Mapana – Journal of Sciences
2025, Vol. 24, No. 3, *171-186*
ISSN 0975-3303|https://doi.org/10.12723/mjs.74.10

Tri-Class Exponential ROC Space

G. Kumarapandiyan* and Sahana T S*

Abstract

Biomarkers play a vital role in detecting the presence of disease or medical condition of interest. The challenging tasks in clinical diagnosis are to interpret the performance of biomarkers. To evaluate the biomarkers efficiency, the most advantageous tool used is Receiver Operating Characteristic (ROC) Curve. For two class problems (abnormal and normal), various models and alternatives are developed to find the biomarker's performance. In this study, the two class problem has been further extended to the threeclass problem i.e. (normal, suspicious and abnormal). The three-class exponential ROC model, Volume Under the ROC Surface (VUS), asymptotic variance and Confidence Interval (CI) for VUS have been derived. The model has been validated using a simulated data and for the real-life dataset from the underlying distribution.

Keywords: Biomarker, Volume Under the ROC, Asymptotic Variance and Exponential Distribution

1. Introduction

Biomarkers are quantifiable signs of a biological process or condition and they are frequently employed in diagnosis, surveillance, and disease progression or therapy response predictions. It must be evaluated to determine their clinical value in discriminating between non-diseased and diseased subjects. One of the most commonly used methods for determining clinical value is the Receiver Operating Characteristic (ROC) curve.[5,6]

Diversified approaches are available for evaluating a biomarker that classifies an individual into one of two classes. Nevertheless, very few parametric approaches exist to evaluate a biomarker that identifies an individual into three (Non-Diseased, Suspicious, and Diseased) or multiclass. For three-class or multi-class, explicit or non-explicit models for the ROC surface need to be obtained [1, 8]. The ROC surface is a graph representing the simultaneous trade-off between the three correct classification rates, and the surface is traced by plotting the three probabilities of correct classification

* Department of Statistics, Madras Christian College, Chennai, India; kumarapandiyan@mcc.edu.in, Sahanaram2926@gmail.com

[7,9]. Extending to a three class model is particularly important in clinical diagnosis, as it allows for the identification of suspicious cases which often occur in many disease like Alzheimer. This not only improves diagnostic accuracy but also supports more informed decision-making. Volume Under the ROC Surface (VUS) is is a diagnostic accuracy measure used to assess how well a three-class or multi-class classification task is performing. It is sometimes referred to as the area underneath the ROC surface (three dimensional). It is an expansion of the two class classification problem in which Area Under the ROC Curve (AUC), which measures the overall discriminatory capacity of biomarkers in distinguishing diseased and non-diseased subjects, whereas VUS captures how well a biomarker or diagnostic test can separate all three (or more) clinical categories at once. The VUS generalizes the AUC from binary to three class classification. The interpretation of VUS of is same as AUC, that 0.5 means the test performs no better than chance for three classes, while values closer to 1 indicate excellent discriminatory power between the classes. This is especially relevant in clinical settings where patients may fall into more than two status categories, such as "non-diseased," "suspicious," and "diseased. The VUS provides a concise and interpretable summary measure for evaluating multi-class discrimination. Furthermore, the derivation of asymptotic variance for tri-class exponential model has been derived for the first time. This helps in the calculation of Confidence Intervals (CI) which showcase the reliability of the estimate of VUS. Together, these metrics provides the observed discriminatory ability of the selected biomarker in diagnostic accuracy studies.

In two-class classification, we have a non-diseased and diseased group in which classification is made using a gold standard along with their biomarker values. Now, the test procedure works as follows:

Combine the biomarkers of non-diseased and diseased individuals to form a single set, say B_i where i ranges from 1 to n. Now, tag the individuals as non-diseased and diseased using the following decision criteria:

$$\begin{cases} B_i > t & diseased \\ B_i \leq t & non-diseased \end{cases} \quad i = 1, 2, \ldots, n \tag{1}$$

where 't' is the threshold

Now, we have two classification criteria, i.e., one is the gold standard and the other is defined by equation (1). With these two criteria, we obtain the following probabilities:

True Positive Rate (TPR): It is the proportion of diseases detected by the test to the total number of diseased people. It is also called sensitivity.

True Negative Rate (TNR): It is the proportion of non-diseased subjects identified correctly by the test. It is sometimes referred to as specificity.

False Positive Rate (FPR): It is the proportion of non-diseased subjects identified as diseased by the test. It is sometimes referred to as 1-specificity.

False Negative Rate (FNR): It is the proportion of diseased subjects identified as non-diseased by the test. It is also referred to as 1-sensitivity.

Each threshold t leads to a 2×2 contingency table given in Table 1 containing the true and false classification probabilities.

Table 1: 2×2 Confusion matrix regarding the performance of a two-class classification

Actual Status (Gold Standard)	**Result of the classifier (diagnostic marker)**	
	Diseased	**Non-Diseased**
Diseased	$TPR = P\ (Y > t)$	$FNR = P(Y \leq t)$
Non-Diseased	$FPR = P\ (X > t)$	$TNR = P(X \leq t)$

***TPR – True Positive Rate *FNR – False Negative Rate**
***FPR – False Positive Rate *TNR – True Negative Rate**

Now a two-class ROC curve can be generated by plotting FPR on the horizontal axis and TPR on the vertical axis. The primary goal is to reduce errors in classification, specifically focusing on minimizing both FPR and FNR. Sensitivity and specificity of the test refer to the probability values where True Classification (TC) is maximized while False Classification (FC) is minimized. This means achieving a high rate of correctly identifying positives (sensitivity) and negatives (specificity) while minimizing the chances of incorrectly classifying them.

1.1. A parametric approach for a two-class ROC approach:

Let X and Y be two random variables that denote the biomarker values from the non-diseased and diseased populations. Also, let X and Y follow some continuous distribution having $F(x)$ and $G(y)$ as cumulative distribution functions, respectively. Then the form of the ROC model for plotting the ROC curves takes the following form [2,3]:

$$ROC(t) = \bar{G}_Y \circ \bar{F}_X^{-1}(t);\ -\infty < t < \infty \quad (2)$$

The summary index of the ROC curve is the AUC, which is the total area covered under the ROC curve towards the x-axis. AUC is defined as the probability that a randomly chosen marker value from a diseased population will have higher values than a randomly chosen marker value from a non-diseased population [3], i.e., as

$$AUC = P(Y > X) = \int_0^1 ROC(t)\ dt = \int_0^1 \bar{G}_Y \circ \bar{F}_X^{-1}(t)\ dt \tag{3}$$

The ROC curve ROC(t) is directly related to the AUC, which is the integral of the ROC curve over all possible thresholds from 0 to 1. The AUC provides a single scalar value that summarizes the overall performance of the classifier in terms of its ability to discriminate between positive and negative instances.

1.2. A parametric approach of ROC surface to three-class diagnostic problem:

A three-class ROC surface can be used in medical diagnosis to assess the effectiveness of a diagnostic test or biomarker that divides the subject into three groups: Non-Diseased, Suspicious, or Diseased, using two thresholds t_1 and t_2. A three-dimensional extension of the AUC that offers a more thorough evaluation of the effectiveness of the biomarker is the VUS. Here the ROC surface is fitted using the true classification rates i.e., (TC_1, TC_2, and TC_3). [7]

Let X, Y and Z be the random variables representing the biomarker values of Non-Diseased, Suspicious, and Diseased groups respectively. These variables assume continuous distribution function, i.e., $X_i \sim F_X(.); i = 1, 2, \ldots, m$ *(For non – diseased)*,, $Y_j \sim F_Y(.); j = 1,2,\ldots, n$ *(For suspicious)* and $Z_k \sim F_Z(.); k = 1,2,..., p$ *(For diseased)*. As in the two-class approach, combine the biomarker values of all the three groups together to form a single set, say B_i $(i = 1,2,\ldots,N)$; where $N = m + n + p$; m, n and p be the number of observations from Non-Diseased, Suspicious, and Diseased groups, respectively. For each set of two ordered thresholds t_1 and t_2 ($t_1 < t_2$), assigns a probability score for each class and they are defined as follows:

$P(B_i \le t_1)$:	Estimate of the probability that the data point B_i belongs to diseased
$P(t_i < B_i \le t_2)$:	Estimate of the probability that the data point B_i belongs to a suspicious group
$P(B_i > t_2)$:	Estimate of the probability that the data point B_i belongs to a non-diseased group

Whenever a gold standard technique is available, then the confusion matrix can be described in Table 2:

Table 2: 3×3 Confusion matrix regarding the performance of a three-class classification

Results from the Gold Standard	**Results from the Diagnostic test**		
	Non-diseased	**Suspicious**	**Diseased**
Non-diseased	$TC_1 = P(X \le t_1)$	$FC_1 = P(t_1 < X \le t_2)$	$FC_2 = P(X > t_2)$
Suspicious	$FC_3 = P(Y \le t_2)$	$TC_2 = P(t_1 < Y \le t_2)$	$FC_4 = P(Y > t_2)$
Diseased	$FC_5 = P(Z \le t_1)$	$FC_6 = P(t_1 < Z \le t_2)$	$TC_3 = P(Z > t_2)$

***TC – True Classification**
***FC- False Classification**

If TC and FC are the correct and false classifications, respectively, then TC_1 is the probability of correctly identifying non-diseased individuals, TC_2 is the probability of correctly identifying suspicious individuals, and TC_3 is the probability of correctly identifying diseased individuals. The remaining six misclassification that occur during the classification procedure using t_1 and t_2 are: FC_1 (non-diseased subjects are misidentified as suspicious), FC_2 (non-diseased subjects are identified as diseased), FC_3 (suspicious are incorrectly identified as non-diseased), FC_4 (suspicious are misidentified as diseased), FC_5 (Diseased are misidentified as non-diseased), and FC_6 (diseased are misidentified as suspicious). The general ROC surface model is given by

$$ROC\ (t_1, t_2) = F_Y(F_Z^{-1}(t_2)) - F_Y(F_Z^{-1}(t_1));\ -\infty < t_1, t_2 < \infty \quad (4)$$

where,

$$P(Z \le t_1) = F_Z\ (t_1) \quad (5)$$

$$P\ (t_1 < Y \le t_2) = F_Y(t_2) - F_Y(t_1) \quad (6)$$

$$\mathrm{P}\ (\mathrm{X} > t_2) = 1 - F_X(t_2) \quad (7)$$

A typical ROC surface plot will look like the one presented in Figure 1. In this plot, the x-axis represents TC_1, the y-axis represents TC_2, and the z-axis represents TC_3.

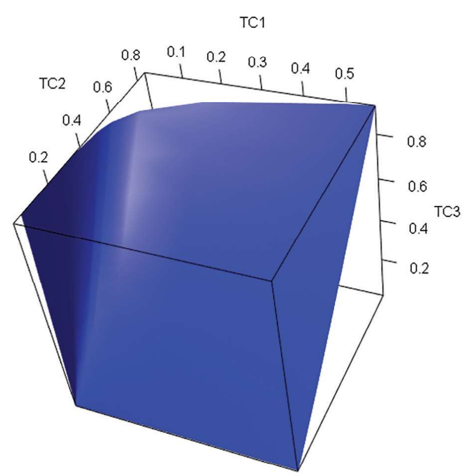

Figure 1: The general ROC plot

1.3. The Volume Under the ROC Surface

The VUS offers a summary index of biomarkers in a three-class approach. The non-parametric form of VUS can be written as

$$P\ (Z > Y\ >\ X) = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{p}\frac{I(X_i,\ Y_j, Z_k)}{mnp} \tag{8}$$

where

$$I(X_i, Y_j, Z_k) = \begin{cases} 1 & if\ Z < Y < X \\ 0.5 & if\ Z = Y < X \\ 0.5 & if\ \mathrm{Z}\ <\ \mathrm{Y}\ = \mathrm{X} \\ 0 & if\ \mathrm{Z} >\ \mathrm{Y}\ >\ \mathrm{X} \\ \dfrac{1}{6} & if\ (Z = Y = X) \end{cases}$$

The VUS ranges from 0 to 1. A VUS value of 1 indicates a perfect model that achieves perfect classification for all classes. It is computationally difficult to obtain for a larger number of classes.

2. Materials and Methods

2.1. Tri-class exponential ROC (Tri-EROC) model

Let us assume that biomarker values from the three groups follows exponential distribution individually, i.e., $X \sim Exp\ (\theta_x)$, $Y \sim Exp\ (\theta_y)$ and $Z \sim Exp\ (\theta_z)$ where *X*, *Y* and *Z* are mutually independent.

where

$$F_X\ (x) =\ 1 - e^{-\theta_x};\ 0 \le \theta_x < \infty$$

$$F_Y(y) =\ 1 - e^{-\theta_y};\ 0 \le \theta_y < \infty$$

$$F_Z(z) = 1 - e^{-\theta_z};\ 0 \le \theta_z < \infty$$

$$VUS\ =\ P(Z > Y > X)\ = \int_0^\infty \int_x^\infty \int_y^\infty f(x)f(y)f(z)\ dz\ dy\ dx \tag{9}$$

$$= \int_0^\infty \int_x^\infty \int_y^\infty \theta_x\ e^{-\theta_x}\ \theta_y e^{-\theta_y}\ \theta_z e^{-\theta_z}\ dz\ dy\ dx \tag{10}$$

The volume under the ROC surface is calculated by solving the integral of equation (10) and we get,

$$VUS = \frac{\theta_x \theta_y}{(\theta_y + \theta_z) + (\theta_x + \theta_y + \theta_z)} \tag{11}$$

2.2. Asymptotic Variance of VUS of Tri-EROC

Asymptotic variance quantifies how much the VUS estimate would fluctuate if you repeatedly sampled from the population. It's essential for assessing the reliability and precision of the VUS estimate. Smaller variance means more confidence in the summary measure.

The log-likelihood function of the joint pdf of the three population is given by

$$L = \prod_{i=1}^{m} \theta_x e^{-\theta x_i} \prod_{j=1}^{n} \theta_y e^{-\theta y_j} \prod_{k=1}^{p} \theta_z e^{-\theta z_k} \tag{12}$$

Then the log likelihood function (ln L) of (12) is given by

$$Log\ L = \mathrm{m}\ell\, \theta_x - \theta_x \sum_{i=1}^{m} x_i + \mathrm{n}\ell\, \theta_y - \theta_y \sum_{j=1}^{n} y_j + \mathrm{k}\ell\, \theta_z - \theta_z \sum_{k=1}^{p} z_k \tag{13}$$

The Fisher Information matrix helps determine the variance-covariance structure of this limiting distribution. The consistent property of Maximum Likelihood Estimator (MLE) has been used. A parametric approach utilizing MLE was also proposed and executed using Mathematica [4]. It shows that $\delta = (\theta_x, \theta_y, \theta_z) \sim N[0, I^{-1}(\delta)]$ and $I(\delta)$ is a Fisher information matrix which is given by

$$I(\delta) = \begin{bmatrix} E\left(\frac{\partial^2 \ell}{\partial \theta_z^2}\right) & E\left(\frac{\partial^2 \ell}{\partial \theta_z^2 \partial \theta_y^2}\right) & E\left(\frac{\partial^2 \ell}{\partial \theta_z^2 \partial \theta_x^2}\right) \\ . & E\left(\frac{\partial^2 \ell}{\partial \theta_y^2}\right) & E\left(\frac{\partial^2 \ell}{\partial \theta_x^2 \partial \theta_y^2}\right) \\ . & . & E\left(\frac{\partial^2 \ell}{\partial \theta_x^2}\right) \end{bmatrix}$$

Now taking the first derivative of ln L with respect to θ_x, θ_y and θ_z

$$\frac{\partial \ln L}{\partial \theta_x} = \frac{m}{\theta_x} - \Sigma x_i, \frac{\partial \ln L}{\partial \theta_y} = \frac{n}{\theta_y} - \Sigma y_j, \frac{\partial \ln L}{\partial \theta_z} = \frac{p}{\theta_z} - \Sigma z_i \tag{14}$$

In order to get the I(δ) the second derivative for the parameters have been taken

$$\frac{\partial^2 L}{\partial \theta_x} = \frac{-m}{\theta_x^2}, \frac{\partial^2 L}{\partial \theta_y} = \frac{-n}{\theta_y^2}, \frac{\partial^2 L}{\partial \theta_z} = \frac{-p}{\theta_z^2} \tag{15}$$

By using the delta method, the estimated asymptotic variance becomes

$$V\left(\widehat{VUS}\right) = 2\,\frac{\theta_y^2\theta_z^2}{\theta_x^2} m^2\,\theta_y\left(\theta_y + \theta_x\right) - \left(\frac{\theta_x^2\theta_z^2}{\theta_y^2}\right)\left(n^2\,\theta_x\left(\theta_x\,\theta_z - \theta_y^2 + \theta_z^2\right) + \left(\frac{\theta_x^{\,2}\theta_y^2}{\theta_z^2}\right)\frac{(p^2\,\theta_x\,\theta_y(\theta_x + 2\theta_y + 2\,\theta_z)}{(\theta_x + \theta_y + \theta_z)2(\theta_y + \theta_z)} \right. \tag{16}$$

3. Results and Discussion

This section uses simulation studies to validate Tri-EROC and Tri-EROC's VUS. The three-class exponential distribution ROC model has been researched through simulated studies. The exponential distribution was used to generate the sample of same sizes (30,30,30) for with varying parametric values. $(\theta_z^2,\ \theta_y^2,\ \theta_x^2)$={(80,65,3), (80,35,3), (80,15,2), (90,20,1)}. The simulation studies have been done and the values are elaborated in Table 3. Simulated data was generated to evaluate the performance of the proposed method under controlled conditions. The data were simulated under the assumption that biomarker values follow an exponential distribution, which is widely applied in survival and reliability analysis and serves as a reasonable model for certain biomedical processes. Using simulated datasets allows for systematic assessment of the VUS estimator across different scenarios and ensures that the results are not overly dependent on a single clinical dataset. Additionally, the exponential distribution's density plot is also fitted for all the parameters in figure 2, 3, 4 and 5.

Table 3: Simulation Studies that present ROC surface for different parametric values

SET	Parameters	Estimated means	VUS parametric and CI	VUS Non-Parametric and CI	Standard Error
1	$\theta_z = 80$ $\theta_y = 65$ $\theta_x = 3$	$\hat{\theta}_z = 79.9$ $\hat{\theta}_y = 42.9$ $\hat{\theta}_x = 2.9$	0.595 [0.294,0.869]	0.619 [0.318,0.920]	0.15352
2	$\theta_z = 80$ $\theta_y = 35$ $\theta_x = 3$	$\hat{\theta}_z = 79.9$ $\hat{\theta}_y = 23.1$ $\hat{\theta}_x = 2.9$	0.6701 [0.626,0.714]	0.666 [0.622,0.710]	0.022287
3	$\theta_z = 80$ $\theta_y = 15$ $\theta_x = 2$	$\hat{\theta}_z = 79.9$ $\hat{\theta}_y = 9.9$ $\hat{\theta}_x = 1.9$	0.729 [0.7284,0.7296]	0.715 [0.7144,0.7156]	0.00031
4	$\theta_z = 90$ $\theta_y = 25$ $\theta_x = 1$	$\hat{\theta}_z = 89.8$ $\hat{\theta}_y = 6.6$ $\hat{\theta}_x = 0.9$	0.805 [0.8049,0.8051]	0.814 [0.8139,0.8141]	0.000056

***ROC – Receivers Operating Characteristics Curve**
***VUS – Volume Under the ROC Surface**

Table 3 presents various sets of assumed parameters and their corresponding evaluation results. Each set includes specific parameter values used in the simulation. The table provides estimated means for these parameters, calculated using the MLE. In addition to the estimated means, the table includes parametric estimates of the VUS. These parametric VUS values quantify the overall performance of a classifier for each set of parameters. Alongside the parametric VUS estimates, the table also provides non-parametric VUS estimates, which serve as an alternative measure that does not rely on the assumptions inherent to parametric methods.

The standard error values listed in the table reflect the precision of the VUS estimates. Lower standard error values indicate more reliable and precise VUS estimates. Together, these metrics offer a comprehensive assessment of classifier performance across different parameter sets, highlighting both the parametric and non-parametric perspectives. The set 4 parameters is the best performer because it has the highest VUS value and the lowest standard error, indicating both superior classifier performance and high precision in the estimates.

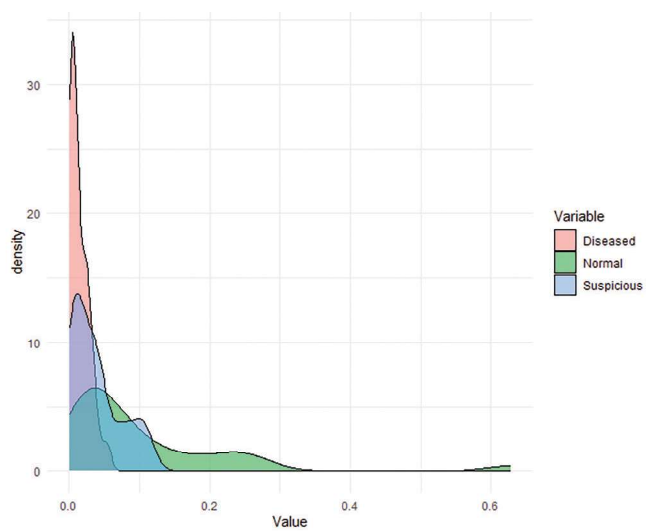

Figure 2: Density Plot for the Parameter (80,65,3)

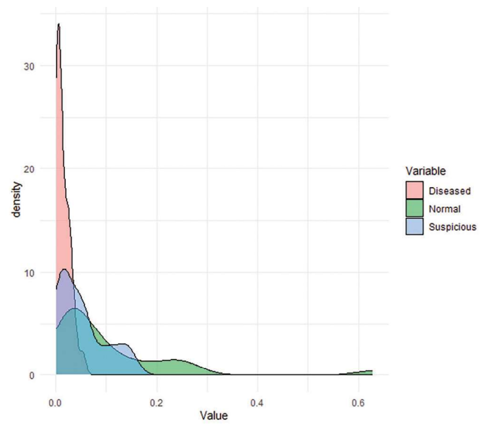

Figure 3: Density Plot for the parameter (80,35,3)

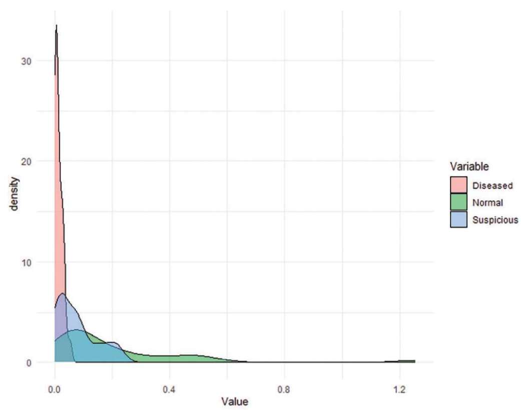

Figure 4: Density plot for the parameter (80,15,2)

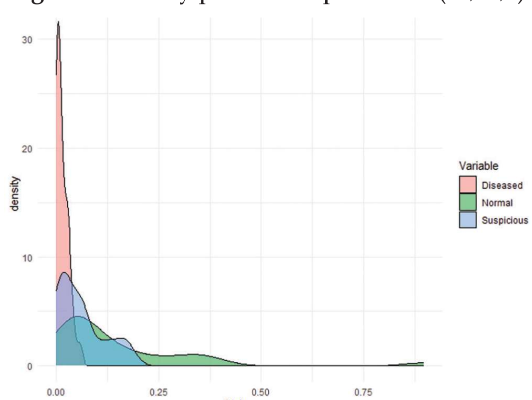

Figure 5: Density plot for the parameter (90,20,1)

Each parameter (1/θ) represents the rate at which Non-Diseased, Suspicious, and Diseased occur. The higher the 1/θ, the more frequent the events. Conversely, a lower 1/θ means events occur less frequently. In the above graph, it is evident that higher a parameter value steeper the decline in the density plot from 0 indicating more of diseased group, while lower parameter leads to a shallower decline, indicating less of non-diseased group and moderate of suspicious group.

The parenthesis in the VUS column represents the non-parametric estimate of VUS for the same set. Both parametric and non-parametric versions of the three-layered VUS are displayed from figure 6 to 13.

a. ROC curve of parametric approach

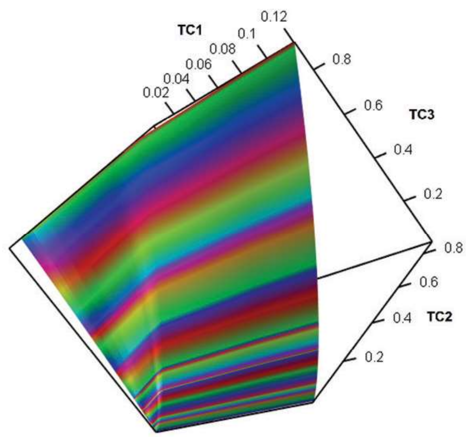

Figure 6: ROC surface of Tri-EROC for (θ_z = 80, θ_y = 65, θ_x = 3)

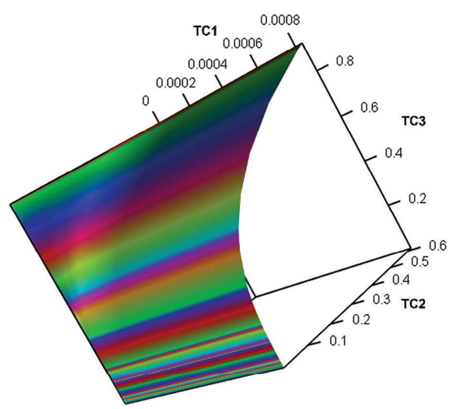

Figure 7: ROC surface of Tri-EROC for (θ_z = 80, θ_y = 35, θ_x = 3)

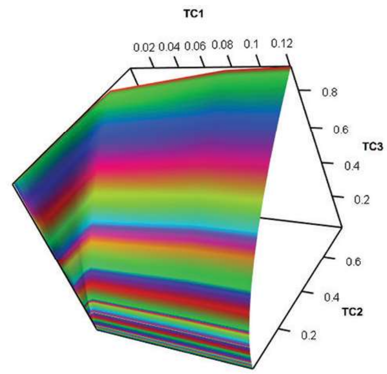

Figure 8: ROC surface of Tri-EROC for (θ_z = 80, θ_y = 15, θ_x = 2)

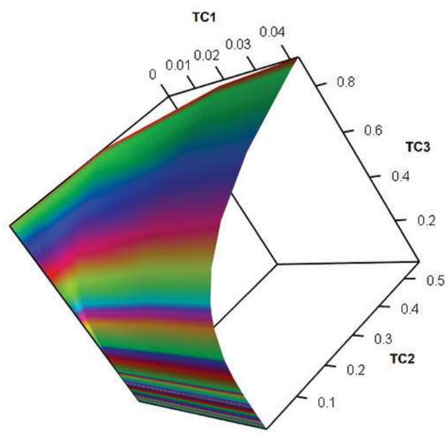

Figure 9: ROC surface of Tri-EROC for (θ_z = 90, θ_y = 20, θ_x = 1)

b. Non-parametric ROC approach

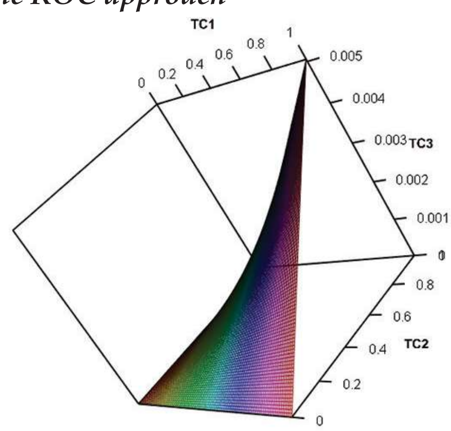

Figure 10: ROC surface of Tri-EROC for (θ_z = 80, θ_y = 65, θ_x = 3)

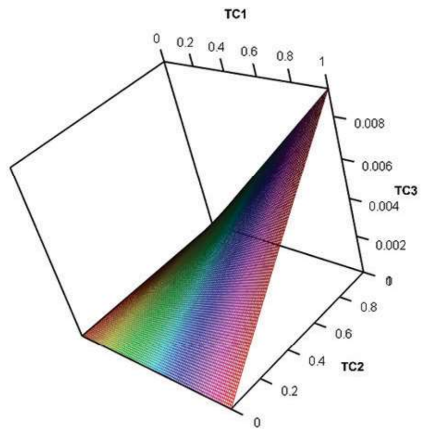

Figure 11: ROC surface of Tri-EROC for (θ_z = 80, θ_y = 35, θ_x = 3)

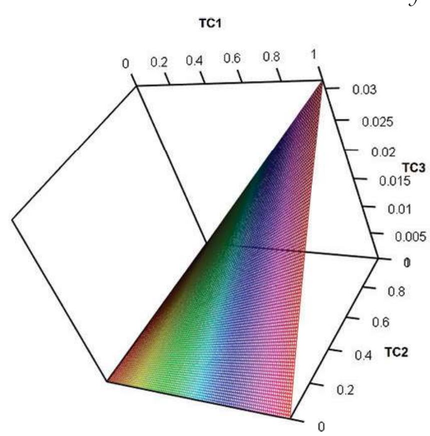

Figure 12: ROC surface of Tri-EROC for (θ_z = 80, θ_y = 15, θ_x = 2)

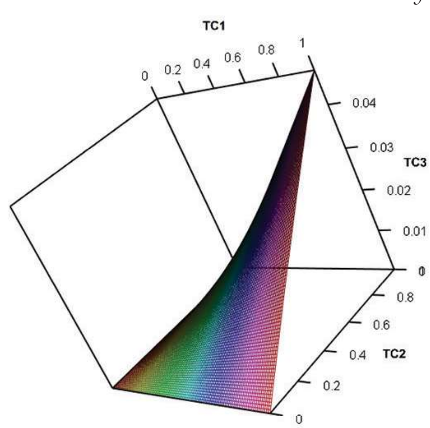

Figure 13: ROC surface of Tri-EROC for (θ_z = 90, θ_y = 20, θ_x = 1)

A larger volume under the ROC curve typically indicates better model performance. It means that the model is performing well across different threshold settings and possibly across different values of the third dimension. ROC curves are useful when you want to visualize and analyze how model performance varies not only with different threshold settings but also with other factors. From the above figures, the volume under the curve increase according to the parameter which means when the difference in the parameter increase the model fits better.

A real-life clinical dataset which was downloaded from kaggle has been used further to evaluate the VUS. This dataset pertains to the identification of early stage cardiovascular diseases and was used to validate the Tri-class exponential ROC space model. The dataset contains patient diagnostic information classified into three categories: *Non-Diseased*, *Suspicious*, and *Diseased*. Incorporating this dataset enables the practical evaluation of the proposed model, ensuring that the VUS estimation is both theoretically rigorous and clinically relevant. This validation enhances the model's credibility and demonstrates its potential applicability in biomarker assessment for diagnostic accuracy studies involving continuous biomarkers. For this particular dataset, the VUS, which measures the effectiveness of the classifier in distinguishing between these groups, is 0.5963 and the SE is 0.00263. The corresponding graphical representation of this analysis is shown in Figure 14.

A VUS value of 0.5963 suggests that the classifier has a moderate capability to differentiate between the Non-Diseased, Suspicious, and Diseased groups in the context of early-stage cardiovascular disease identification. This value indicates a fair level of accuracy in identifying the various stages of cardiovascular disease within the dataset.

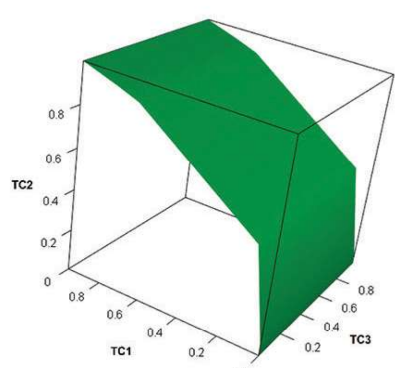

Figure 14: The ROC graph for the cardiovascular disease

4. Conclusion

We have derived the asymptotic variance and VUS for the exponential distribution in this study. We used varying degrees of parameters to simulate the data for the suggested model, and the ROC model was created for both parametric and non-parametric data with the aid of those parameters. According to the simulation study, the VUS varies with the parameters to varying degrees. The VUS reaches its maximum (0.805) and the standard error reaches its lowest (0.000056) when there is a large discrepancy in the parameters, and vice versa. Furthermore, we found that the parametric ROC model has a stronger influence on the parameters than the non-parametric model. Therefore, we deduce that the VUS achieves its maximum and SE achieves its minimum with a large parameter difference (Non-Diseased, Suspicious, and Diseased). Additionally, the VUS for the real-life dataset is 0.5963, which suggests that the classifier possesses a moderate ability to accurately identify and differentiate between various stages of cardiovascular disease. This value indicates that the classifier performs better than random guessing but is not highly accurate. It can reasonably categorize individuals into Non-Diseased, Suspicious, and Diseased groups based on their risk factors and early signs of cardiovascular disease.

5. Conflicts of Interest

This paper doesn't have any potential conflicts of interest. There are no competing interest to declare.

Reference

1. Mossman, D. (1999). Three-way rocs. Medical Decision Making, 19(1), 78-89. *https://doi.org/10.1177/0272989X9901900110*

2. Krzanowski, W. J., & Hand, D. J. (2009). ROC curves for continuous data. CRC Press. *https://doi.org/10.1201/9781439800225*

3. Pundir, Sudesh, and R. Amala. "Evaluation of biomarker using two parameter bi-exponential ROC curve." Pakistan Journal of Statistics and Operation Research (2015): 481-496 *https://doi.org/10.18187/pjsor.v11i4.992*

4. Heckerling, P. S. (2001). Parametric three-way receiver operating characteristic surface analysis using mathematica. Medical Decision Making, 21(5), 409-417. https://doi.org/10.1177/0272989X0102100507

5. Everson, Richard M., and Jonathan E. Fieldsend. ”Multi-class ROC analysis from a multi-objective optimization perspective.” Pattern Recognition Letters 27.8 (2006): 918-927. *https://doi.org/10.1016/j.patrec.2005.10.016*

6. Yang, H., and Carlin, D. (2000), “ROC Surface: A Generalization of ROC Curve Analysis,” Journal of Biopharmaceutical Statistics, 10, *https://doi.org/10.1081/BIP-100101021*

7. Fieldsend, Jonathan E., and Richard M. Everson. ”Visualisation of multiclass ROC surfaces.” (2005). *http://hdl.handle.net/10871/11691*

8. Nakas, Christos T. “Developments in ROC surface analysis and assessment of diagnostic markers in three-class classification problems.” REVSTAT-Statistical Journal 12.1 (2014): 43-65. *https://doi.org/10.57805/revstat.v12i1.143*

9. Kang, Le, and Lili Tian. ”Estimation of the volume under the ROC surface with three ordinal diagnostic categories.” Computational Statistics Data Analysis 62 (2013): 39-51. *https://doi.org/10.1016/j.csda.2013.01.004*