The Hybrid ROC (HROC) Curve and its Divergence Measures for Binary Classification

S. Balaswamy[1], R. Vishnu Vardhan[1,*] and K.V.S. Sarma[2]

[1]*Department of Statistics, Pondicherry University, Puducherry – 605 014, India*

[2]*Department of Statistics, Sri Venkateswara University, Tirupati – 517 502, India*

Abstract: In assessing the performance of a diagnostic test, the widely used classification technique is the Receiver Operating Characteristic (ROC) Curve. The Binormal model is commonly used when the test scores in the diseased and healthy populations follow Normal Distribution. It is possible that in real applications the two distributions are different but having a continuous density function. In this paper we considered a model in which healthy and diseased populations follow half normal and exponential distributions respectively, hence named it as the Hybrid ROC (HROC) Curve. The properties and Area under the curve (AUC) expressions were derived. Further, to measure the distance between the defined distributions, a popular divergence measure namely Kullback Leibler Divergence (KLD) has been used. Simulation studies were conducted to study the functional behavior of Hybrid ROC curve and to show the importance of KLD in classification.

Keywords: AUC, Exponential distribution, Half-Normal distribution, Hybrid ROC Curve, Kullback-Leibler Divergence.

1. INTRODUCTION

In the recent years, the Receiver Operating Characteristic (ROC) curve analysis has become a popular statistical technique in the field of medical diagnosis. Even though it originated during Second World War, (Green and Swets, [1]), many researchers highlighted its significance in Medicine, Experimental Psychology, Finance, Banking, data mining etc., in later years. A considerable amount of work has been carried out on the methods such as estimation of Area Under the Curve (AUC), Maximum Likelihood Estimation, Regression Methods for estimating ROC and its related measures in the past seven decades of which few are mentioned here, Green and Swets [1], Oglive and Creelman [2], Dorfman and Alf [3,4], Lusted L.B. [5], Bamber [6], Egan [7], Metz CE [8], Swets *et al.* [9], Hanley and Mc Neil [10,11], Hanley [12], Gaddard and Hindberg [13], Pepe [14-16], Alonzo and Pepe [17], Zhang and Pepe [18], Krazanowski and Hand [19], R Vishnu Vardhan and KVS Sarma [20,21]. Further, the framework of ROC curve is formulated basing on some distributions, Kernel based methods, Bayesian approach, Meta Analysis, censored and truncated data. The summary measure AUC allows us to compare two diagnostic tests and also acts as a measure to compare two statistical tools.

Apart from well-known statistical classification procedures like Logistic Regression and Discriminant analysis, the ROC curve has its mathematical formulation which helps in fitting and estimating the parameters of the curve. The entire classification will be carried out on the basis of a threshold value often referred to as *Gold Standard* and determines the true condition status. If the condition status is true, it indicates the presence of disease, otherwise.

Two basic measures of ROC curve are sensitivity (S_n) and specificity (S_p). Sensitivity refers to the ability of a test to detect the condition when it is present and Specificity refers to the ability of test to exclude patients without the condition. An ROC curve is a plot of $1\text{-}S_p$ versus S_n. The construction of ROC curve primarily depends on the four possible states which are obtained on the basis of a threshold value i.e., TP, TN, FN and FP. The resulting curve is called empirical ROC. Conventionally, it is assumed that diseased (Y) and the healthy (X) individuals follow Normal distribution and hence the name *Binormal ROC curve,* with unknown monotonic transformation (Farraggi and Reiser [22]).

There are two main objectives of ROC curve analysis. The first one is to identify the best cutoff in some sense and the other one is to choose the best test/procedure among several procedures (called biomarkers) in terms of AUC.

In the following section, the ROC methodology, AUC and some properties of the binormal model are discussed.

2. BINORMAL MODEL OF ROC CURVE

Let D denote the individuals in the diseased group and H the individuals in the healthy group. Let X and Y

*Address correspondence to this author at the Department of Statistics, Pondicherry University, Puducherry – 605 014, India;
E-mail: rvvcrr@gmail.com

E-ISSN: 1929-6029/15 © 2015 Lifescience Global

denote the random variables denoting the test value in the groups H and D respectively. Further assume that X~N(μ_H, σ_H^2) and Y~N(μ_D, σ_D^2), where the parameters have their usual meaning.

Let 'S' be the test scores of a diagnostic test and 't' be the threshold value or cutoff, which will classify the unlabelled individuals into one of the two groups. In order to assess the accuracy of this classifier, we need to calculate the probability of making an incorrect allocation, since such probability provides the rate at which future individuals requiring classification will be misallocated. Now, we define four possible probabilities at this cutoff.

i. The probability that an individual from D is correctly classified.

   True Positive Rate, TP = P(S>t|D)
   (Sensitivity)

ii. The probability that an individual from H is misclassified.

   False Positive Rate, FP = P(S>t|H)
   (1-Specificity)

iii. The probability that an individual from H is correctly classified.

   True Negative Rate, TN = P(S ≤ t|H)
   (Specificity)

iv. The probability that an individual from D is misclassified.

   False Negative Rate, FN = P(S ≤ t|D).

These four probabilities describe the performance of the test at this cutoff. It is to be noted that for a good performance, we require "high" true and "low" false rates.

It is assumed that the mean test score of D group will be greater than the mean of the H group i.e., $\mu_D > \mu_H$, but no constraints are placed on the standard deviations. Now, define S as the total score of the test value, then $(S-\mu_D)/\sigma_D$ has a standard normal distribution in D and $(S-\mu_H)/\sigma_H$ has a standard normal distribution in H.

Suppose FPR is x(t), with corresponding classifier or cutoff t, then

$$x(t) = P(S > t / H) = P\left(Z > \frac{(t - \mu_H)}{\sigma_H} \right)$$

$$x(t) = \Phi\left(\frac{\mu_H - t}{\sigma_H} \right)$$

where Z is the standard normal deviate and $\Phi(.)$ is the normal cumulative distribution function.

On further simplification, one can get the threshold t as,

$$t = \mu_H - \sigma_H Z_x; \text{ where } Z_x = \Phi^{-1}[x(t)]$$

The TPR is defined as y(t) at each x(t) with threshold t is as follows,

$$y(t) = P(S > t|D) = P\left(Z > \left(\frac{t - \mu_D}{\sigma_D} \right) \right) = \Phi\left[\frac{\mu_D - c}{\sigma_D} \right]$$

$$y(t) = \Phi\left[a + b\Phi^{-1}(x) \right] \quad (1)$$

where $a = \frac{\mu_D - \mu_H}{\sigma_D}, b = \frac{\sigma_H}{\sigma_D}$

Now, here are some properties of ROC curve (Krzanowski & Hand [19]).

Properties of the ROC

i. Y = h(x) is the mathematical model of the ROC curve, where y denotes the true positive rate and x denotes the false positive rate. The curve is a monotonic increasing function in the positive quadrant, lying between y=0 at x=0 and y=1 at x=1.

ii. The ROC curve is unaltered if the classification scores undergo a strictly increasing transformation.

iii. The slope of the ROC curve at threshold value 't' is given by

$$\frac{dy}{dx} = \frac{P(S > t|D)}{P(S > t|H)}$$

The formal definition of AUC is, $AUC = \int_0^1 y(t)dx(t)$, which shows that the total area under the ROC curve (or) domain is 1.0. If A and B are two thresholds such that the ROC curve for A nowhere lies below the ROC curve for B, then AUC for A must be greater than or equal to AUC for B, but the reverse implication is not

true because of the possibility that the two curves can cross each other. In Figure **1**, typical forms of ROC curves are presented with varying cutoffs. Higher the AUC, better will be the discriminating ability of the test. As the distance between the diagonal and the upper left corner is more, then that test is said to be the best test and can be used for classification. The accuracy of a diagnostic test can be explained by using the Area under the Curve (AUC) of an ROC curve. AUC describes the ability of the test to discriminate between diseased and non-diseased. AUC gives us information about the general "goodness" of a test and not the interpretation of a test result. ROC curve starts at the point (0, 0) and ends at (1, 1) and the diagonal line separates the area into two halves. AUC can be interpreted as a probability that a randomly selected subject with disease will have a higher test result compared to a normal subject.

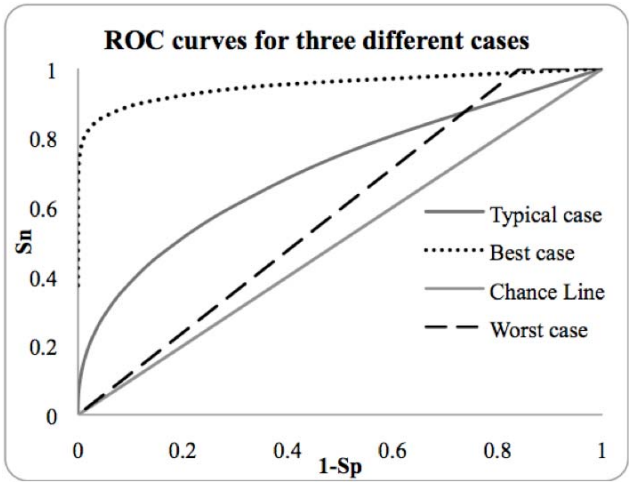


Figure 1: Different shapes of ROC curves.

Bamber [6] interpreted AUC as, "the probability that the threshold will allocate a higher score to a randomly chosen individual from population D than it will to a randomly and independently chosen individual from population H". That is, if S_D and S_H are the scores allocated to randomly and independently chosen individuals from D and H respectively, then

$$AUC=P(S_D > S_H) \quad (2)$$

The AUC expression for Binormal model is,

$$AUC = \Phi\left(\frac{a}{\sqrt{1+b^2}}\right) \quad (3)$$

Hanley [13], discussed the robustness of the binormal model by mentioning two methods for estimating the parameters of the ROC curve. One is by plotting the ROC points on a Binormal deviate paper and the other method which is a formal procedure (Maximum Likelihood Estimation) given by Dorfman and Alf [4], Ogilive and Creelman [2]. With the Binormal form, the ROC curve is a graph generated by two overlapping normal distributions and hence it can also be referred to as the finite mixture of distributions.

In next section, a newer version of ROC curve and Area under the Curve is proposed and its properties are derived. Here, the test scores of healthy (H) and diseased (D) populations follow Half-Normal and Exponential distributions respectively, hence named it as *Hybrid ROC (HROC) Curve*.

3. HYBRID ROC (HROC) CURVE METHODOLOGY

Let us assume that the test scores X and Y of healthy and diseased populations follow half-normal and exponential distributions respectively. The cumulative distribution functions of half-normal and exponential distributions are as follows,

$$F(x) = \int_0^x \frac{1}{\sigma}\sqrt{\frac{2}{\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx; x > 0, \sigma^2 > 0 \quad (4)$$

Using the change of variable $z = \frac{x}{\sigma\sqrt{2}}$ then the CDF can be rewritten as

$$F(x) = \frac{2}{\sqrt{\pi}} \int_0^{x/\sigma\sqrt{2}} e^{-z^2} dz = erf\left(\frac{x}{\sigma\sqrt{2}}\right) \quad (5)$$

where 'σ' is the scale parameter and erf(x) is the error function.

$$G(y) = 1 - \exp\left(-\frac{y}{\sigma}\right); y \geq 0, \sigma > 0 \quad (6)$$

where 'σ' is the scale parameter.

Let x(t) denote the FPR, function for the horizontal coordinate and y(t) denote the TPR, vertical coordinate, i.e., x(t) = 1-F(t) and y(t) = 1-G(t).

In the conventional process of classification, higher values of test scores attribute to diseased population, otherwise. As it is well known that the ROC Curve is a function of S_n as a function of 1-S_p. Therefore, to derive the proposed HROC Curve, the FPR is defined as

$$x(t) = P(S > t|H) = 1 - erf\left(\frac{t}{\sigma_H\sqrt{2}}\right)$$

$$x(t) = 2\left\{1 - \Phi\left(\frac{t}{\sigma_H}\right)\right\} \quad (7)$$

The threshold value can be obtained by using the expression (7) and is given below

$$t = \sigma_H \Phi^{-1}\left(1 - \frac{x(t)}{2}\right) \tag{8}$$

Similarly, TPR is defined as,

$$y(t) = P(S > t|D) = 1 - \left(1 - \exp\left(\frac{-t}{\sigma_D}\right)\right)$$

$$y(t) = \exp\left(-\frac{\sigma_H}{\sigma_D}\Phi^{-1}\left(1 - \frac{x(t)}{2}\right)\right) \tag{9}$$

Here, σ_H, σ_D are the scale parameters of healthy (Half-Normal) and diseased (Exponential) populations. The expression obtained in equation (9) is the Hybrid ROC (HROC) Curve.

The AUC expression for the Hybrid ROC curve can be obtained by integrating (9) over [0, 1].

$$i.e., \text{AUC} = \int_0^1 \exp\left(-\frac{\sigma_H}{\sigma_D}\Phi^{-1}\left(1 - \frac{x(t)}{2}\right)\right)dx(t)$$

$$\text{AUC} = \int_0^1 \exp\left(\frac{-\sqrt{2\sigma_H}}{\sigma_D}\text{erf}^{-1}(1 - x(t))\right)dx(t)$$

Let $t = erf - 1(x(t) - 1)$

$$dt = \frac{\sqrt{\pi}}{2} exp\left((erf^{-1}(x(t) - 1))^2\right)dx(t)$$

By substituting the above expression in AUC, one can get,

$$AUC = \frac{2}{\sqrt{\pi}}\int_{-\infty}^{0} exp(kt - t^2)dt$$

On further simplification using Mathematica, one can get the accuracy measure (AUC) as follows,

$$AUC = \left[1 - erf\left(\frac{\sigma_H}{\sqrt{2\sigma_D}}\right)\right]exp\left(\frac{\sigma_H^2}{2\sigma_D^2}\right) \tag{10}$$

Once the ROC Curve and AUC expressions are obtained, next step is to derive the properties which exhibit the functionality of the HROC Curve. The three basic properties of Binormal ROC model (Krazanowski and Hand [19]) have been verified for the proposed HROC Curve.

3.1. Properties

Property 1: Hybrid ROC Curve is Monotonically Increasing

Proof: Let us consider two false positive values P_1 and P_2 such that $P_1 < P_2$ and $\Phi^{-1}(.)$ be a strictly increasing function.

Since $P_1 < P_2$ which implies that

$$1 - \frac{P_1}{2} > 1 - \frac{P_2}{2} \Rightarrow \Phi^{-1}\left[1 - \frac{P_1}{2}\right] \geq \Phi^{-1}\left[1 - \frac{P_2}{2}\right]$$

$$\Rightarrow \exp\left\{-\frac{\sigma_H}{\sigma_D}\Phi^{-1}\left(1 - \frac{P_1}{2}\right)\right\} \leq \exp\left\{-\frac{\sigma_H}{\sigma_D}\Phi^{-1}\left(1 - \frac{P_2}{2}\right)\right\}$$

$$\Rightarrow \text{ROC}(P_1) \leq \text{ROC}(P_2)$$

Hence, the HROC Curve is monotonically increasing with its FPR.

Property 2: Slope of the Hybrid ROC Curve Equals the likelihood Ratio

Proof: The derivative of ROC curve at a given pair of coordinates equals the likelihood ratio. Let us parameterize x and y in terms of 't' and the derivative can be written as

$$\frac{dy}{dx} = \frac{dy/dt}{dx/dt} = \frac{g(t)}{f(t)}$$

The derivatives of the cumulative distribution functions F(t) and G(t) are the probability distribution functions f(t) and g(t). Therefore the derivative of the Hybrid ROC Curve is

$$\frac{dy}{dx} = \frac{g(t)}{f(t)}$$

$$\Rightarrow \frac{dy}{dx} = \frac{\frac{1}{\sigma_D}\exp\left(-t/\sigma_D\right)}{\frac{\sqrt{2}}{\sigma_H\sqrt{\pi}}\exp\left(-t^2/2\sigma_H^2\right)} \tag{11}$$

where, $\sigma_H > 0$, $\sigma_D > 0$; $t \geq 0$ and $\sigma_D > \sigma_H$

On simplifying the above equation (11), we have

$$\frac{dy}{dx} = \left(\frac{\sigma_H}{\sigma_D}\right)\sqrt{\frac{\pi}{2}}\exp\left(\frac{t(\sigma_D t - 2\sigma_H^2)}{2\sigma_D\sigma_H^2}\right) \geq 0$$

which is the ratio of the distribution of diseased scores to healthy scores of the two probability densities at the

value of 't'. This is referred to as *likelihood ratio of HROC Curve*.

Property 3: The Hybrid ROC Curve is Invariant under Strictly Increasing Transformation

Proof: Let 'S' denote the set of scores with $s \subset \Re$ and h(.) is strictly increasing function. Let a, b ϵ S and a < b, then by using the strictly increasing function, we can write h(a) < h(b).

The transformed random variables U and V from the respective healthy and diseased classes are

$P(U \leq t) = P[h(U) \leq h(t)$ & $P(V \leq t) = P[h(V) \leq h(t)]$

Let us consider the points (x*(t), y*(t)) on the ROC Curve for the transformed scores,

$x^*(t) = P\{h(U) > h(t) \mid H\} = 1-P\{h(U) \leq h(t)\}$

$=1-P(U \leq t) = x(t)$

$y^*(t) = P(h(V) > h(t)|D\} = 1-P\{h(V) \leq h(t)\}$

$=1-P(V \leq t) = y(t)$

hus the Hybrid ROC Curve is invariant to transformation.

In the following section, the mathematical and practical importance of Kullback-Leibler Divergence (KLD) is highlighted. Further simulation studies are conducted to study the behavior of Hybrid ROC curve and how KLD can be used to explain the proximity between the two populations in terms of ROC curve.

In next section, a brief introduction about the divergence measure is given.

4. KLD AS A MEASURE OF CLASSIFICATION

The Kullback – Leibler Divergence (KLD) is a fundamental equation of information theory that quantifies the proximity of two probability distributions. KLD is popular because it arises from likelihood theory and provides the relative entropy of distribution to a reference measure. It is always non-negative and equals zero if and only if the two distributions are identical. Fisher [23] gave a meaningful introduction about the *criterion of sufficiency*; means that the statistic chosen should summarize the whole of the relevant information supplied by the sample and concern was on the statistical problem of discrimination by considering a measure of *distance or divergence* between statistical populations in terms of the measure of information.

Let f(x) and g(x) be two probability density functions and it is usually defined as (Kullback and Leibler [24], Cover and Thomas [25]),

$$KL\left[f\|g\right] = E_f\left[\log\frac{f(x)}{g(x)}\right] = \int f(x)\log\frac{f(x)}{g(x)}dx \qquad (12a)$$

$$KL\left[g\|f\right] = E_g\left[\log\frac{g(x)}{f(x)}\right] = \int g(x)\log\frac{g(x)}{f(x)}dx \qquad (12b)$$

It is well known that KL[f||g] ≠ KL[g||f] and KL[f||g] ≥ 0 and equality holds if and only if f=g (Burnham and Anderson [26]). The smaller KL[f||g] means that "**f**" is preferred and large values of KLD favor "**g**." KLD is sometimes called the *information gain* by *X* if 'f' can be used instead of 'g'. It is also called the *relative entropy* for using *g* instead of 'f'. *KL* [g||f] measures how easy it is to tell apart the two probability distributions (Henson & Douglas [27]). Here, a brief review about the theoretical developments and practical implications of KLD is highlighted.

Dumonceaux and Antle [28], Kundu and Manglick [29] and Pascual [30] worked on the problem of testing whether some given observations follow one of the two possible distributions. Further, the idea has been extended to discriminate between Gamma and Weibull distributions (Bain and Englehardt [31], Fearn and Nebenzahl [32] and Mohd Saat *et al*. [33]), between Gamma and Log-Normal distributions (Kundu and Manglick [34]). Arizono and Ohta [35] used order statistics and KLD to test for the normality of a distribution based on sampling. Clarke [36] applied KLD for stochastic complexity and sample size calculation. Song [37] made use of order statistics and KLD and proposed a new nonparametric based goodness of fit test. Volkau *et al*. [38], Cabella *et al*. [39] promoted KLD as an application tool for analyzing the magnetic resonance images and also to compare the performance of two separate tests respectively. Hughes and Bhattacharya [40] characterized the symmetry properties of Bi-Normal and Bi-Gamma ROC curves in terms of KLD between two probability distributions which are of cases and controls.

Further, the major objective is to make use of a test statistic that should summarize the whole of the relevant information supplied by the sample, namely Kullback-Leibler Divergence (KLD) to discriminate between one of two possible distributions. Because, the KLD will produce better information since it incorporates information contained in both the populations (distributions). The probability density

functions of Half Normal and Exponential distributions respectively are

$$f(x) = \frac{\sqrt{2}}{\sigma_H\sqrt{\pi}} e^{-\frac{x^2}{2\sigma_H^2}}$$

$$g(x) = \frac{1}{\sigma_D} e^{-\frac{x}{\sigma_D}}$$

The KLD expressions using above density functions are as follows

$$KL\left[f\|g\right] = \ln\left[\frac{\sigma_D}{\sigma_H}\sqrt{\frac{2}{\pi}}\right] - \frac{1}{2} + \frac{\sigma_H}{\sigma_D}\sqrt{\frac{2}{\pi}} \quad (13)$$

$$KL\left[g\|f\right] = \frac{\sigma_D^2}{\sigma_H^2} - 1 + \ln\left[\frac{\sigma_H}{\sigma_D}\sqrt{\frac{\pi}{2}}\right] \quad (14)$$

5. SIMULATION STUDIES

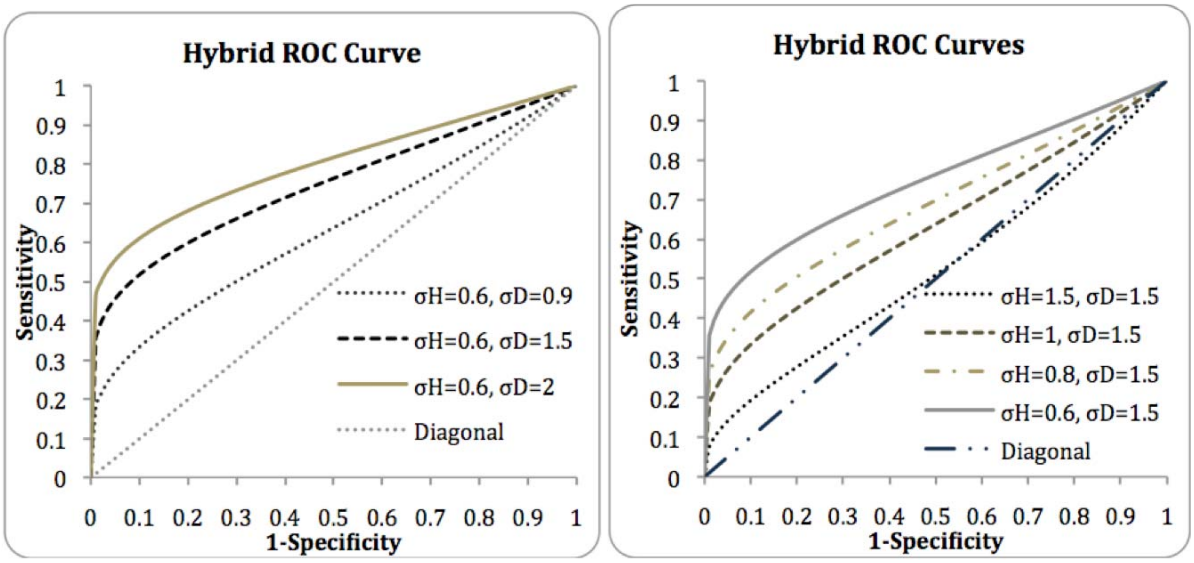
Simulation studies were conducted to illustrate the behavior of the Hybrid ROC (HROC) Curve. Samples of sizes {50, 100, 200, and 600} were generated at various combinations of σ_H = *{0.6, 0.8, 1, 1.5}* and σ_D = *{0.9, 1.5, 2}* from their respective densities. To observe the variation in the HROC Curve, two different cases were considered, (i) $\sigma_D = \sigma_H$, where the scale values of diseased and healthy populations will be equal. (ii) $\sigma_D > \sigma_H$, assuming that the scale parameter of diseased population will have greater variation than that of the healthy population.

Table **1** reports the AUC values for two different cases of scale parameter along with its KLD values. An interesting fact was observed that AUC remain same for all sample sizes at various combinations of σ_D *&* σ_H. On considering the first case of the simulation study i.e. $\sigma_D = \sigma_H$, the AUC obtained was almost equal to the chance line. As we increase the variability in diseased population by fixing the scale value of healthy population, it is observed that there is a gradual increase in the AUC value. Similar kind of experiment was carried out by fixing the scale parameter of diseased population and varying the values from higher to lower of healthy population. Under this also, the AUC was increasing gradually, starting its value from 0.5 and more. Further, the scale parameters of both populations were made equal to show that AUC is equal to 0.5. The above explained phenomenon is visualized in the form of smooth Hybrid ROC Curves.

In Table **1**, the value of KLD reveals the fact that the proximity between two probability distributions will be closer if the KLD value is closer to zero. More over increasing values of KLD imply that the distance or divergence between two density curves or distributions also increases. If $\sigma_D = \sigma_H$, the KLD value is 0.0717, this means that the two distributions are closer to each other and more over the AUC value is observed nearer to 0.5. In the context of ROC if AUC attains 0.5 then the density curves of two distributions will get overlapped and as the AUC value increases, the discrepancy between two densities also increases. Considering the first experiment in Table **1**, the KL [g||f] takes value 1.0706, which means that the two distributions are far away from each other with 1.0706 units of bits. Basing on the critical values of the reference (Exponential) distribution, it can be inferred that the samples tend to lie more from diseased (exponential population) rather than the healthy (half normal population). Hence, a consideration can be made that most of the samples are from exponential distribution. In the ROC phenomenon, the same meaning can be expressed in terms of AUC, i.e., higher values of AUC indicates that the samples or individuals

Table 1: Results of Hybrid ROC (HROC) Curve

Experiment No.	Half-normal Distribution	Exponential Distribution	AUC	KL[f‖g]	KL[g‖f]
	σ_H	σ_D			
1	0.6	0.9	0.6251	0.8761	1.0706
2	0.6	1.5	0.7351	2.1846	4.5596
3	0.6	2	0.7844	3.1371	9.1332
4	1.5	1.5	0.5210	0.0717	0.2260
5	1	1.5	0.6251	0.8761	1.1119
6	0.8	1.5	0.6766	1.3984	2.1130
7	0.6	1.5	0.7351	2.1846	4.5596



a **b**

Figure 2: a: Plot of HROC curves at different combinations of σ_D and σ_H when $\sigma_H = 0.6$.
b: Plot of HROC curves at different combinations of σ_D and σ_H when $\sigma_D = 0.6$.

are being classified correctly with less percentage of misclassification.

Figures **2a** & **b** depict the first case and the latter case of the phenomenon. These curves will deviate from chance line as the variability between the scales of both diseased and healthy populations becomes larger, otherwise. On observing the shapes of the Hybrid ROC Curves in Figures **2a** & **b**, it is clear that the curve is monotonically increasing.

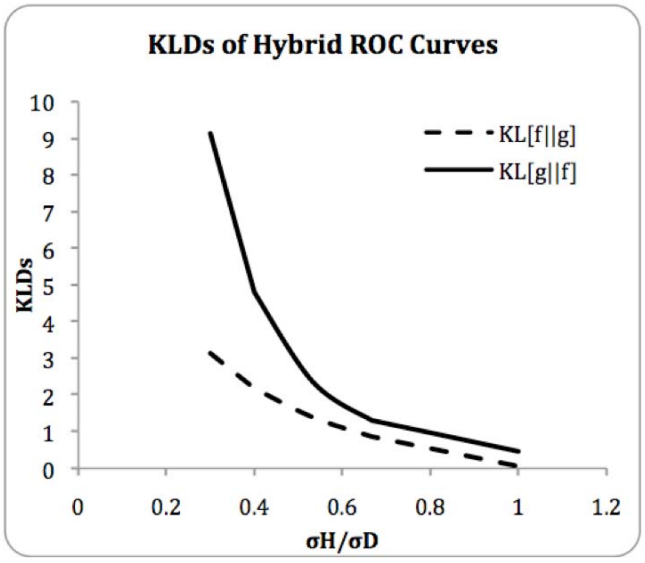


Figure 3: Plot of Kullback - Leibler Divergence Measures.

Figure **3** shows the Kullback Leibler divergences KL[f||g] (the dashed line) and KL[g||f] (the solid line) for two densities of Half Normal and Exponential. Whenever $\sigma_D > \sigma_H$, then KL[g||f] > KL[f||g]. It is clear that as the ratio increases the KL[g||f] attains Exponentiality and this indicates that the samples will follow exponential distribution, which is the criteria of interest.

6. CONCLUSIONS

The present work focused on proposing a new form of ROC model, by assuming that the healthy and diseased populations follow Half-Normal and Exponential distributions respectively. Further, the AUC expression, basic properties of proposed model were derived and the functional behavior of the proposed model namely Hybrid ROC Curve was studied by considering equality in scale values of two distributions as well as varying them in both the distributions. The fact we observed that the scale parameter influences in showing the variant forms of the Hybrid ROC Curve. Simulation studies were conducted to highlight the typical forms of Hybrid ROC Curve. It is proved that the proposed hybrid ROC curve is a monotonically increasing function, slope is a function of likelihood ratio and it is invariant under strictly increasing transformations. Another attempt made in this paper is to show that how the Kullback Leibler Divergence measure can also be used for classification. The proximity between the two distributions will depend upon the ratio of scale parameters of both exponential and half normal distributions. Figure **3** depicts the information that the samples generated over different values of σ_H and σ_D. have the pattern of Exponentiality.

Thus, KLD can also be used as a divergence measure in the context of ROC for binary classification problems.

REFERENCES

[1] Green DM, Swets JA. Signal Detection theory and Psychophysics. Wiley, New York 1966.

[2] Ogilive and Creelman. Maximum Likelihood Estimation of Receiver Operating Characteristic Curve Parameters. Journal of Mathematical Psychology 1968; 5: 377-391. http://dx.doi.org/10.1016/0022-2496(68)90083-7

[3] Dorfman and Alf. Maximum Likelihood Estimation of parameters of signal detection theory-a direct solution. Psychometrika 1968; 33: 117-124. http://dx.doi.org/10.1007/BF02289677

[4] Dorfman and Alf. Maximum-Likelihood Estimation of parameters of signal detection theory and determination of confidence interval-rating method data. Journal of Mathematical Psychology 1969; 6: 487-496. http://dx.doi.org/10.1016/0022-2496(69)90019-4

[5] Lusted LB. Signal detectability and medical decision making. Science 1971; 171: 1217-1219.

[6] Bamber D. The area above the ordinal dominance graph and the area below the receiver operating characteristic graph. Journal of Mathematical Psychology 1975; 12: 387-415. http://dx.doi.org/10.1016/0022-2496(75)90001-2

[7] Egan. Signal Detection Theory and ROC analysis. New York, Academic Press 1975.

[8] Metz CE. Basic Principles of ROC analysis. Seminars in Nuclear Medicine 1978; 8: 283-298. http://dx.doi.org/10.1016/S0001-2998(78)80014-2

[9] Swets JA, *et al*. Assessment of Diagnostic Technologies. Science 1979; 205: 753-759. http://dx.doi.org/10.1126/science.462188

[10] Hanley JA, Mc Neil BJ. A Meaning and Use of the area under a Receiver Operating Characteristics (ROC) Curves. Radiology 1982; 143: 29-36. http://dx.doi.org/10.1148/radiology.143.1.7063747

[11] Hanley JA, Mc Neil BJ. A method of Comparing the Areas Under Receiver Operating Characteristics Analysis derived from the same cases. Radiology 1983; 148: 839-843. http://dx.doi.org/10.1148/radiology.148.3.6878708

[12] Hanley JA. The Robustness of the Binormal Assumption used in fitting ROC curves. Medical Decision Making 1988; 8: 197-203. http://dx.doi.org/10.1177/0272989X8800800308

[13] Goddard MJ, Hindberg I. Receiver operating characteristic (ROC) curves and non-normal data: An empirical study. Statistics in Medicine 1990; 9: 325-337. http://dx.doi.org/10.1002/sim.4780090315

[14] Pepe MS. Three approaches to regression analysis of receiver operating characteristic for continuous test results. Biomterics 1998; 54: 124-135. http://dx.doi.org/10.2307/2534001

[15] Pepe MS. A regression modeling framework for receiver operating characteristic curves in medical diagnostic testing. Biometrika 1997; 84: 595-608. http://dx.doi.org/10.1093/biomet/84.3.595

[16] Pepe MS. Interpretation, estimation and regression for ROC curves. Biometrics 2000; 56: 352-359. http://dx.doi.org/10.1111/j.0006-341X.2000.00352.x

[17] Alonzo TA, Pepe MS. Distribution free ROC analysis using binary regression techniques. Biostatistics 2002; 3(3): 421-432. http://dx.doi.org/10.1093/biostatistics/3.3.421

[18] Zhang Z, Pepe MS. A Linear Regression Framework for Receiver Operating Characteristic (ROC) Curve Analysis. UW Biostatistics Working Paper Series, Paper 253, 2005, http://bepress.com/uwbiostat/paper253

[19] Krzanowski WJ, Hand DJ. ROC curves for continuous data, Monographs on Statistics and Applied Probability. CRC Press, Taylor and Francis Group; NY 2009.

[20] Vardhan RV, Sarma KVS. On the Relationship between the Odds Ratio and the Area under the ROC Curve in the context of Logistic Regression for Comparing Several Biomarkers. International Journal of Statistics and Systems 2010; 5: 165-172.

[21] Vardhan RV, Sarma KVS. Estimation of the Area under the ROC curve using Confidence Intervals of Mean. ANU Journal of Physical Sciences 2010; 2(1): 29-39.

[22] Farragi and Benjamin Raiser. Estimation of the Area under the ROC Curve, Statistics in Medicine 2002; 21: 3093-3106. http://dx.doi.org/10.1002/sim.1228

[23] Fisher RA. On the Mathematical Foundations of Theoretical Statistics. Philosophical Transactions of the Royal Society A 1921; 222: 309-368. http://dx.doi.org/10.1098/rsta.1922.0009

[24] Kullback S, Leibler RA. On Information and Sufficiency. The Annals of Mathematical Statistics 1951; 22(1): 79-86. http://dx.doi.org/10.1214/aoms/1177729694

[25] Cover T, Thomas J. Elements of Information Theory. John Wiley & Sons, Inc 1991.

[26] Burnham KP, Anderson DR. Model selection and multi model inference: a practical information-theoretic approach. 2nd eds., Springer, New York 2002.

[27] Henson R, Douglas J. Test construction for cognitive diagnosis. Applied Psychological Measurement 2005; 29(4): 262-277. http://dx.doi.org/10.1177/0146621604272623

[28] Dumonceaux R, Antle CE. Discrimination between the log normal and weibull distributions. Technometrics 1973; 15(4): 923-926. http://dx.doi.org/10.1080/00401706.1973.10489124

[29] Kundu, D. and Manglick, A. Discriminating between the weibull and log normal distributions. Naval Research Logistics 2004; 51: 893-905. http://dx.doi.org/10.1002/nav.20029

[30] Pascual FG. Maximum likelihood estimation under misspecified Log-Normal and Weibull distributions. Communications in Statistics-Simulation and Computations 2005; 34: 503-524. http://dx.doi.org/10.1081/SAC-200068380

[31] Bain LJ, Englehardt M. Probability of correct selection of Weibull versus gamma based on likelihood ratio. Communications in Statistics Ser A 1980; 9: 375-381.

[32] Fearn DH, Nebenzahl E. On the maximum likelihood ratio method of deciding between the Weibull and gamma distributions. Communications in Statistics Ser A 1991; 20(2): 579-593. http://dx.doi.org/10.1080/03610929108830516

[33] Mohd Saat NZ, Jemain AA, Al-Mashoor SH. A Comparison of Normal and generalized exponential distributions. Journal of Statistical Planning and Inference 2008; 127: 213-227.

[34] Kundu D, Manglick A. Discriminating between the log normal and gamma distributions. Journal of the Applied Statistical Sciences 2005; 14: 175-187.

[35] Arizono I, Ohta H. A test for normality based on Kullback Leibler information. The American Statistician 1989; 43: 20-22.

[36] Clarke B. Asymptotic normality of the post error in relative entropy. IEEE Transactions on Information Theory 1999; 45: 165-176. http://dx.doi.org/10.1109/18.746784

[37] Song KS. Goodness of fit tests based on the Kullback-Leibler discrimination information. IEEE Transactions on Information Theory 2002; 48: 1103-1117. http://dx.doi.org/10.1109/18.995548

[38] Volkau I, Bhanu Prakash KN, Anand A, Aziz A, Nowinski W. L. Extraction of the midsagittal plane from morphological neuroimages using the Kullback-Leibler's measure. Medical Image Analysis 2006; 10: 863-874. http://dx.doi.org/10.1016/j.media.2006.07.005

[39] Cabella BCT, Sturzbecher MJ, Tedeschi W, Filho OB, de raujo DB, Neves UPC. A numerical study of the Kullback-Leibler distance in functional magnetic resonance imaging. Brazilian Journal of Physics 2008; 38: 20-25. http://dx.doi.org/10.1590/S0103-97332008000100005

[40] Hughes, Bhaskar Bhattacharya. Symmetry Properties of Bi-Normal and Bi-Gamma Receiver Operating Characteristic Curves are Described by Kullback-Leibler Divergences. Entropy 2013; 15: 1342-1356. http://dx.doi.org/10.3390/e15041342

Received on 25-10-2014 Accepted on 16-01-2015 Published on 27-01-2015

http://dx.doi.org/10.6000/1929-6029.2015.04.01.11