This means to that a subject can be classified in the following manner.

$$\text{It is classified as} = \begin{cases} P_1, \text{ if } S \leq 12.8556 \\ P_2, \text{ if } 12.8556 < S \leq 417.2446 \\ P_3, \text{ if } S > 417.2446 \end{cases}$$

where P_1, P_2 and P_3 are the three respective classes. For method I, the d_0 and c_2 is taken as 0.6 and 0.4 respectively, and $\hat{c}_1$ which is the corresponding $\widehat{mFPR}$ at d_0 is 0.206488. Altogether, method I results $\widehat{mA}$ and $\widehat{mTpAUC}$ as 0.1873 and 0.0711 respectively. Coming to method II, by using equations (8) and (9) the obtained value for $\hat{d}_0$ and $\hat{c}_2$ are about 0.6948 and 0.2837, $\hat{c}_1$ which is the corresponding $\widehat{mFPR}$ at $\hat{d}_0$ is obtained as 0.2516. By method results the $\widehat{mA}$ and $\widehat{mTpAUC}$ are 0.0381 and 0.0158 respectively. Since, the difference between c_2 and c_1 is too small, the TpAUC portion on the mEROC curve is difficult to depict. However, the TpAUC is shown for breast cancer data (Figure 8).



(a) Mice data **(b) Breast Cancer data**

Figure 6: Density plots of real datasets

Data 2: Breast Cancer data

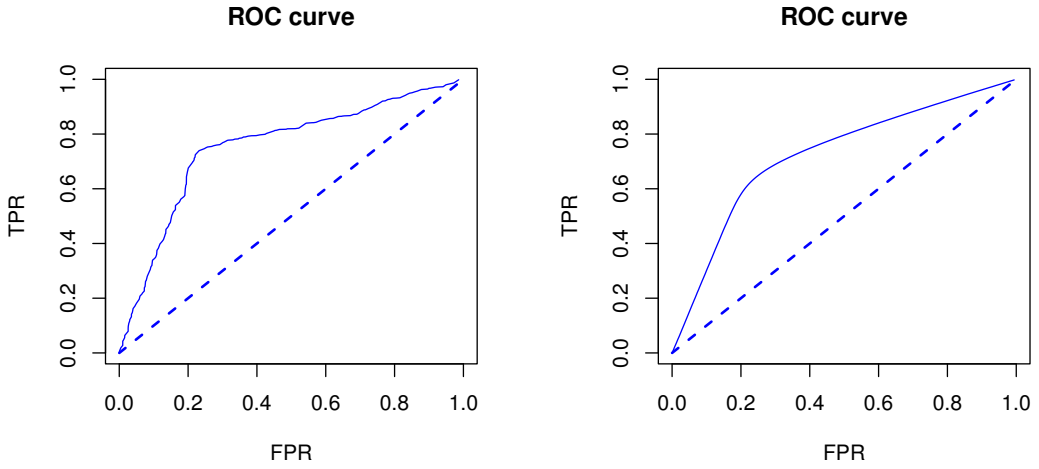
The real dataset represent the survival times of 121 patients with *breast cancer* obtained from a large hospital in a period from 1929 to 1938 (Lee and Wang, 2003). The p-value of K-S test for exponential distribution is 0.06024 (test statistics (D)=0.12031) which indicates that the data follows exponential distribution. The estimation is done by using both the methods and respective results are shown in Tables 4 and 5.

The density plot of the breast cancer data is given in Figure 6 (b), and is very clear that there exists multi-modality which indicates the presence of sub-populations. By using EM algorithm we identified that there are three classes and the estimated the parameters of the respective populations are $\hat{\theta}_0 = 0.4010$, $\hat{\theta}_1 = 0.0280$ and $\hat{\theta}_2 = 0.0202$. The optimal thresholds, $\hat{t}_1$ and $\hat{t}_2$ are 7.7311 and 44.6911, which gives accuracy about 75.26% with false positive rates about 20.63% and true positives of 63.93%. This means to that a subject can be classified in the following manner

$$\text{It is classified as} = \begin{cases} P_1, \text{ if } S \leq 7.7311 \\ P_2, \text{ if } 7.7311 < S \leq 44.6911 \\ P_3, \text{ if } S > 44.6911 \end{cases}$$

where P_1, P_2 and P_3 are the three respective classes with low, medium and high survival rate respectively.

For method I, the $\hat{d}_0$ and $\hat{c}_2$ are taken as 0.6 and 0.5 and the $\hat{c}_1$ corresponding to $\hat{d}_0$ is obtained as 0.1986, which results $\widehat{mA}$ and $\widehat{mTpAUC}$ about 0.1935 and 0.01266. By method II the values obtained for $\hat{d}_0$ and $\hat{c}_2$ are about 0.6123 and 0.2837, the corresponding $\hat{c}_1$ to $\hat{d}_0$ is 0.2142, altogether results $\widehat{mA}$ and $\widehat{mTpAUC}$ of 0.0509 and 0.0084 respectively. The TpAUC for breast cancer data is depicted in Figure 8.



(a) Mice data **(b) Breast Cancer data**

Figure 7: ROC curves for real datasets

Table 4: ROC curve measures of real datasets

$\hat{\pi}_1$	$\hat{\pi}_2$	$\hat{t}_1$	$\hat{t}_2$	$\widehat{mFPR}$	$\widehat{mTPR}$	$Var(\widehat{mFPR})$	$Var(\widehat{mTPR})$	$\hat{J}$	$\widehat{mAUC}$
Mice data									
0.4998	0.5002	12.8556	417.2446	0.2435	0.6973	0.000333	0.000308	0.45378	0.726739
Breast Cancer data									
0.4999	0.5001	7.7311	44.6911	0.20627	0.6393	0.000371	0.000286	0.432985	0.752647

Table 5: Partial area estimates of real datasets

Dataset	Method	$\hat{d}_0$	$\hat{c}_1$	$\hat{c}_2$	$\widehat{mA}_{(c_1,c_2)}$	$\widehat{mTpAUC}_{(c_1,c_2,d_0)}$
Mice	I	0.6	0.206488	0.4	0.187254	0.071147
	II	0.69483	0.251559	0.283712	0.038135	0.015794
Breast Cancer	I	0.6	0.19857	0.5	0.193517	0.012659
	II	0.61229	0.214157	0.283665	0.050896	0.008337

4. Summary

In this work, we made an attempt to explain the need and importance of analyzing a portion of the ROC curve for multi-class non-normal data. Methodological descriptions are given in detail for one-way and two-way pAUC. Expressions for mpAUC and mTpAUC are also derived, and the terms d_0 and c_2 involved in these expressions are obtained using asymptotic confidence intervals of mFPR and mTPR. Two real datasets and considerable simulations are used to demonstrate the proposed work. From the results it is observed that for a multi-class ROC curve, whose area is maximum (minimum), the areas within the mFPR range and mTpAUC will also have a larger (smaller) portion in the entire area.

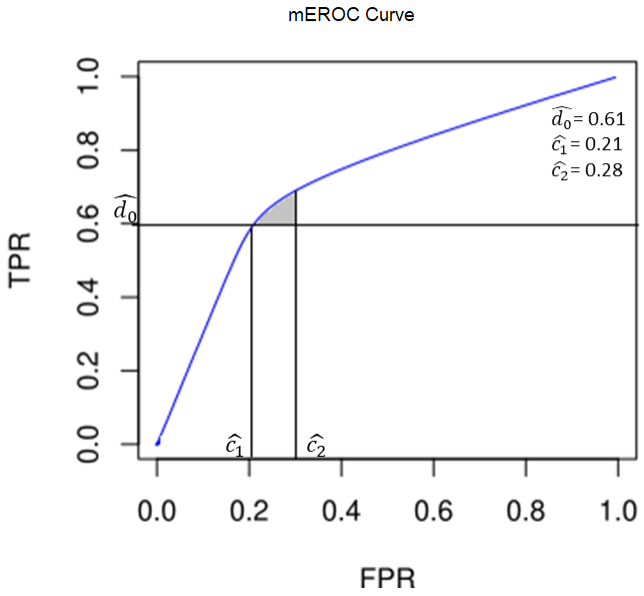


Figure 8: Two way pAUC for Breast Cancer data

Acknowledgements

The authors indeed grateful to the Editors for their guidance and counsel. The authors are grateful to the reviewer for valuable comments and suggestions of generously listing many useful references.

References

Arunima, S. K. and Vishnu, R. V. (2022). Estimation of area under the roc curve in the framework of gamma mixtures. *Communications in Statistics: Case Studies, Data Analysis and Applications*, **8**, 1–14.

Arunima, S. K. and Vishnu, R. V. (2023). Estimation of area under the multi-class roc for non-normal data. *Statistics and Applications*, **21**, 113–121.

Baker, S. G. and Pinsky, P. F. (2001). A proposed design and analysis for comparing digital and analog mammography: special receiver operating characteristic methods for cancer screening. *Journal of the American Statistical Association*, **96**, 421–428.

Cheam, A. S. and McNicholas, P. D. (2016). Modelling receiver operating characteristic curves using gaussian mixtures. *Computational Statistics and Data Analysis*, **93**, 192–208.

Elandt-Johnson, R. C. and Johnson, N. L. (1980). *Survival models and data analysis*. John Wiley and Sons.

Gönen, M. (2013). Mixtures of receiver operating characteristic curves. *Academic Radiology*, **20**, 831–837.

Hillis, S. L. and Metz, C. E. (2012). An analytic expression for the binormal partial area under the roc curve. *Academic Radiology*, **19**, 1491–1498.

Jiang, Y., Metz, C. E., and Nishikawa, R. M. (1996). A receiver operating characteristic partial area index for highly sensitive diagnostic tests. *Radiology*, **201**, 745–750.

Lee, E. T. and Wang, J. (2003). *Statistical Methods for Survival Data Analysis*, volume 476. John Wiley and Sons.

Lusted, L. B. (1971). Signal detectability and medical decision-making: Signal detectability studies help radiologists evaluate equipment systems and performance of assistants. *Science*, **171**, 1217–1219.

McClish, D. K. (1989). Analyzing a portion of the roc curve. *Medical Decision Making*, **9**, 190–195.

Obuchowski, N. A. (2003). Receiver operating characteristic curves and their use in radiology. *Radiology*, **229**, 3–8.

Obuchowski, N. A. and Bullen, J. A. (2018). Receiver operating characteristic (roc) curves: review of methods with applications in diagnostic medicine. *Physics in Medicine and Biology*, **63**, 1–17.

Peterson, W., Birdsall, T., and Fox, W. (1954). The theory of signal detectability. *Transactions of the IRE Professional Group on Information Theory*, **4**, 171–212.

Ricamato, M. T. and Tortorella, F. (2011). Partial auc maximization in a linear combination of dichotomizers. *Pattern Recognition*, **44**, 2669–2677.

Siva, G. and Vishnu, V. R. (2022). Multi-class classification using mixtures of univariate and multivariate roc curves. *Journal of Biostatistics and Epidemiology*, **8**, 209–233.

Tanner, J. W. P. and Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review*, **61**, 401.

Thompson, M. L. and Zucchini, W. (1989). On the statistical analysis of roc curves. *Statistics in Medicine*, **8**, 1277–1290.

Zhang, L., Jie, Z., Xu, S., Zhang, L., and Guo, X. (2018). Use of receiver operating characteristic (roc) curve analysis for tyrer-cuzick and gail in breast cancer screening in jiangxi province, china. *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research*, **24**, 5528.

Zhou, X.-H., McClish, D. K., and Obuchowski, N. A. (2009). *Statistical Methods in Diagnostic Medicine*. John Wiley and Sons.