Hindawi Publishing Corporation
Journal of Probability and Statistics
Volume 2015, Article ID 934362, 8 pages
http://dx.doi.org/10.1155/2015/934362

*Research Article*

# Confidence Interval Estimation of an ROC Curve: An Application of Generalized Half Normal and Weibull Distributions

**S. Balaswamy and R. Vishnu Vardhan**

*Department of Statistics, Pondicherry University, Pondicherry 605014, India*

Correspondence should be addressed to R. Vishnu Vardhan; rvvcrr@gmail.com

Received 25 June 2015; Revised 22 September 2015; Accepted 7 October 2015

Academic Editor: Shesh N. Rai

Copyright © 2015 S. Balaswamy and R. Vishnu Vardhan. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the recent past, the work in the area of ROC analysis gained attention in explaining the accuracy of a test and identification of the optimal threshold. Such types of ROC models are referred to as bidistributional ROC models, for example Binormal, Bi-Exponential, Bi-Logistic and so forth. However, in practical situations, we come across data which are skewed in nature with extended tails. Then to address this issue, the accuracy of a test is to be explained by involving the scale and shape parameters. Hence, the present paper focuses on proposing an ROC model which takes into account two generalized distributions which helps in explaining the accuracy of a test. Further, confidence intervals are constructed for the proposed curve; that is, coordinates of the curve (FPR, TPR) and accuracy measure, Area Under the Curve (AUC), which helps in explaining the variability of the curve and provides the sensitivity at a particular value of specificity and vice versa. The proposed methodology is supported by a real data set and simulation studies.

## 1. Introduction

In classification analysis, the Receiver Operating Characteristic (ROC) curve is a widely used tool to evaluate the performance of a test. Further, the intrinsic measures such as sensitivity, specificity, and accuracy are essential to describe a diagnostic test's ability to classify an individual into one of the two groups/populations. Sensitivity provides an estimate of how good the test is at predicting a disease. Specificity estimates how likely patients without disease can be correctly identified. ROC curve is a graphical representation of 1 − specificity and sensitivity. That is, the points of the curve are obtained by moving the classification threshold from the most positive classification value to the most negative. For a random classification, the ROC curve is a straight line connecting the origin $(0, 0)$ to top right corner of the graph $(1, 1)$. Further, the accuracy measure is defined as the area under the ROC curve. Therefore, the criterion widely used to measure the accuracy of a test in ROC context is the area under an ROC curve (AUC).

In classification, the main aim is to discriminate between normal and abnormal populations with better accuracy. In the literature so far many ROC models exist based on bidistributional assumptions such as binormal (Egan [1]), bilogistic and bilognormal (Dorfman and Alf Jr. [2, 3]), bibeta and biexponential (Zou et al. [4]; Tang et al. [5]; Tang and Balakrishnan [6]), and bigamma etcetera (Hussain [7]). If the test scores of normal and abnormal populations follow different distributions, then these ROC forms will not produce reliable outputs. For instance, consider that a marker, namely, APACHE (Acute Physiology and Chronic Health Evaluation) II, is used to predict the mortality status of patients who gets admitted into ICU. The pattern of APACHE scores for live and dead patient's does not possess the normality and explains skewed nature of the data. Here, the conventional binormal ROC model will fail to produce reliable outputs in terms of AUC, threshold, sensitivity, and specificity. However, the distribution of scores may follow any skewed distributions. Hence, the main concentration of the paper lies in handling the situations when distributions of two populations

are different and the data skewed nature of the data. We propose an ROC model that takes into account Generalized Half Normal (normal population) and Weibull (abnormal population) distribution with shape and scale parameters. In medical, engineering, and life studies, data tend to have extended tails; in this situation, the conventional binormal ROC curve fails to explain the hidden accuracy of the test considered. Recently, Balaswamy et al. [8] addressed this issue and developed a Hybrid ROC (HROC) curve which is based on Half Normal and Exponential distributions. However, this model is restricted by considering only scale parameters to illustrate the accuracy. But there are other statistical measures which accounts the information about the tail property of the data. In this paper, an extended version of the HROC curve is proposed by considering the Generalized Half Normal and Weibull distributions with both scale and shape parameters corresponding to normal as well as abnormal populations. A bootstrap study is used to construct the 95% confidence intervals and other measures of the proposed ROC curve. Further, the proposed methodology is demonstrated using simulation studies as well as a real data set.

The present paper is organized as follows. The ROC curve is developed based on Generalized Half Normal (GHN) and Weibull distributions with scale ($\sigma$) and shape ($\alpha$) parameters of both functions and GHROC curve accuracy measure, Area Under the Curve, is derived. Further, the confidence intervals for AUC and proposed ROC curve are estimated through bootstrap method. Finally, the results obtained using proposed methodology are illustrated in Results and Discussion.

## 2. Methodology

Let $\{x_1, x_2\} \in S$ be the test scores, which are observed in normal ($H$) and abnormal ($D$) populations, respectively. Here, it is assumed that $H$ and $D$ populations follow Generalized Half Normal (GHN) and Weibull distributions with shape and scale parameters as $\alpha > 0$ and $\sigma > 0$, respectively. The probability density function and cumulative distribution function of GHN (Cooray and Ananda [9]) and Weibull distributions are given as follows:

$$f(t) = \sqrt{\frac{2}{\pi}} \frac{\alpha}{t} \left(\frac{t}{\sigma}\right)^{\alpha} e^{(-1/2)(t/\sigma)^{2\alpha}},$$
$$F(t) = 2\left[\Phi\left(\frac{t}{\sigma}\right)^{\alpha}\right] - 1, \tag{1}$$

where $\Phi(\cdot)$ is the c.d.f. of the standard normal distribution:

$$g(t) = \frac{\alpha}{t} \left(\frac{t}{\sigma}\right)^{\alpha-1} e^{-(t/\sigma)^{\alpha}}; \quad t > 0,$$
$$G(t) = 1 - e^{-(t/\sigma)^{\alpha}}. \tag{2}$$

In classification, ROC curve is a graphical plot that illustrates the performance of a binary classifier as its discrimination threshold varies (Green and Swets [10]). The curve is obtained by plotting the false positive rate (FPR) against the true positive rate (TPR).

The expression for FPR is derived by using its probabilistic definition as

$$x(t) = P(S > t \mid H) = 2\left\{1 - \Phi\left[\left(\frac{t}{\sigma_H}\right)^{\alpha_H}\right]\right\}; \tag{3}$$

on further simplification, the expression for $t$ can be obtained by the formula

$$t = \sigma_H \left[\Phi^{-1}\left(1 - \frac{x(t)}{2}\right)\right]^{1/\alpha_H}, \tag{4}$$

where $\Phi^{-1}(\cdot)$ is the inverse cumulative standard normal distribution function.

Similarly, the expression for TPR is derived by using its probabilistic definition from Weibull distribution as

$$y(t) = P(S > t \mid D) = 1 - \left[1 - e^{-(t/\sigma_D)^{\alpha_D}}\right], \tag{5}$$

$$y(t) = e^{-(t/\sigma_D)^{\alpha_D}}; \tag{6}$$

on substituting (4) into (6), the expression for TPR can be written as

$$y(t) = e^{-\beta^{\alpha_D}[\Phi^{-1}(1-x(t)/2)]^{\varsigma}}; \tag{7}$$

here $\beta = \sigma_H/\sigma_D$, $\varsigma = \alpha_D/\alpha_H$, and (7) is the expression of ROC Curve based on Generalized Half Normal and Weibull distributions. This expression (7) can be referred to as the *Generalized Hybrid ROC (GHROC) curve*, since the ROC curve is developed based on two generalized distributions.

In ROC methodology, the statistical measure which helps in explaining the overlapping area and the accuracy of a classifier is the Area Under the Curve (AUC). It can be interpreted as the probability that a subject randomly selected from the group with the condition will have a discriminating score indicating greater likelihood than that of a randomly selected subject from the group without condition (Bamber [11]). The AUC can take values between 0 and 1 with practical lower bound value of 0.5 (chance diagonal). The expression for the accuracy measure AUC can be obtained by integrating the ROC expression (7) over the range [0, 1] with respect to the false positive rate as

$$\text{AUC} = \int_0^1 e^{-\beta^{\alpha_D}[\Phi^{-1}(1-x(t)/2)]^{\varsigma}} dx(t). \tag{8}$$

The above expression has no closed form; hence it has to be solved using numerical integration. In the next subsection, the variance and confidence intervals for AUC are estimated through bootstrapping method.

*2.1. Confidence Intervals for AUC.* The $100(1-\alpha)\%$ confidence interval for AUC can be defined as

$$\widehat{\text{AUC}} \pm Z_{1-\alpha/2} \sqrt{\text{Var}\left(\widehat{\text{AUC}}\right)}, \tag{9}$$

where $Z_{1-\alpha/2}$ is the $1-\alpha/2$ standard normal percentile and $\mathrm{Var}(\widehat{\mathrm{AUC}})$ is the estimated variance of $\widehat{\mathrm{AUC}}$, which is obtained using bootstrapping. Let "$B$" be the number of bootstraps obtained from the data with the sample sizes $n_H$ and $n_D$, respectively, from normal and abnormal populations. Then the bootstrapped AUC estimate and its variance are

$$\widehat{\mathrm{AUC}} = \frac{1}{B}\sum_{b=1}^{B}\mathrm{AUC}_b,$$
$$\mathrm{Var}\left(\widehat{\mathrm{AUC}}\right) = \frac{1}{B-1}\sum_{b=1}^{B}\left(\mathrm{AUC}_b - \widehat{\mathrm{AUC}}\right)^2, \tag{10}$$

where $\mathrm{AUC}_b$ is the $b$th bootstrap estimate of AUC. The next subsection deals with the construction of confidence intervals for the proposed ROC curve to explain the variability of the curve at each and every threshold value.

*2.2. Confidence Intervals for GHROC Curve.* The $100(1-\alpha)\%$ confidence intervals for the GHROC curve are estimated using delta method. This confidence interval for the ROC Curve represents the range at each point of false positive rate and its corresponding true positive rate. Therefore, the $100(1-\alpha)\%$ confidence intervals for FPR and TPR are as follows:

$$\widehat{\mathrm{FPR}} \pm Z_{1-\alpha/2}\sqrt{\mathrm{Var}\left(\widehat{\mathrm{FPR}}\right)},$$
$$\widehat{\mathrm{TPR}} \pm Z_{1-\alpha/2}\sqrt{\mathrm{Var}\left(\widehat{\mathrm{TPR}}\right)}, \tag{11}$$

where $\widehat{\mathrm{FPR}}$ and $\widehat{\mathrm{TPR}}$ are the estimated FPR and TPR, respectively, and their variances are

$$\mathrm{Var}\left(\widehat{\mathrm{FPR}}\right) = \left(\frac{\partial \mathrm{FPR}}{\partial \sigma_H}\right)^2 \mathrm{Var}\left(\sigma_H\right) + \left(\frac{\partial \mathrm{FPR}}{\partial \alpha_H}\right)^2 \mathrm{Var}\left(\alpha_H\right),$$
$$\mathrm{Var}\left(\widehat{\mathrm{TPR}}\right) = \left(\frac{\partial \mathrm{TPR}}{\partial \sigma_D}\right)^2 \mathrm{Var}\left(\sigma_D\right) + \left(\frac{\partial \mathrm{TPR}}{\partial \alpha_D}\right)^2 \mathrm{Var}\left(\alpha_D\right). \tag{12}$$

Further, the confidence intervals for FPR and TPR can be obtained using the following expression:

$$\widehat{\mathrm{FPR}} \pm Z_{1-\alpha/2}\sqrt{\left[\left(\frac{2\alpha_H t^{\alpha_H}}{{\sigma_H}^{\alpha_H+1}}\right)\phi\left(\frac{t}{\sigma_H}\right)^{\alpha_H}\right]^2 \mathrm{Var}\left(\sigma_H\right) + \left[\left(\frac{-2t^{\alpha_H}}{{\sigma_H}^{\alpha_H}}\right)\phi\left(\frac{t}{\sigma_H}\right)^{\alpha_H}\log\left(\frac{t}{\sigma_H}\right)\right]^2 \mathrm{Var}\left(\alpha_H\right)},$$
$$\widehat{\mathrm{TPR}} \pm Z_{1-\alpha/2}\sqrt{\left[\left(\frac{\alpha_D t^{\alpha_D}}{\sigma_D^{\alpha_D+1}}\right)e^{-(t/\sigma_D)^{\alpha_D}}\right]^2 \mathrm{Var}\left(\sigma_D\right) + \left[-e^{-(t/\sigma_D)^{\alpha_D}}\left(\frac{t}{\sigma_D}\right)^{\alpha_D}\log\left(\frac{t}{\sigma_D}\right)\right]^2 \mathrm{Var}\left(\alpha_D\right)} \tag{13}$$

(for complete proof, refer to appendix). These confidence interval lines show the variability of the proposed ROC curve at each and every point on the ROC curve.

In the next section, the results are carried out using simulation studies and a real data set to explain the proposed methodology. Further, the confidence intervals are evaluated for the summary measure AUC and the intrinsic measures FPR and TPR.

## 3. Results and Discussion

The proposed methodology is demonstrated using simulation studies and real data set (SAPS III).

*3.1. Simulation Studies.* Simulation studies are conducted with different combinations of scale and shape parameters of both normal and abnormal populations and the entire simulations are done at various sample sizes $\{50, 100, 200, 300 \text{ and } 500\}$ with $B = 100$ bootstraps. At every parameter combination and sample size, the AUC and its confidence intervals are obtained. The main purpose of conducting simulations is to show how the AUC of GHROC curve possesses different values as the scale and shape parameters of the normal and abnormal distributions change. The variations in the parameter values of both populations are used to explain the overlapping area in terms of AUC; this mean that the higher the AUC, the lesser the overlapping area and vice versa. Further, to demonstrate the behavior of AUC, the entire simulation work is carried out with three different experiments. In the first experiment, the shape parameter of abnormal population is varied by fixing the other parameters as constant; in second experiment, the scale parameter of abnormal population is varied by fixing the other parameters as constant and, in the third experiment, the shape parameters of both populations are considered to be equal with varying scale in abnormal population. The results so obtained from these experiments are reported in Table 1.

In the first experiment, when $\alpha_D = 2$ with $\sigma_H = 1$, $\sigma_D = 1.5$, and $\alpha_H = 0.6$, the AUC is observed to be around 0.6791 (67.91% of accuracy) and, as $\alpha_D$ takes higher values as 3 and 5, the AUC is observed to have a better value indicating high level of accuracy, thus, reflecting the scenario that as the discrepancy between shape parameters of both normal and abnormal population's increases, AUC attains a larger value indicating a better extent of correct classification with minimum percentage of overlapping area. Suppose that

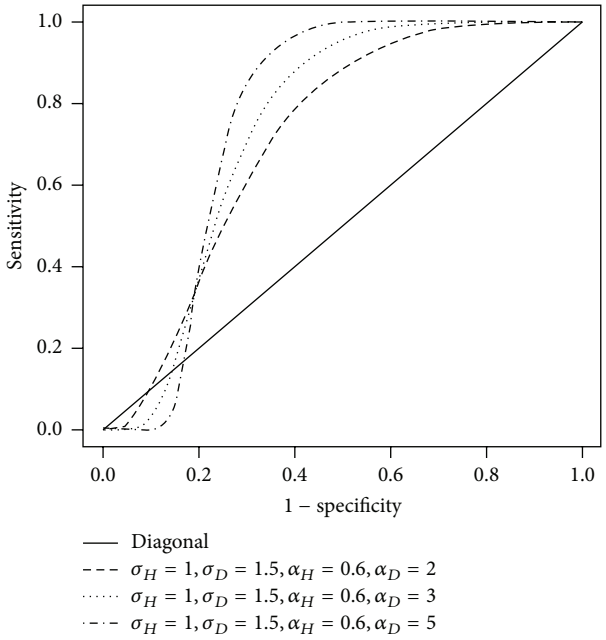
Table 1: Confidence intervals for AUC at various combinations of scale and shape parameters of normal and abnormal populations with different sample sizes.

| Experiment | Parameter values | Sample size | | | | |
|---|---|---|---|---|---|---|
| | | 50 | 100 | 200 | 300 | 500 |
| 1 | $\sigma_H = 1, \sigma_D = 1.5, \alpha_H = 0.6$ | | | | | |
| | $\alpha_D = 2$ | 0.6791 (0.5728, 0.7854) | 0.724 (0.6503, 0.7978) | 0.7036 (0.6564, 0.7509) | 0.6810 (0.6407, 0.7212) | 0.7186 (0.6864, 0.7509) |
| | $\alpha_D = 3$ | 0.7609 (0.6651, 0.8568) | 0.7336 (0.6623, 0.8049) | 0.7385 (0.6873, 0.7897) | 0.7513 (0.7116, 0.7910) | 0.7249 (0.6954, 0.7545) |
| | $\alpha_D = 5$ | 0.7311 (0.6384, 0.8239) | 0.7655 (0.7033, 0.8277) | 0.7408 (0.6971, 0.7846) | 0.7811 (0.7526, 0.8096) | 0.7807 (0.7586, 0.8028) |
| 2 | $\sigma_H = 1.5, \alpha_H = 1.5, \alpha_D = 3$ | | | | | |
| | $\sigma_D = 1.5$ | 0.5953 (0.4814, 0.7092) | 0.5588 (0.4881, 0.6294) | 0.5765 (0.5241, 0.6290) | 0.5592 (0.5198, 0.5986) | 0.5739 (0.5416, 0.6061) |
| | $\sigma_D = 2.25$ | 0.8033 (0.7230, 0.8836) | 0.7684 (0.7061, 0.8307) | 0.769 (0.7353, 0.8027) | 0.7656 (0.7328, 0.7983) | 0.7871 (0.7604, 0.8139) |
| | $\sigma_D = 3.5$ | 0.9133 (0.8737, 0.9528) | 0.8973 (0.8606, 0.9339) | 0.9132 (0.8944, 0.9321) | 0.9058 (0.8898, 0.9218) | 0.9144 (0.9028, 0.9259) |
| 3 | $\sigma_H = 1.5, \alpha_H = 2, \alpha_D = 2$ | | | | | |
| | $\sigma_D = 1.5$ | 0.5395 (0.4340, 0.6450) | 0.5212 (0.4480, 0.5943) | 0.4886 (0.4347, 0.5425) | 0.5395 (0.4949, 0.5841) | 0.5178 (0.4840, 0.5516) |
| | $\sigma_D = 2.25$ | 0.7303 (0.6296, 0.8310) | 0.6904 (0.6146, 0.7663) | 0.7222 (0.6803, 0.7641) | 0.7101 (0.6756, 0.7447) | 0.7084 (0.6776, 0.7391) |
| | $\sigma_D = 3.5$ | 0.8319 (0.7562, 0.9076) | 0.8193 (0.7689, 0.8697) | 0.8458 (0.8076, 0.8840) | 0.8584 (0.8352, 0.8815) | 0.8563 (0.8400, 0.8726) |

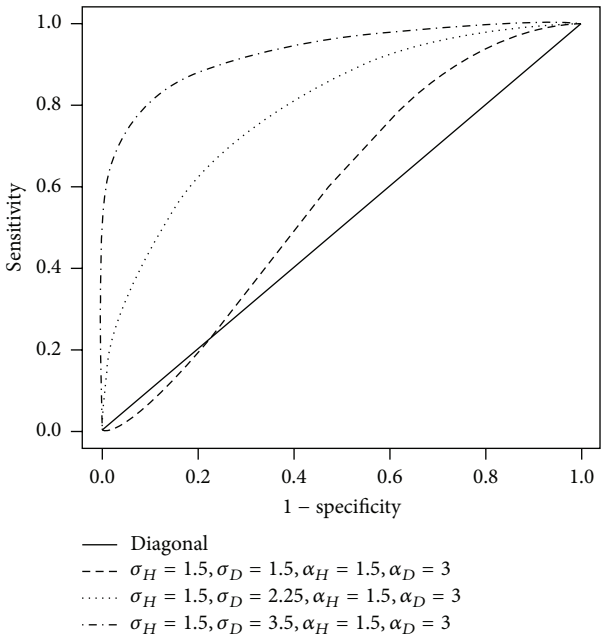
if we have real data set with these parameter values then that particular test will provide a better accuracy. Along with the shape, scale parameter also influences the measure AUC. Further, in experiment 2, scale parameter of abnormal population ($\sigma_D$) is varied by keeping all the other parameters ($\alpha_D, \alpha_H, \sigma_H$) as constant. Moderate levels of discrepancy in the shape values and scale parameters influence the accuracy of the classification. As $\sigma_D$ attains a larger value, the AUC of GHROC curve tend to have better values of accuracy. So this reveals that along with discrepancy in shape parameters of both populations, scale parameter also tends to explain better variability in the data giving rise to talk about the exact performance of the test considered. The accuracy of the test needs to be examined when there is an equal discrepancy in the shape parameter with varying scale parameters. This is addressed by conducting another experiment (third). Here, the first part is defined by considering the scale and shape parameters of both populations to be equal and, in the second part of this experiment, the scale parameters are varied by taking equal shape parameters. The first part reveals the finding that, when all parameters tend to be equal to unit value, then two populations get overlapped giving rise to having AUC nearer to 0.5. The results of the second part outline the observation that even though the shape parameters are equal, the discrepancy in scale parameters of abnormal population tends to explain the hidden accuracy and when the discrepancy between the scale values of two populations is larger, the explanation about the accuracy of the test can be given better. Thus, from three experiments it is noticed that shape parameter has its major influence in explaining better accuracy of a test than that observed with scale parameter alone. However, scale parameter also has its role in explaining the accuracy and it should not be neglected.

To demonstrate the proposed methodology with the help of graphical visualization, ROC curves are drawn for three experiments (Figure 1). From Figure 1(a), it is visualized that the curve moves towards the top left corner of the plot with increasing accuracy as the shape of abnormal population tends to have a larger value. Further, Figure 1(b) explains the effect of scale in abnormal population and it can be seen that the curve moves away from chance line with high accuracy as the scale attains larger value in abnormal population. Figure 1(c) illustrates the effect of scale parameter in presence of equal shape parameter and it is observed that the shape of the ROC curve is affected as the scale changes. From Figure 1, it is reasonable to say that the proposed ROC curve completely depends on the shape and scale parameters of normal and abnormal populations.

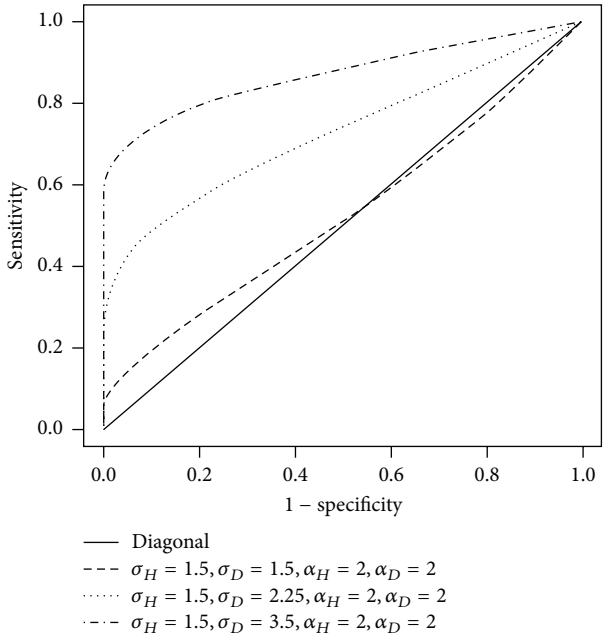
Apart from explaining the importance and the influence of the scale and shape parameters in GHROC context, it is essential to construct the confidence intervals for the measures of GHROC curve. This attempt is to illustrate the changing behavior of the estimates of the proposed ROC curve. In statistical literature, the theory of interval estimation has gained its importance over point estimation because it reveals the true information of the estimate within the potential uncertainties. Hence, it is very important to address the position of the true estimate in the presence of sample size within the range of potential uncertainties. The $100(1 - \alpha)\%$ confidence intervals are constructed for




(a) Effect of shape parameter in abnormal population (experiment 1)




(b) Effect of scale parameter in abnormal population (experiment 2)




(c) Effect of scale parameter in abnormal population with constant shape parameter (experiment 3)

Figure 1: Plot of ROC curves with various combinations of scale and shape parameters of both normal and abnormal populations.

all the combinations which are defined as three different experiments.

With respect to the approach of confidence interval, the perception about the impact of sample size on the width of the confidence intervals and the graphical visualization of the true estimates of GHROC curve along with its confidence intervals is more important and to be addressed. From the results, it is evident that the sample size effect can be witnessed in terms of the width of the confidence interval, notifying that the true estimate is independent from the effect of sample size and its corresponding confidence interval possesses a narrowing-down phenomenon. These simulation studies points out the information that, irrespective of the sample size and width of the confidence interval, the information about the true estimate of the ROC curve lies within the potential uncertainties. Even though this is a generally observed phenomenon but the fact to be noticed is that the variability in the populations will get diminished

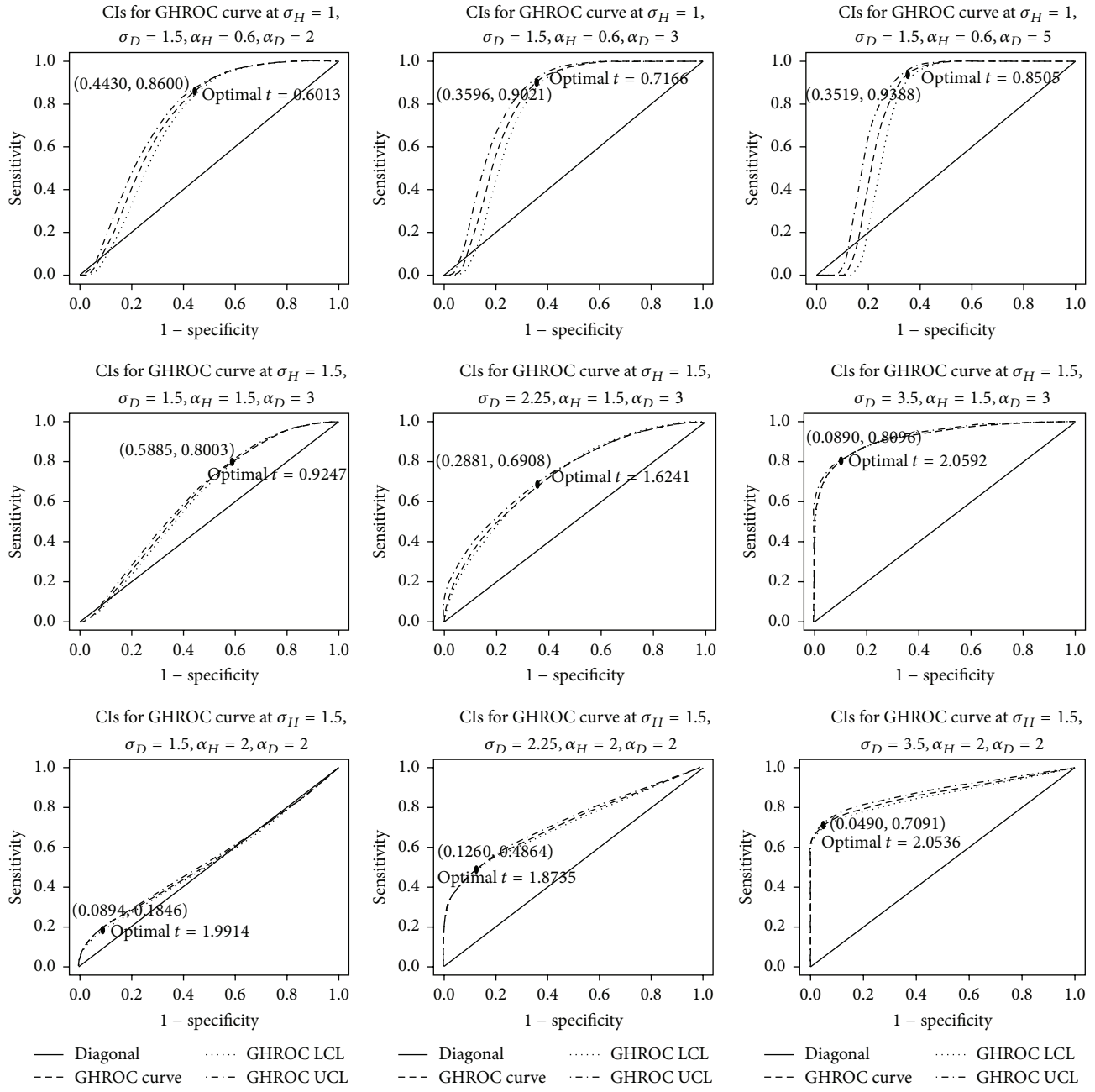



Figure 2: The confidence intervals for GHROC curve at various combinations of scale and shape parameters of both populations.

as the sample size takes a larger number, giving rise to a shortened confidence interval.

Figure 2 clearly explains the variability of GHROC curve at each and every point on the ROC curve. This means the lower control limit and the upper control limit for the proposed ROC curve are plotted at a particular sample size $n = 200$ (Figure 2) and these curves explain the range of false positive rate and true positive rate at each and every threshold. Further, the optimal threshold is also depicted in Figure 2 along with the pair (FPR, TPR) obtained at that particular optimal threshold. Further, this optimal threshold is used to classify the subjects with better accuracy and this can be used as a reference value for future classification. For example, consider the combination $\sigma_H = 1.5$, $\sigma_D = 3.5$, $\alpha_H = 1.5$, and $\alpha_D = 3$ and, at this combination, the optimal threshold is found to be 2.0592 with true positive rate 0.8096. This explains the identification of abnormal subjects as abnormal with 80.96% of correct classification at the optimal threshold value 2.0592 for the considered combination. At the case of worst classification (equal scale and shape parameters), the optimal threshold is observed to be 1.9914 with very less value of true positive rate 0.1846

Table 2: Results for SAPS III using GHROC curve methodology.

| $\widehat{\sigma}_H$ | $\widehat{\sigma}_D$ | $\widehat{\alpha}_H$ | $\widehat{\alpha}_D$ | $t$ | AUC (LCL, UCL) |
|---|---|---|---|---|---|
| 32.6927 | 37.9974 | 1.1831 | 1.9872 | 22 | 0.6278 (0.5530, 0.6917) |

(Figure 2). Similarly, the interpretation can be given for all the remaining combinations which are considered in the study using Figure 2.

*3.2. Real Data Set.* The real data set is about the ICU scoring system; SAPS III is a system for predicting mortality (dead or alive) status of a patient in ICU. SAPS III has been designed to provide a real-life predicted mortality for a patient by following a well defined procedure, based on a mathematical model that needs calibration. This data consists of a total of 111 respondents of which 66 (59.45%) are alive and 45 (40.54%) are dead.

From this data set it is observed that the SAPS III scores for dead patients follow Weibull distribution (KS − Statistic = 0.1280; $p$ value = 0.4165 at 0.05 level of significance) whereas the scores for patients who are alive follow GHN distribution (KS − Statistic = 0.0901; $p$ value = 0.6243 at 0.05 level of significance). The results for the prognosis of disease are reported in Table 2. It is observed that the accuracy of the test is 62.78% indicating that the SAPS III score is able to identify the status of mortality about 62.78%. The optimal threshold value is observed to be 22.00 which means that when the SAPS III score exceeds the optimal threshold 22.00, the patient will have 71.35% chance of death. Further, the confidence interval of AUC is (0.5530, 0.6917) and the proposed ROC curve for SAPS III uniformly lies above the chance line to explain the mortality rate (Figure 3) depicting lower and upper confidence intervals for proposed ROC curve along with its optimal threshold.

## 4. Conclusion

The present paper is focused on addressing the practical issue where the populations with and without condition underlie two different generalized skewed distributions with scale and shape parameters which are useful in explaining and handling skewed nature of the data. Simulation studies are conducted at various combinations of the parameters. The entire exercise is done using three experiments and the effect of sample size is also noted. Further, it is observed that the width of the confidence interval is affected by the size of the sample in turn providing shortened confidence intervals as sample size is considered to be large. Moreover, from the proposed methodology it is feasible to identify the sensitivity at a specific false positive rate and vice versa.

Further, the proposed methodology is applied to a real data set, namely, SAPS III, which is used to predict the mortality status of the patient in ICU. The accuracy of SAPS III system in predicting the mortality event, death, is 62.78%. The optimal threshold is identified to be 22.00 which can be used to identify the status of a new individual whose SAPS III score is calculated.

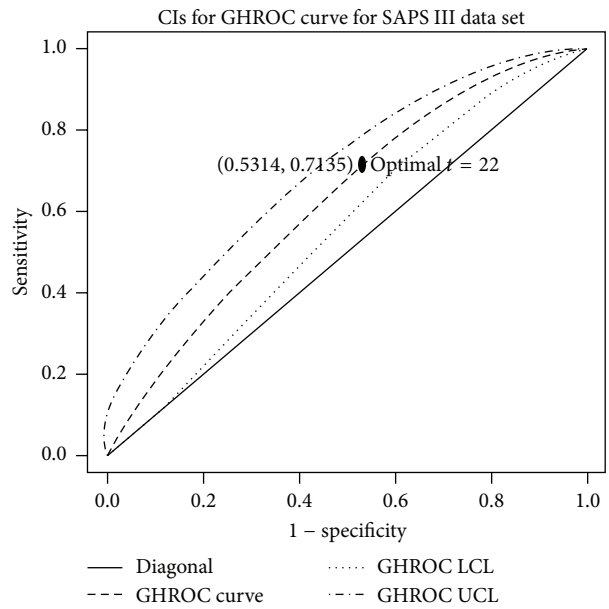



Figure 3: GHROC curve for SAPS III with its confidence intervals.

## Appendix

The partial differentiations of FPR and TPR with respect to their parameters are

$$\begin{aligned}
\frac{\partial \mathrm{FPR}}{\partial \sigma_H} &= \left(\frac{2\alpha_H t^{\alpha_H}}{{\sigma_H}^{\alpha_H+1}}\right)\phi\left(\frac{t}{\sigma_H}\right)^{\alpha_H},\\
\frac{\partial \mathrm{FPR}}{\partial \alpha_H} &= \left(\frac{-2t^{\alpha_H}}{{\sigma_H}^{\alpha_H}}\right)\phi\left(\frac{t}{\sigma_H}\right)^{\alpha_H}\log\left(\frac{t}{\sigma_H}\right),\\
\frac{\partial \mathrm{TPR}}{\partial \sigma_D} &= \left(\frac{\alpha_D t^{\alpha_D}}{\sigma_D^{\alpha_D+1}}\right)e^{-(t/\sigma_D)^{\alpha_D}},\\
\frac{\partial \mathrm{TPR}}{\partial \alpha_D} &= -e^{-(t/\sigma_D)^{\alpha_D}}\left(\frac{t}{\sigma_D}\right)^{\alpha_D}\log\left(\frac{t}{\sigma_D}\right).
\end{aligned} \tag{A.1}$$

Now, by substituting the above partial derivatives in (12), we have

$$\begin{aligned}
&\mathrm{Var}\left(\widehat{\mathrm{FPR}}\right)\\
&= \left[\left(\frac{2\alpha_H t^{\alpha_H}}{{\sigma_H}^{\alpha_H+1}}\right)\phi\left(\frac{t}{\sigma_H}\right)^{\alpha_H}\right]^2 \mathrm{Var}\left(\sigma_H\right)\\
&\quad + \left[\left(\frac{-2t^{\alpha_H}}{{\sigma_H}^{\alpha_H}}\right)\phi\left(\frac{t}{\sigma_H}\right)^{\alpha_H}\log\left(\frac{t}{\sigma_H}\right)\right]^2 \mathrm{Var}\left(\alpha_H\right),
\end{aligned} \tag{A.2}$$

$$\begin{aligned}
&\mathrm{Var}\left(\widehat{\mathrm{TPR}}\right)\\
&= \left[\left(\frac{\alpha_D t^{\alpha_D}}{\sigma_D^{\alpha_D+1}}\right)e^{-(t/\sigma_D)^{\alpha_D}}\right]^2 \mathrm{Var}\left(\sigma_D\right)\\
&\quad + \left[-e^{-(t/\sigma_D)^{\alpha_D}}\left(\frac{t}{\sigma_D}\right)^{\alpha_D}\log\left(\frac{t}{\sigma_D}\right)\right]^2 \mathrm{Var}\left(\alpha_D\right).
\end{aligned} \tag{A.3}$$

The bootstrapped estimates and their variances of the parameters $\sigma_H$, $\alpha_H$, $\sigma_D$, and $\alpha_D$ are

$$\widehat{\sigma_H} = \frac{1}{B}\sum_{b=1}^{B}\sigma_{Hb},$$

$$\mathrm{Var}\left(\widehat{\sigma_H}\right) = \frac{1}{B-1}\sum_{b=1}^{B}\left(\sigma_{Hb} - \widehat{\sigma_H}\right)^2,$$

$$\widehat{\alpha_H} = \frac{1}{B}\sum_{b=1}^{B}\alpha_{Hb},$$

$$\mathrm{Var}\left(\widehat{\alpha_H}\right) = \frac{1}{B-1}\sum_{b=1}^{B}\left(\alpha_{Hb} - \widehat{\alpha_H}\right)^2,$$

$$\widehat{\sigma_D} = \frac{1}{B}\sum_{b=1}^{B}\sigma_{Db},$$

$$\mathrm{Var}\left(\widehat{\sigma_D}\right) = \frac{1}{B-1}\sum_{b=1}^{B}\left(\sigma_{Db} - \widehat{\sigma_D}\right)^2,$$

$$\widehat{\alpha_D} = \frac{1}{B}\sum_{b=1}^{B}\alpha_{Db},$$

$$\mathrm{Var}\left(\widehat{\alpha_D}\right) = \frac{1}{B-1}\sum_{b=1}^{B}\left(\alpha_{Db} - \widehat{\alpha_D}\right)^2, \tag{A.4}$$

where $\sigma_{Hb}$, $\sigma_{Db}$, $\alpha_{Hb}$, and $\alpha_{Db}$ are the $b$th bootstrap estimates of $\sigma_H$, $\sigma_D$, $\alpha_H$, and $\alpha_D$, respectively.

Now, by substituting the above variances of the parameters of two considered distributions in (A.2) and (A.3), we can obtain the expressions for the variances of FPR and TPR, respectively. Further, using (A.2) and (A.3), the confidence intervals can be estimated for the intrinsic measures which results in producing the confidence intervals for the proposed ROC curve as follows:

$$\widehat{\mathrm{FPR}} \pm Z_{1-\alpha/2}\sqrt{\left[\left(\frac{2\alpha_H t^{\alpha_H}}{{\sigma_H}^{\alpha_H+1}}\right)\phi\left(\frac{t}{\sigma_H}\right)^{\alpha_H}\right]^2 \mathrm{Var}\left(\sigma_H\right) + \left[\left(\frac{-2t^{\alpha_H}}{{\sigma_H}^{\alpha_H}}\right)\phi\left(\frac{t}{\sigma_H}\right)^{\alpha_H}\log\left(\frac{t}{\sigma_H}\right)\right]^2 \mathrm{Var}\left(\alpha_H\right)},$$

$$\widehat{\mathrm{TPR}} \pm Z_{1-\alpha/2}\sqrt{\left[\left(\frac{\alpha_D t^{\alpha_D}}{\sigma_D^{\alpha_D+1}}\right)e^{-(t/\sigma_D)^{\alpha_D}}\right]^2 \mathrm{Var}\left(\sigma_D\right) + \left[-e^{-(t/\sigma_D)^{\alpha_D}}\left(\frac{t}{\sigma_D}\right)^{\alpha_D}\log\left(\frac{t}{\sigma_D}\right)\right]^2 \mathrm{Var}\left(\alpha_D\right)}. \tag{A.5}$$

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgment

The authors would like to thank and acknowledge Dr. Vimal Kumar, Department of Public Health and Medicine, SRM Medical College Hospital and Research Centre, Chennai, India, for sharing the data to carry out the results.

## References

[1] J. P. Egan, *Signal Detection Theory and ROC Analysis*, Academic Press, New York, NY, USA, 1975.

[2] D. D. Dorfman and E. Alf Jr., "Maximum likelihood estimation of parameters of signal detection theory—a direct solution," *Psychometrika*, vol. 33, no. 1, pp. 117–124, 1968.

[3] D. D. Dorfman and E. Alf Jr., "Maximum-likelihood estimation of parameters of signal-detection theory and determination of confidence intervals-rating-method data," *Journal of Mathematical Psychology*, vol. 6, no. 3, pp. 487–496, 1969.

[4] K. H. Zou, W. J. Hall, and D. E. Shapiro, "Smooth non-parametric receiver operating characteristic (ROC) curves for continuous diagnostic tests," *Statistics in Medicine*, vol. 16, no. 19, pp. 2143–2156, 1997.

[5] L. Tang, P. Du, and C. Wu, "Compare diagnostic tests using transformation-invariant smoothed ROC curves," *Journal of Statistical Planning and Inference*, vol. 140, no. 11, pp. 3540–3551, 2010.

[6] L. L. Tang and N. Balakrishnan, "A random-sum Wilcoxon statistic and its application to analysis of ROC and LROC data," *Journal of Statistical Planning and Inference*, vol. 141, no. 1, pp. 335–344, 2011.

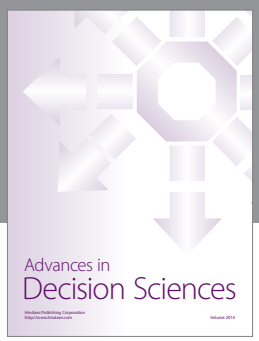
[7] E. Hussain, "The bi-gamma roc curve in a straightforward manner," *Journal of Basic & Applied Sciences*, vol. 8, no. 2, pp. 309–314, 2012.

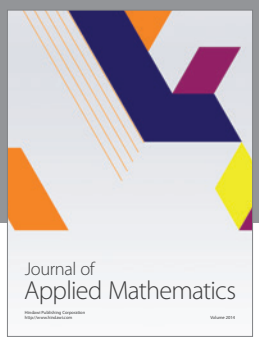
[8] S. Balaswamy, R. V. Vardhan, and K. V. S. Sarma, "The hybrid ROC (HROC) curve and its divergence measures for binary classification," *International Journal of Statistics in Medical Research*, vol. 4, no. 1, pp. 94–102, 2015.

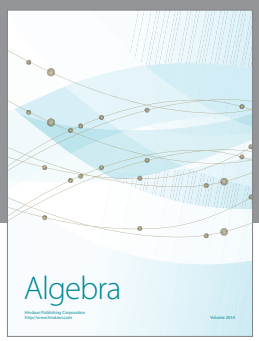
[9] K. Cooray and M. M. A. Ananda, "A generalization of the half-normal distribution with applications to lifetime data," *Communications in Statistics—Theory and Methods*, vol. 37, no. 9, pp. 1323–1337, 2008.

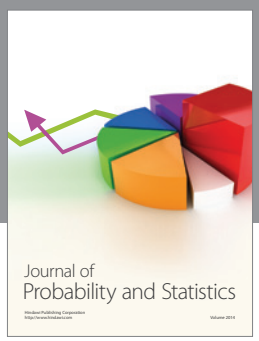
[10] D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics*, John Wiley & Sons, New York, NY, USA, 1966.

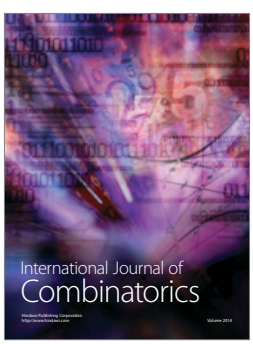
[11] D. Bamber, "The area above the ordinal dominance graph and the area below the receiver operating characteristic graph," *Journal of Mathematical Psychology*, vol. 12, no. 4, pp. 387–415, 1975.

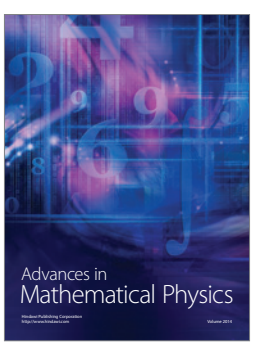


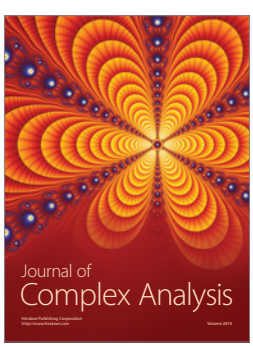



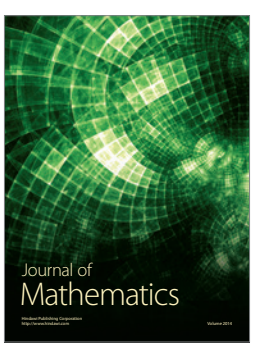



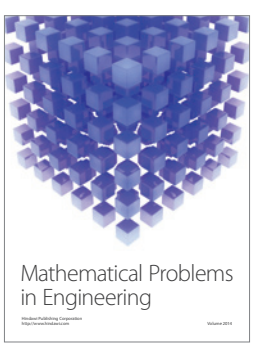



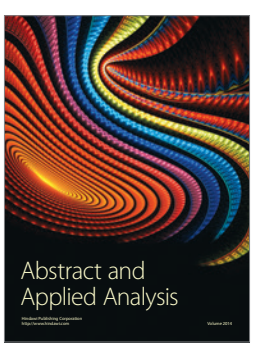


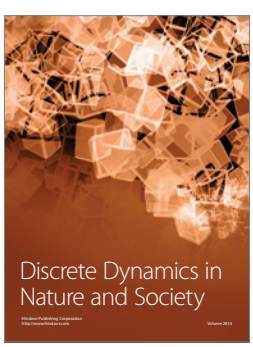



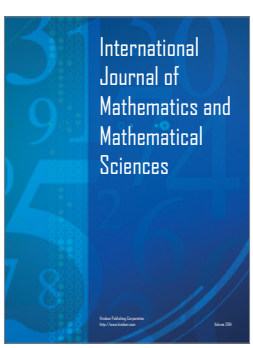



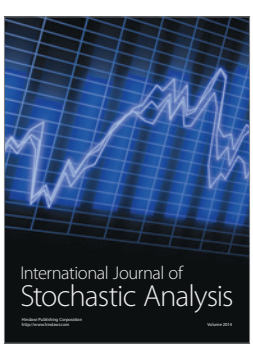





Hindawi

Submit your manuscripts at
http://www.hindawi.com