International Journal of Advanced Computer and Mathematical Sciences
ISSN 2230-9624. Vol4, Issue4, 2013, pp285-294
http://bipublication.com

ESTIMATION OF AUC OF THE BINORMAL ROC CURVE USING CONFIDENCE INTERVALS OF VARIANCES

***Suresh Babu N. and Sarma K.V.S.**

Department of Statistics, S.V.University, Tirupati – 517 502, India,
*sureshbabu.nellore@gmail.com, sarma_kvs@rediffmail.com

[Received-11/11/2013, Accepted-01/12/2013]

ABSTRACT

In medical diagnosis, testing of the performance of a particular diagnostic test is a common and good practice, after classifying the subjects in to different groups by various classification techniques, in particularly binary classification. Assessment of the performance of a diagnostic test can be achieved by Area under the Receiver Operating Characteristic (AUROC) Curve, simply denoted by AUC. For Diseased (D) and Healthy (H) normal populations, Binormal model gives a closed form expression to the Area under the Curve (AUC). In this paper, we proposed a new method in the estimation of Area under the Binormal ROC curve using confidence intervals for all possible combinations of the combined estimates of pooled variance, obtained from the chi-square test and taking weighted average of possible AUC values with different patterns of weights. Kolmogorov–Smirnov (K-S) test has been used to test the normality of the possible AUCs. Numerical illustrations for the proposed method are given with simulated data.

Key words: ROC, AUROC, AUC, Binormal model, Sensitivity, Specificity

Introduction

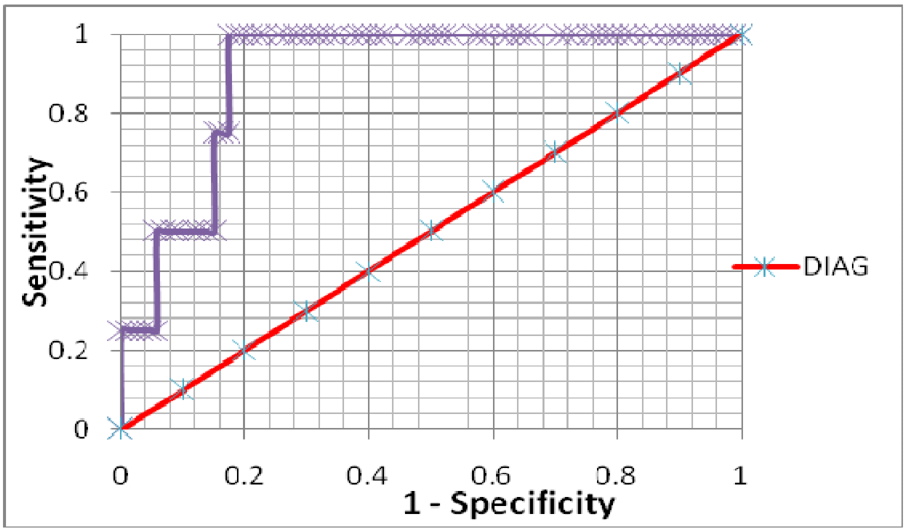
A generalized assessment of the performance of binary classifiers is typically carried out by a tool called Receiver Operating Characteristic (ROC) curve. It is a graphical relationship between False Positive Rate (FPR) verses True Positive Rate (TPR). The term Receiver Operating Characteristic (ROC) curve had its origin from statistical Decision theory as well as signal detection theory (SDT) and was used during world war-II for the analysis of radar images. It is an effective method of evaluating the quality or performance of a diagnostic test and is widely used in medical related fields like radiology, epidemiology etc to evaluate the performance of many medical related tests.

The phrase ROC, is tradeoff between sensitivity (TPR) on Y-axis and 1-specificity (FPR) on X-axis. The Area under the ROC (AUROC) curve lies between 0 to 1, which enables us performance of a diagnostic test accuracy especially in all medical diagnosis tests. Sensitivity and specificity are the most popularly used measurements of the performance accuracy, can be evaluated basing on the four possible states TP, FP, TN and FN from a *confusion* matrix [11]. With all four possible states, sensitivity and specificity are calculated as

$$Sensitivity = S_n = \frac{\text{Number of True Positives}}{\text{Number of True Positives+Number of False negatives}} = \frac{TP}{TP+FN}$$

$$Specificity = S_p = \frac{\text{Number of True nagatives}}{\text{Number of True nagatives+Number of False positives}} = \frac{TN}{TN+FP}$$

Both sensitivity and specificity are used as performance indicators of the area under the receiver operating characteristic curve. An illustration of the empirical ROC curve is shown in the following figure.



An empirical ROC curve

If sensitivity will increase then 1-specificity will decrease and vice-versa. In the above diagram diagonal line represents discrimination of the area under the ROC curve. Suppose roc curve falls above the diagonal line may discriminate about a diagnostic test otherwise no discrimination is described. So the discrimination can be explained by the area under the ROC curve. One of the very useful consequence of binormal ROC model is that its AUC can be derived very easily, and has a very simple form [3, 7, 14].

For the binormal case, [1] have used a test to compare the sensitivities (true positive rates) of two classifiers for a given common level of specificity (1-false positive rate). A method of comparing the Areas under Receiver Operating Characteristics Analysis derived from the same cases [3].

In medical diagnosis problems, one may be interested to know the performance of a diagnostic test accuracy (or classifier performance) . more generally, for a binary scale data, to assess the diagnostic test accuracy by sensitivity (S_n), specificity (S_p), accuracy, area under the ROC curve (AUC) etc,. Usually, all performance evaluation measures are the estimates (or proportions). Naturally estimates do not possess (represent) any characteristics of the population parameters. In this scenario we don't know the true value of the parameter. This problem arises the necessity of the confidence interval (CI). This confidence interval gives lower and upper limits basing on an estimated value at certain level of significance (α %). So the estimated value certainly lies within these two confidence limits. Basing on the confidence intervals [16, 17] developed a new method of approach on the comparison of Binormal ROC curves using confidence interval of means.

Auc of The Binormal ROC Model

In the binormal model two normal populations 'X' (Positive) and 'Y' (Negative) have the means μ_x and μ_y with standard deviations σ_x and σ_y from the two normal populations viz., Disease (D) and Healthy (H),such that $X \sim N(\mu_x, \sigma^2_x)$ and $Y \sim N(\mu_y, \sigma^2_y)$. Let '*C*' be the "cut off" or "optimal cut off" and if P (X>C) then a variable is classified into D group otherwise into H group.

In the parametric approach, the binormal ROC model is given by

$$\text{ROC}(c) = a + b\phi^{-1}(c) \qquad (1)$$

Where a and b are the parameters which can be estimated by using the method of Maximum Likely hood then $a = \left(\frac{\mu x - \mu y}{\sigma x}\right)$ and $b = \left(\frac{\sigma x}{\sigma y}\right)$. In the binormal model, the estimation of AUC given by the Faraggi and Reiser (2002) is

$$\text{AUC} = \Phi\left(\frac{\mu_X - \mu_y}{\sqrt{\sigma_X^2 + \sigma_Y^2}}\right) \qquad (2)$$

where σ^2_x and σ^2_y are the variances of the two groups D and H respectively.

For instance one wishes to get more appropriate AUC which is predetermined say AUC* like 0.9, 0.85, 0.8 etc., and Z* can be defined as

$$Z^* = \Phi^{-1}(AUC^*) \qquad (3)$$

be the standard normal deviate corresponding to AUC*. The Z* is used in the estimation of the group D mean i.e. $\hat{\mu}_X$ based on the assumed group H mean i.e. $\hat{\mu}_Y$. Faraggi and Reiser (2002) have shown that the estimated mean of the group D is

$$\hat{\mu}_X = \hat{\mu}_y + Z^*\sqrt{s_X^2 + s_Y^2}. \qquad (4)$$

with the estimated $\hat{\mu}_X$ and $\hat{\mu}_Y$, AUC can be estimated.. Now the equation (2) will become

$$\widehat{\text{AUC}} = \Phi\left(\frac{\hat{\mu}_x - \hat{\mu}_y}{\sqrt{s_X^2 + s_Y^2}}\right) \qquad (5)$$

Suresh Babu et al., (2011) have developed a new method of estimation of AUC of the binormal ROC model through interval estimation of the differences between the means while sample variances can be considered as point estimates. In this paper we derive a new approach to the estimation of AUC using confidence intervals for variances. So as we derive 100(1-α) % confidence interval basing on combined sample variances from the upper and lower limits of the D and H normal populations. And, now the confidence interval of the pooled variance (σ^2) can be obtained by χ^2- test separately for two populations as

$$\frac{(n_1-1)S_x^2}{\chi^2_{\alpha/2,n_1-1}} \le \sigma^2 \le \frac{(n_1-1)S_x^2}{\chi^2_{1-\alpha/2,n_1-1}} \qquad (\textit{for Diseased population})$$

$$\frac{(n_2-1)S_y^2}{\chi^2_{\alpha/2,n_2-1}} \le \sigma^2 \le \frac{(n_2-1)S_y^2}{\chi^2_{1-\alpha/2,n_2-1}} \qquad (\textit{for Healthy population})$$

Possible Combinations Of Combined Estimates Of Pooled Variance

To obtain various possible combinations of the pooled variance, we define the lower and upper limits for D and H populations as given below.

Define $L_1 = \frac{(n_2-1)S_y^2}{\chi^2_{\alpha/2,n_2-1}}$, $U_1 = \frac{(n_2-1)S_y^2}{\chi^2_{1-\alpha/2,n_2-1}}$ and

$$L_2 = \frac{(n_1-1)S_x^2}{\chi^2_{\alpha/2,n_1-1}} \quad , \quad U_2 = \frac{(n_1-1)S_x^2}{\chi^2_{1-\alpha/2,n_1-1}}$$

Then the pooled variance is obtained by considering 9 different combinations as shown in the below Table 1.

Combination (i)	δ (combined estimate)	Pooled SD
1	$L_1 + L_2$	$\delta_1 = \sqrt{L1 + L2}$
2	$L_1 + S_x^2$	$\delta_2 = \sqrt{L1 + S_x^2}$
3	$L_1 + U_2$	$\delta_3 = \sqrt{L1 + U2}$
4	$S_y^2 + L_2$	$\delta_4 = \sqrt{S_y^2 + L2}$
5	$S_y^2 + S_x^2$	$\delta_5 = \sqrt{S_y^2 + S_x^2}$
6	$S_y^2 + U_2$	$\delta_6 = \sqrt{S_y^2 + U2}$
7	$U_1 + L_2$	$\delta_7 = \sqrt{U1 + L2}$
8	$U_1 + S_x^2$	$\delta_8 = \sqrt{U1 + S_x^2}$
9	$U_1 + U_2$	$\delta_9 = \sqrt{U1 + U2}$

Table-1: Different combinations of combined estimates

Since the location parameters are replaced by their point estimates, the numerator on the right hand side of (5) will be fixed as $\Delta = (\hat{\mu}_x - \hat{\mu}_y)$ for each of the 9 combinations. The different pooled standard deviations are considered to be the 9 combinations of the combined estimates i.e. δ_1, δ_2 ……… δ_9. If we denote by A_j the AUC obtained by the j^{th} pooled variance, the resulting AUCs will be calculated as shown in Table 2.

A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9
$\frac{\Delta_1}{\delta_1}$	$\frac{\Delta_2}{\delta_2}$	$\frac{\Delta_3}{\delta_3}$	$\frac{\Delta_4}{\delta_4}$	$\frac{\Delta_5}{\delta_5}$	$\frac{\Delta_6}{\delta_6}$	$\frac{\Delta_7}{\delta_7}$	$\frac{\Delta_8}{\delta_8}$	$\frac{\Delta_9}{\delta_9}$

Table-2 : List of possible AUCs using CIs of Variances

The expression for A_j is given by

$A_j = \Phi\left(\frac{\Delta}{\delta_j}\right)$ for j = 1,2,…,9

Lasko et al (2005) have shown that the estimated variance of the AUC is given by

$$\hat{V}(A_j) = \frac{1}{n_1+n_2}\left[\frac{1}{s_X+s_Y}\left(\frac{s_X^2}{n_1}+\frac{s_Y^2}{n_2}\right)+\frac{A_j^2}{2\left(s_X^2+s_Y^2\right)^2}\left(\frac{s_X^4}{n_1-1}+\frac{s_Y^2}{n_2-1}\right)\right]$$

After finding the 9 different estimates, we derive three new weighted averages with different weights as discussed below.

i) Simple Average Method

Among three methods, it is the simplest one. The estimation of new AUC through this method is basing on weights. For this method, the weights are given as

$$W_j = 1/9 \quad \forall\, j = 1, 2, \ldots, 9$$

Then the estimated AUC is the sum of the product of the weights and the corresponding AUCs and is given by

$$AUC_{AVG} = \sum_{j=1}^{9} W_j A_j$$

where A_j for j = 1, 2, ….., 9 are the AUCs obtained from the confidence intervals of means given in (3.5). It can be seen that $\sum_{j=1}^{9} W_j$

The variance of estimated AUC is computed as below

$$V(AUC_{AVG}) = \sum_{j=1}^{9} W_j^2 V(A_j)$$

ii) Fixed Weights Method (FW-Method)

In this method, among all weights, one is to be fixed as $W_5 = 0.5$ because Δ_5 occurs with high probability and the remaining 8 weights are to be calculated by the following formula

$$W_j = \frac{0.5}{8} \quad \forall\, j \neq 5$$

The new estimator by the fixed weights method (FW-Method) is the sum of the product of the weights including W_5 with their corresponding AUCs and is defined as below

$$AUC_{FW} = \sum_{j\neq 1}^{9} W_j A_j + \frac{0.5}{8} A_5$$

where A_j ; j = 1, 2, ….., 9 are s are the different AUCs obtained from the confidence intervals of means. It can be seen that $\sum_{j=1}^{9} W_j = 1$. For fixed weights method (FW-Method), the estimation of the variance is

similar to that of the computation of the variance in the simple average method.
so, the variance of the estimated AUC i.e. $V(AUC_{FW})$ is the given by

$$V\,(AUC_{FW}) = \sum_{j=1}^{9} W_j^2 V(A_j)$$

iii) Proportional Weights Method (PW-Method)

This method differs from both the simple average method and fixed weight method. In this method the weights are taken as proportional to the difference in the means used in (3.5) such that

$$W_j = \frac{1}{\Delta_j} \ \forall\, j = 1, 2, \ldots, 9$$

It means the weights are increase by proportional to Δ. The new estimator by propositional weight method (PW-Method) is defined by the following formula

$$AUC_{PW} = \frac{\sum_{j=1}^{9} W_j\, A_j}{\sum_{j=1}^{9} W_j}$$

where Aj for j = 1, 2, ….., 9 are the different AUCs obtained from the confidence intervals of means.
The variance of the estimated new estimator of the AUC by propositional weights method will be $V\,(AUC_{PW}) = \sum_{j=1}^{9} r_j^2 V(A_j)$

where $r_j = W_j \{\sum_{j=1}^{9} W_j\}^{-1}$

ESTIMATION OF AUC FROM THE SIMULATED DATA

Illustration 1: Fixing the required AUC as 0.90 we have estimated μ_x by using (4). By changing the values of n_1,n_2, Sx, Sy and μ_y we have calculated the AUC by using (5) with the help of the Excel function NORMSDIST( ). The resulting 9 AUC estimates are tested from normality using K-S test and the location and scale parameters $\hat{a}$ and $\hat{b}$ are estimated. The numerical results and the testing of assumption of the normality by K-S test are shown in Table 3.

Target AUC = 0.90									
Trail : 1	n_1 = 5, n_2 = 5, μ_x = 103.12, S_x = 10, S_y = 10								
A_i	0.98	0.94	0.73	0.94	0.90	0.72	0.73	0.72	0.67
$V(A_i)$	0.2060	0.2055	0.2033	0.2055	0.2051	0.2033	0.2033	0.2033	0.2028
$AUC_{AVG\text{-}S}$ = 0.8164 ±0.4518 Bias = 0.0836			$AUC_{FW\text{-}S}$ = 0.8530 ±0.4523 Bias = 0.0471				$AUC_{PW\text{-}S}$=0.8771±0.4526 Bias = 0.0229		
K-S test for Normality			Z = 0.933, p = 0.349				Normality accepted		

Trail : 2	$n_1 = 10, n_2 = 15, \mu_x = 103.12, S_x = 10, S_y = 10$								
A_i	0.96	0.93	0.82	0.93	0.90	0.81	0.85	0.83	0.77
$V(A_i)$	0.0342	0.0341	0.0339	0.0341	0.0341	0.0339	0.0340	0.0340	0.0339
AUC_{AVG-S} = 0.8685 ±0.1844 Bias = 0.0315			AUC_{FW-S} = 0.8823±0.1845 Bias = 0.0177				AUC_{PW-S}=0.8855±0.1845 Bias = 0.0145		
K-S test for Normality			Z = 0.501, p = 0.963				Normality accepted		

Trail : 3	$n_1 = 20, n_2 = 30, \mu_x = 103.12, S_x = 10, S_y = 10$								
A_i	0.95	0.92	0.86	0.93	0.90	0.85	0.88	0.86	0.82
$V(A_i)$	0.0085	0.0085	0.0085	0.0085	0.0085	0.0085	0.0085	0.0085	0.0085
AUC_{AVG-S} =0.8851±0.0922 Bias = 0.0149			AUC_{FW-S} = 0.8916±0.0922 Bias = 0.0084				AUC_{PW-S} =0.8922±0.0922 Bias = 0.0078		
K-S test for Normality			Z = 0.512, p = 0.955				Normality accepted		

Trail : 4	$n_1 = 50, n_2 = 50, \mu_x = 108.10, S_x = 15, S_y = 10$								
A_i	0.94	0.91	0.87	0.93	0.90	0.86	0.90	0.88	0.85
$V(A_i)$	0.0027	0.0026	0.0026	0.0027	0.0026	0.0026	0.0026	0.0026	0.0026
AUC_{AVG-S} =0.8934±0.0514 Bias = 0.0066			AUC_{FW-S} =0.8963±0.0514 Bias = 0.0037				AUC_{PW-S}=0.8969±.0.055 Bias = 0.0031		
K-S test for Normality			Z = 0.424, p = 0.994				Normality accepted		

Trail : 5	$n_1 = 50, n_2 = 100, \mu_x = 119.51, S_x = 25, S_y = 10$								
A_i	0.94	0.90	0.86	0.93	0.90	0.85	0.93	0.89	0.85
$V(A_i)$	0.0026	0.0026	0.0026	0.0026	0.0026	0.0026	0.0026	0.0026	0.0026
AUC_{AVG-S} =0.8948±0.0511 Bias = 0.0052			AUC_{FW-S} = 0.8971± 0.0511 Bias = 0.0029				AUC_{PW-S}=0.8993±0.0511 Bias = 0.0029		
K-S test for Normality			Z = 0.535 , p = 0.937				Normality accepted		

Table 3: Estimated AUCs with target value 0.90

Illustration 2: The experiment is repeated 5 times with target AUC = 0.80 with different combination of input parameters and the results are shown in Table 4.

Target AUC = 0.80

Trail : 1	$n_1 = 5, n_2 = 5, \mu_x = 96.90, S_x = 10, S_y = 10$								
A_i	0.92	0.85	0.66	0.85	0.80	0.65	0.66	0.65	0.62
$V(A_i)$	0.2053	0.2045	0.2027	0.2045	0.2040	0.2027	0.2027	0.2027	0.2024
AUC_{AVG-S} = 0.7386 ± 0.4510 Bias =0.0614			AUC_{FW-S} = 0.7654 ± 0.4513 Bias = 0.0346				AUC_{PW-S} = 0.7964 ± 0.4516 Bias = 0.0036		
K-S test for Normality			Z = 0.944, p = 0.335				Normality accepted		

Trail : 2	n_1 = 10, n_2 = 15, μ_x = 96.90, S_x = 10, S_y = 10								
A_i	0.88	0.83	0.73	0.84	0.80	0.72	0.76	0.74	0.69
$V(A_i)$	0.0340	0.0340	0.0338	0.0340	0.0339	0.0338	0.0339	0.0338	0.0338
AUC_{AVG-S} = 0.7752 ± 0.1841 Bias = 0.0248			AUC_{FW-S} =0.7860 ± 0.1841 Bias = 0.0140				AUC_{PW-S} = 0.7923 ± 0.1841 Bias = 0.0077		
K-S test for Normality			Z = 0.485, p = 0.973				Normality accepted		

Trail : 3	n_1 = 20, n_2 = 30, μ_x = 96.90, S_x = 10, S_y = 10								
A_i	0.86	0.82	0.76	0.83	0.80	0.75	0.78	0.76	0.73
$V(A_i)$	0.0085	0.0085	0.0085	0.0085	0.0085	0.0085	0.0085	0.0085	0.0084
AUC_{AVG-S} =0.7879±0.0920 Bias = 0.0121			AUC_{FW-S} =0.7932±0.0920 Bias = 0.0068				AUC_{PW-S}=0.7953±0.0920 Bias = 0.0047		
K-S test for Normality			Z = 0.561, p = 0.911				Normality accepted		

Trail : 4	n_1 = 50, n_2 = 50, μ_x = 100.17, S_x = 15, S_y = 10								
A_i	0.84	0.81	0.77	0.83	0.80	0.76	0.80	0.78	0.75
$V(A_i)$	0.0026	0.0026	0.0026	0.0026	0.0026	0.0026	0.0026	0.0026	0.0026
AUC_{AVG-S} =0.7948±0.0514 Bias = 0.0052			AUC_{FW-S} = 0.7971±0.0514 Bias = 0.0029				AUC_{PW-S}=0.7984±0.0514 Bias = 0.0016		
K-S test for Normality			Z = 0.424, p = 0.994				Normality accepted		

Trail : 5	n_1 = 50, n_2 = 100, μ_x = 107.66, S_x = 25, S_y = 10								
A_i	0.84	0.80	0.76	0.84	0.80	0.76	0.83	0.79	0.75
$V(A_i)$	0.0026	0.0026	0.0026	0.0026	0.0026	0.0026	0.0026	0.0026	0.0026
AUC_{AVG-S} = 0.7967±0.0510 Bias = 0.0033			AUC_{FW-S} = 0.7981±0.0510 Bias = 0.0019				AUC_{PW-S}=0.8013±0.0510 Bias = -0.0013		
K-S test for Normality			Z = 0.558, p = 0.915				Normality accepted		

Table 4: Estimated AUCs with target value 0.80

Hypothesis Testing Of New Estimates with Target Auc

All three *summarized estimates* can be compared for their significance of difference by t-test.

Since the 9 AUC estimates can be treated as a sample of binormal AUC's and each of the three estimates is an average it is possible to compare them with AUC^* by using Students t- test.

The null hypothesis is that *there is no significant difference between the target AUC and the estimated AUC.*

For example, the test statistic for comparing AUC_{AVG-S} with AUC* is given by

$$t = \left\{ \frac{AUC_{AVG} - AUC^*}{S/\sqrt{n-1}} \right\}; \quad \text{where } S = \sqrt{V(AUC_{AVG})} \quad \text{and } n-1 = 8.$$

As explained above, the other two summarized statistics are also compared by the same method. In the following section, the calculations are made of three new estimates for different values of n_1, n_2, S_x, and $S_{y.}$ The statistical significance of the difference between the target AUC and the AUC obtained by the new estimators is tested and the results are shown in table 5.

Trial	t-test with $AUC^* = 0.9$			t-test with $AUC^* = 0.8$		
	AUC_{AVG}	AUC_{FW}	AUC_{PW}	AUC_{AVG}	AUC_{FW}	AUC_{PW}
1	0.5233 (0.6149)	0.2941 (0.7762)	0.1432 (0.8897)	0.3853 (0.7101)	0.2166 (0.8340)	0.0225 (0.9826)
2	0.4834 (0.6418)	0.2718 (0.7926)	0.2225 (0.8295)	0.3816 (0.7127)	0.2146 (0.8355)	0.1178 (0.9091)
3	0.4571 (0.6597)	0.2571 (0.8036)	0.2377 (0.8181)	0.3720 (0.7196)	0.2092 (0.8395)	0.1452 (0.8882)
4	0.3629 (0.7261)	0.2041 (0.8434)	0.1704 (0.8689)	0.2871 (0.7814)	0.1615 (0.8757)	0.0870 (0.9328)
5	0.2871 (0.7813)	0.1615 (0.8757)	0.0398 (0.9692)	0.1824 (0.8598)	0.1026 (0.9208)	0.0735 (0.9432)

Table 5: t-test for comparing the summarized AUCs with the Target

(Figures in the brackets indicate the p-value)

CONCLUSIONS

In this paper, we have proposed a new method basing on weights for estimating AUC of a classifier by the binormal ROC model based on confidence intervals of estimated variances obtained from the chi-square test. All the 9 estimated values of the AUCs at different trails are found to follow normal distribution and hence proportional weights method has been observed least absolute bias. However in all the cases the estimated AUC values by the three weighted average methods, do not differ significantly from the target AUC.

REFERENCES

[1] Bradley, (1997), The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recognition, 30, 1145-1159.

[2] Cai, T. and Moskowitz, C.S, (2004), Semi parametric estimation of binormal ROC curve for a continuous diagnostic test, Biostatistics 5(4): 573 - 586.

[3] David Farragi and Benjamin Reiser (2002), Estimation of the area under the ROC curve, Statistic in Medicine, 21: 3093 - 3106.

[4] DeLong, E. R, et al., (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics, 44, 837-845.

[5] Greenhouse and Mantel (1950), The evaluation of diagnostic tests. Biometrics, 6, 399-412

[6] James, A. Hanley, et al., (1982), A Meaning and Use of the area under a Receiver Operating Characteristic (ROC) curves, Radiology, 143: 29 - 36.

[7] James, A. Hanley, et al., (1983), A method of Comparing the Areas Under Receiver Operating Characteristics Analysis derived from the same cases, Radiology, 141: 839 - 843.

[8] Lloyd, C. J, (1998), The Use of smoothed ROC curves to summarize and compare diagnostic systems, Journal of American Statistical Association, 93: 1356 – 1364.

[9] Margaret Pepe, et al., (2008), Estimation and Comparison of Receiver Operating Characteristic Curves , paper 323, UW Biostatistics working paper series, University of Washington. http://www.bepress.com/uwbiostat/paper323.
[10] Metz, C.E, (1978), Basic Principles of ROC analysis, Seminars in Nuclear Medicine, 8: 283 – 298.
[11] Metz, et al., (1998), Maximum Likelihood Estimation of receiver operating characteristic (ROC) curves continuously distributed data, Statistics in Medicine, 17: 1033-1053.
[12] Pepe, M. S, (2000), Receiver operating characteristic methodology, Journal of American Statistical Association, 95:308 – 311.
[13] Sarma, K. V. S, et al., (2010), Estimation of the Area under the ROC Curve using Confidence Intervals of Mean, ANU Journal of Physical Sciences, ISSN: 0954-0954.
[14] Sarma, K. V. S, et al., (2012), Determining the Optimal Cut-point in an ROC Curve, International Journal of Statistics and Analysis, vol-2, pp. 219-225.
[15] Suresh Babu, N, et al., (2011), On the Comparison of Binormal ROC Curves using Confidence Interval of Means, International Journal of Statistics and Analysis, vol-1, pg 213-221.
[16] Lasko, T. A, et al., (2005), The use of receiver operating characteristics curves in biomedical informatics, Journal of Biomedical Informatics, 38: 404 – 415.
[17] Wojtek J. Krzanowski and David J. Hand (2009), ROC Curves for Continuous Data.
[18] Zweig and Campbell (1993), Receiver operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clinical Chemistry, 39, pg 561-577.