Available Online at http://www.recentscientific.com

ISSN: 0976-3031

International Journal of Recent Scientific Research
Vol. 7, Issue, 7, pp. 12330-12336, July, 2016

International Journal of Recent Scientific Research

Research Article

FUNCTIONAL RELATIONSHIP BETWEEN BRIER SCORE AND AREA UNDER THE CONSTANT SHAPE BI-WEIBULL ROC CURVE

Leo Alexander T* and lavanya A

Department of Statistics, Loyola College, Chennai–34, India

ARTICLE INFO

Article History:

Received 20th April, 2016
Received in revised form 29th May, 2016
Accepted 30th June, 2016
Published online 28th July, 2016

Key Words:

AUC, Brier Score, Classification, Probabilistic Judgments, ROC Curve.

ABSTRACT

In classification, the Receiver Operating Characteristics (ROC) curve analysis is one of the most familiar techniques and it will provide accuracy for the extent of correct classification of a test. The conventional way of expressing the true accuracy of test is by using its summary measure Area Under the Curve (AUC) and intrinsic measures Sensitivity and Specificity. Brier Score ($\overline{\mathrm{B}}$) is shown as another summary measure in the context of ROC Curve to make the probabilistic judgments as well as to identify the extent of classification. Further, the Functional relationship between the Brier Score and AUC of ROC Curve is provided using the parameters **b** (ratio of standard deviations of signal and noise) and **α** (priori probability).The influence of slope on ROC curve is highlighted to explain the behavior of Brier Curves and its relationship with AUC. To demonstrate the proposed methodology Simulation Study is conducted at different combinations of scale parameters of both populations. If parameters **b** and **α** are constants, $\overline{\mathrm{B}}$ values in relation to given AUC values if $\overline{\mathrm{B}}$ values monotonically decreases as AUC values increases, and these relationship curves have monotonically decreasing slopes. An illustrative example is also provided to explain the concepts.

INTRODUCTION

In several situations, physicians are required to express probabilistic judgments in numerical terms, and there is some evidence that such judgments have operational meaning to physicians [5]. Thus, it is very important to evaluate these judgments properly. The term "*Classification*" indicates the method of allocating a group of objects or individuals to one of the predefined status of health (with condition/without condition or healthy/diseased or alive/dead).

In classification, the Receiver Operating Characteristics (ROC) curve analysis is one of the most familiar techniques and it will provide accuracy for the extent of correct classification of a test. The conventional ways expressing the true accuracy of test is by using its summary measure Area Under the Curve (AUC) and intrinsic measures Sensitivity and Specificity. Brier Score is shown as another summary measure in the context of ROC Curve to make the probabilistic judgments as well as to identify the extent of classification.

The present work focuses on establishing a functional relationship between Brier Score and AUC of the ROC Curve using the parameters **b** and **α.** The influence of slope ROC Curve is highlighted to explain the behavior of Brier Curves and its relationship with AUC. The proposed work is supported by simulation studies with different combinations of scale parameters of both populations.

Roc methodology

Let x be the test score observed from two populations with (abnormal individuals) and without (normal individuals) condition respectively which follow Constant Shape Bi-Weibull distributions. The density functions of Constant Shape Bi-Weibull distributions are as follows,

$$f(x|n) = \frac{\beta}{\sigma_n} x^{\beta-1} e^{-\left[\frac{x^\beta}{\sigma_n}\right]}. \tag{1}$$

and

*Corresponding author: **Leo Alexander T**
Department of Statistics, Loyola College, Chennai–34, India

$$f(x|s) = \frac{\beta}{\sigma_s} x^{\beta-1} e^{-\left[\frac{x^\beta}{\sigma_s}\right]}. \quad (2)$$

Let S be a continuous biomarker. The probabilistic definitions of the measures of ROC Curve are as follows:

$$\text{Sensitivity}(s_n) = P(S|s) = \int_t^\infty f(x|s)dx,$$

$$1 - \text{Specificity}(s_p) = P(S|n) = \int_t^\infty f(x|n)dx \ .$$

In this context, the (1-Specificity) and Sensitivity can be defined using equations (1) and (2) and are given in equations (3) and (4) respectively,

$$P(S|n) = x(t) = e^{-\left[\frac{t^\beta}{\sigma_n}\right]}. \quad (3)$$

and

$$P(S|s) = y(t) = e^{-\left[\frac{t^\beta}{\sigma_s}\right]}. \quad (4)$$

The ROC Curve is defined as a function of (1-Specificity) with scale parameters of distributions and is given as,

$$ROC(t) = y(t) = x(t)^b, \quad (5)$$

where $t = -[\sigma_n log x(t)]^{\frac{1}{\beta}}$ is the threshold and $b = \frac{\sigma_n}{\sigma_s}$

The accuracy of a diagnostic test can be explained using the Area Under the Curve (AUC) of an ROC Curve.AUC describes the ability of the test to discriminate between abnormal and normal populations. A natural measure of the performance of the classifier producing the curve is AUC. This will range from 0.5 for a random classifier to 1 for a perfect classifier. The AUC is defined as,

$$AUC = \int_0^1 x(t)^b dx(t). \quad (6)$$

The closed form of AUC is as follows

$$AUC = \frac{1}{1+b}. \quad (7)$$

Slope of the ROC Curve

The slope of an ROC Curve can be defined in three ways [1]: first, as the tangent at a particular point on the ROC Curve corresponding to a test value *x*, (That is tangent(*x*)); Second, the slope between origin 0 (That is point (0,0)) and the point on the ROC Curve corresponding to the a test value (That is slope(0-*x*)); and Third, the slope between two points on the ROC Curve corresponding to the test values *x* and *y* (That is slope(*x*-*y*)).

First, the tangent at a point *x* on the ROC Curve, That is, tangent(*x*) corresponds to the likelihood ratio (LR) for a single test value corresponding to that point on the ROC Curve for a continuous test, That is, LR(*x*).Here, the LR is defined as the ratio between the probability of a defined test result given the presence of condition and the probability of the same result given the absence of condition [8]. If a test generates results on a continuous scale, then a likelihood ratio can theoretically be defined for each test value *x* as

$$\text{L(x)} = \frac{\text{Probability density function of continuous test in variable x as f(x|s) in the presence of condition}}{\text{Probability density function of continuous test in variable x as f(x|n) in the absence of condition}}$$

That is $L(x) = \frac{f(x|s)}{f(x|n)}$. (8)

Now, using equations (1) and (2), the slope of ROC Curve which is called as likelihood ratio can be defined as follows

$$\frac{f(x|s)}{f(x|n)} = \frac{dP(S|s)}{dP(S|n)} = bexp\left(\left[\frac{t^\beta}{\sigma_n}\right] - \left[\frac{t^\beta}{\sigma_s}\right]\right). \quad (9)$$

The inverse slope of ROC Curve is defined as

$$\frac{f(x|n)}{f(x|s)} = \frac{dP(S|n)}{dP(S|s)} = \frac{1}{b} exp\left(\left[\frac{t^{\beta}}{\sigma_s}\right] - \left[\frac{t^{\beta}}{\sigma_n}\right]\right). \tag{10}$$

Brier Score

It is a well known evaluation measure for probabilistic classifiers and is proposed by Brier in 1950. In literature, procedures for calibrating classifiers have been proposed in different contexts such as Weather Prediction Tasks [2], Game Theory [3].

Further, Brier Score was considered in the context of signal Detection Theory by assuming that the calibration in the observers probability estimate is perfect and provided the theoretical relationship between Brier Score and Area Under the Binormal ROC Curve [4]. So Functional relationship between Brier Score and Area Under the ROC Curve was discussed in the Context of Skewed Distributions [9].

Now consider a set of M signal detection tasks with αM signal events and (1-α) M non-signal events (0≤α≤1) and αM, (1-α)M and α denote a priori probability of signal events.

Let $y_i = 0$ if the event is non-signal
$y_i = 1$ if the event is signal.

Let p_i denote the observers (or subject's) probability estimate that the i^{th} event will be the signal one, where the subscript i indicates the individual event [6] and the Brier Score ($\bar{B}$) can be defined as [7] using the expression

$$\bar{B} = \frac{1}{M}\sum_{i=1}^{M}(y_i - p_i)^2 , \tag{11}$$

where p_i is a function of x_i and is defined using Bayes theorem as follows,

$$p_i = p(x_i) = Pr(S|x_i) = \frac{\alpha f(x_i|s)}{\alpha f(x_i|s) + (1-\alpha) f(x_i|n)} . \tag{12}$$

Now, we consider the expected value of $(y_i - p_i)^2$ in the equation (11), when the calibration in the observers probability estimate is perfect. In this case, p_i in the equation (11) is obtained by equation (12) from the Bayes theorem. Therefore, expected value of $(y_i - p_i)^2$is given as

$$E[(y_i - p_i)^2] = \int_{-\infty}^{\infty} (1 - p(x))^2 \alpha f(x|s) dx + \int_{-\infty}^{\infty} p(x)^2 (1-\alpha) f(x|n) dx$$

On simplification, we get,

$$E[(y_i - p_i)^2] = \int_{-\infty}^{\infty} \frac{\alpha(1-\alpha) f(x|n) f(x|s)}{(1-\alpha) f(x|n) + \alpha f(x|s)} dx.$$

Further, we assume that the convergence in probability of $\bar{B}$ given by the law of large numbers as M tends to infinity is the Brier score. Make the assumption that the calibration in the observer's probability estimate is perfect. That is, Expected Brier Score is equal to $\bar{B}$.

Therefore, [4] considered the Brier Score in the context of Signal Detection Theory by assuming the calibration in the observers probability estimate is perfect and provided the theoretical relationship between Brier Score and AUC. Therefore, Brier Score can be defined in the context of ROC Curve analysis as,

$$\bar{B} = \int_{-\infty}^{\infty} \frac{\alpha(1-\alpha) f(x|s)}{(1-\alpha) + \alpha \frac{f(x|s)}{f(x|n)}} dx. \quad \text{(with slope of the ROC Curve)} \tag{13}$$

or

$$\bar{B} = \int_{-\infty}^{\infty} \frac{\alpha(1-\alpha) f(x|n)}{\alpha + (1-\alpha) \frac{f(x|n)}{f(x|s)}} dx . \quad \text{(with inverse slope of the ROC Curve)} \tag{14}$$

Functional Relationship between AUC and Brier Score

Now we can describe $\bar{B}$ and AUC as the function of **b** and **α** as

$$\bar{B} = \int_{-\infty}^{\infty} \frac{\alpha(1-\alpha)}{(1-\alpha) + \alpha \frac{dP(x|s)}{dP(x|n)}} dP(S|s) \quad . \quad \text{(with slope of the ROC Curve)} \tag{15}$$

or

$$\bar{B} = \int_{-\infty}^{\infty} \frac{\alpha(1-\alpha)}{\alpha + (1-\alpha) \frac{dP(x|n)}{dP(x|s)}} dP(S|n) \ . \ \text{(with inverse slope of the ROC Curve)} \tag{16}$$

Therefore, the Brier Score can be defined as the function of likelihood ratio under a Bayesian set up. Now on substituting equation (10) in equation (16), the expression for $\bar{B}$ as a function of **b** and **α** in the context of ROC Curve is follows,

$$\bar{B} = \int_{0}^{1} \frac{\alpha(1-\alpha)}{\alpha + (1-\alpha)\left[\frac{1}{b} exp\left\{\left[\frac{t^{\beta}}{\sigma_s}\right] - \left[\frac{t^{\beta}}{\sigma_n}\right]\right\}\right]} dt \ . \tag{17}$$

Now by substituting the threshold 't' value in the above equation (17) and on further simplification, the Brier Score ($\bar{B}$) expression which is a function of **b** and **α** reduces to

$$\bar{B} = \int_{0}^{1} \frac{\alpha(1-\alpha)}{\alpha + (1-\alpha)\left[\frac{1}{b} exp\left\{\left[\frac{-[\sigma_n logx(t)]^{\frac{1}{\beta}}}{\sigma_s}\right]^{\beta} - \left[\frac{-[\sigma_n logx(t)]^{\frac{1}{\beta}}}{\sigma_n}\right]^{\beta}\right\}\right]} dx(t) \ .$$

Therefore

$$\bar{B} = \int_{0}^{1} \frac{\alpha(1-\alpha)}{\alpha + (1-\alpha)\left[\frac{1}{b} exp\left\{(1-b)\left[[logx(t)]^{\frac{1}{\beta}}\right]^{\beta}\right\}\right]} dx(t) \ . \tag{18}$$

Therefore, the theoretical relationship between $\bar{B}$ and AUC of ROC Curve can be described by equation (18) and equation (7) using the two parameters **b** and **α**. The above expression (18) does not have a closed form and has to be evaluated using numerical integration methods.

Simulation Study

Simulation study is conducted to demonstrate how the proposed methodology works in a simulated environment for explaining the behavior of ROC Curve in terms of Brier Score. **b,** which is the ratio of scale parameters of Constant Shape Bi-Weibull distributions is considered between the range 0 and 1 with an increment of 0.1, since as **b** varies between zero and one. The ROC Curve possesses different shapes. Further Table 1 depicts the average Brier Scores by varying **b** values between zero and one with fixed values **α={0.1 to 0.99}** with an increment of 0.1.

Table 1 Average Brier Score for different values of b with fixed α

b	AUC	α=0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.99
0.1	0.9090	0.0118	0.0232	0.0339	0.0438	0.0525	0.0592	0.0629	0.0611	0.0475	0.0080
0.2	0.8333	0.0224	0.0433	0.0620	0.0781	0.0906	0.0982	0.0987	0.885	0.0606	0.0083
0.3	0.7692	0.0320	0.0607	0.0856	0.1057	0.1197	0.1260	0.1221	0.1043	0.0669	0.0084
0.4	0.7142	0.0406	0.0761	0.1057	0.1284	0.1426	0.1468	0.1385	0.1145	0.0705	0.0084
0.5	0.6667	0.0484	0.0897	0.1231	0.1473	0.1611	0.1629	0.1506	0.1216	0.0728	0.0085
0.6	0.6250	0.0554	0.1018	0.1381	0.1634	0.1764	0.1758	0.1599	0.1269	0.0745	0.0085
0.7	0.5882	0.0619	0.1125	0.1512	0.1771	0.1891	0.1862	0.1672	0.1309	0.0757	0.0085
0.8	0.5555	0.0677	0.1222	0.1628	0.1888	0.1997	0.1948	0.1731	0.1341	0.0766	0.0085
0.9	0.5263	0.0731	0.1308	0.1729	0.1991	0.2088	0.2019	0.178	0.1366	0.0774	0.0085
1	0.5000	0.078	0.1386	0.182	0.208	0.2166	0.208	0.182	0.1386	0.078	0.0085

From above Table 1, when **b**>1, the ROC Curve lies below the chance line which indicates less accuracy and possibly the test scores of both populations are completely overlapped.

Further, the **α** priori probability is partitioned into two sets, one set between 0.1 and 0.5 the other 0.6 and 0.99. This is established to show that how the Brier Score behaves with increment and decrement pattern as **b** changes. This means that the Brier Score values are found to be increasing as a priori probability **α** increases from 0.1 to 0.5 and then Brier Score is found to be decreasing as **α** increases from 0.6 to 0.99.

Thus, a priori probability **α** should posses a lesser value to explain the extent of correct classification. It is noticed that at **α**=0.5 with **b**=1, the Brier Score is found to 0.2166 and its corresponding AUC is 0.5.

Similarly, when at **α**=0.5 with **b**=0.1, the Brier Score is found to 0.0525 and its corresponding AUC is 0.909091, this reflects the meaning that the Brier Score is found to be similar to AUC of ROC Curve, where AUC approaches to one then $\bar{B}$ approaches zero and vice versa. Not only at **α**=0.5, but also at all values, the Brier Score and AUC values are observed to be inversely related to each other.

Further, the graphical plots are made in two ways, first to the effect of **b** on the behavior of Brier Scores at **α**={0.1 to 0.5} and **α**={0.6 to 0.99}. The first case is figured displayed in Figures (1) and (2) and the later one in Figures (3) and (4) respectively.

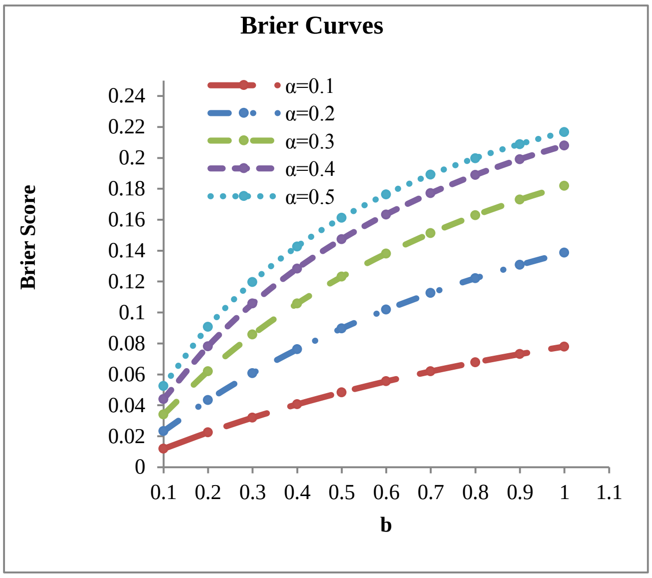

Fig. 1 Effect of b in Brier Curves at fixed value of α={0.1 to 0.5}

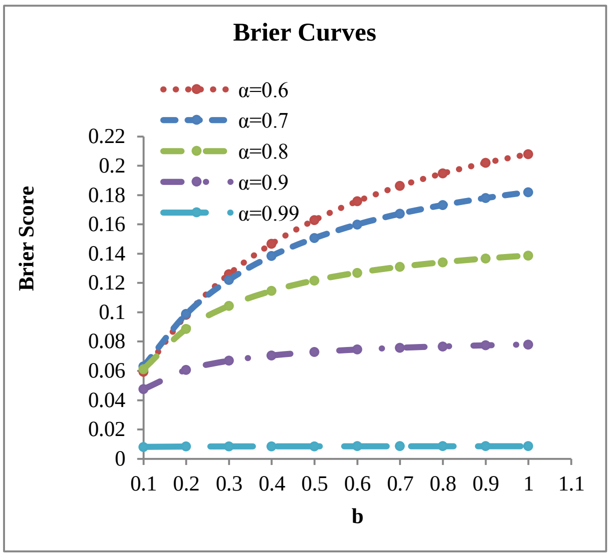

Fig. 2 Effect of b in Brier Curves at fixed value of α={0.6 to 0.99}

From Figures (1) and (2), it is visualized that as **b** increases from 0.1 to 1, Brier Curve moves towards the top right corner of the plot and if **α** increases from 0.1 to 0.5 then the curve approaches the top left corner of the plot. This is because of the influence of slope on HROC Curve, when the ratio **b** increases, the slope of HROC Curve will decrease.

Therefore, the decreased values of slope will influence the Brier Score to attain a larger value, since the expression for Brier Score involve the slope in its denominator.

Further Figures (3) and (4) show the one to one functional and relationship between $\bar{B}$ and AUC using the parameters **b** and **α**. As it is shown in the methodology that the entire behavior of ROC Curve depends on the ratio **b**. As **b** varies between zero and one the ROC curve gets different shapes.

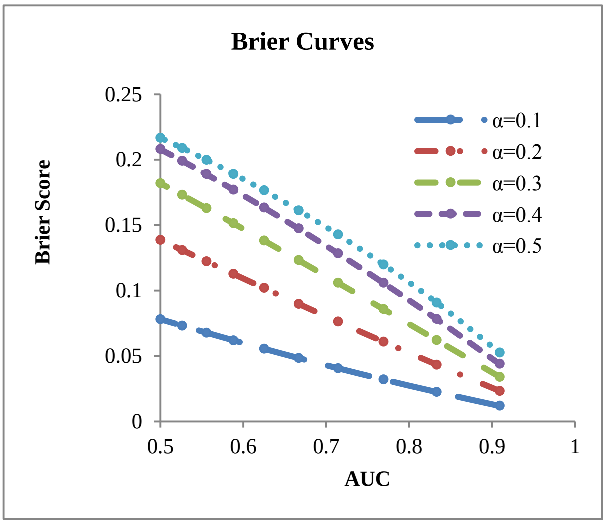

Fig.3 Relationship between Brier Score and AUC as b varies with fixed value of α={0.1 to 0.5}

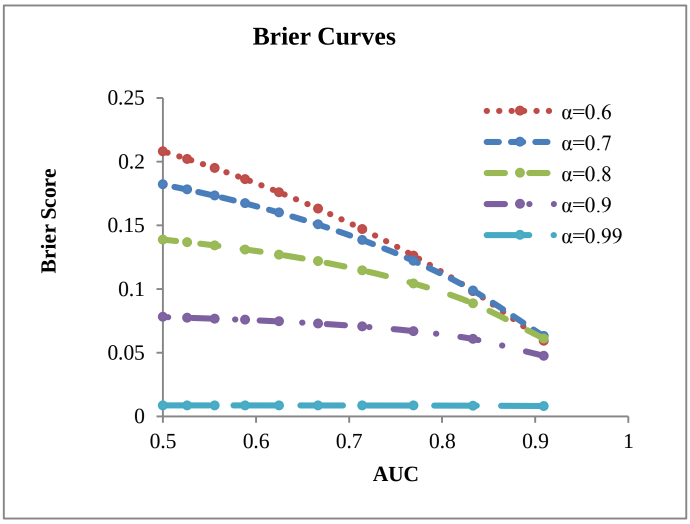

Fig.4 Relationship between Brier Score and AUC as b varies with fixed value of α={0.5 to 0.99}

From Figures (3) and (4), it is clearly seen that as AUC increases, the Brier Curve decreases with less accuracy at different values of **α**. Which means that as **b** varies between zero and one the AUC increases with decreasing values of Brier Score. As it is shown that when **b** decreases, the corresponding slope increases there by this increased slope will in turn gives rise to decreased values for the Brier Score, since slope lies in the denominator of the expression for Brier Score.

From Figure 3, it can be seen that when **α**=0.1, the Brier Curve is seen to be lower than the other values of **α.** Higher the value of **α** then the curve placed above all other values of **α**. Figure 4, it is seen that when **α**=0.6 the corresponding Brier Curve is found to be above all other values of **α** and as higher the value of **α**=0.99, the corresponding Brier Curve is below other curves.

From these experiments, it is observed that the Brie Score attains a maximum value at **α**=0.5 and **α**=0.6 and it increase till **α**=0.4 and then decreasing pattern after **α**=0.6.

Illustration

The Lung Cancer data was extracted from North Central Cancer Treatment Group [NCCTG] [10]. This data set contains 228 patients and censored subjects are 47 and died cases are 132. The Comparison of four parameters namely, Time, Sex, Karnofsky performance score rated by physician and Karnofsky performance score as rated by patient are considered for diagnosis using the Statistical methods AUC and Brier Score. Statistical prediction tools are same, but there is considerable disagreement about how they should be evaluated. These tools have been proposed for the status in patients with clinically localized prostate Lung cancer who are candidates for a radical prostatectomy. We evaluated four variables for status in Lung cancer patients.

Results and limitations: Traditional statistical methods (that is, ROC plots and Brier scores) could clearly determine which one of the variable tools should be preferred. The independent validation cohort consisting of 228 patients treated surgically for clinically localized Lung cancer at North Central Cancer Treatment Group.

Table 2 shows the characteristics of the study cohort (N=228). The patients' characteristics are

Table 2 Case Processing Summary

Status	Valid N (list wise)
Positive	47
Negative	132
Missing	49

Here larger values of the test result variables indicate the stronger evidence for a positive actual state. Here positive is censored and Negative is dead.

Table 3 shows statistics for each variable. The Karnofsky performance score as rated by patient had the highest area under the curve (AUC): 0.630 versus 0.625 for sex and 0.623 for Time, and 0.606 for Karnofsky performance score rated by physician. The Sex had Brier Score 0.225 and it had better Brier score than the Time (0.106), Karnofsky performance score as rated by patient(0.121), Karnofsky performance score rated by physician (0.078). No calibration plot could be provided for sex because it only provides binary results.

Table 3 Evaluation of the prediction models using AUC and Brier Score

Variable	AUC	Brier Score
Time	0.623	0.106
Sex	0.625	0.225
Karnofsky performance score rated by physician	0.606	0.121
Karnofsky performance score as rated by patient	0.630	0.078

AUC=Area Under the Curve

It is observed that when predicting status, surgeons should prefer the Karnofsky performance score as rated by patient over the Time, the Sex and the, Karnofsky performance score rated by physician because this variable provided the highest net benefit and more accuracy. In contrast to traditional statistical methods, decision curve analysis gave an unambiguous result applicable to both continuous and binary models, providing an insight into clinical utility.

CONCLUSION

The present work is carried out to establish a functional relationship between Brier Score and the Area under the ROC Curve .The entire relation depends on the ration **b**. As **b** attains smaller and larger values, the AUC and Brier Scores attain different set of values with smaller and larger values. If **b**>1, then the corresponding ROC Curve lies below the chance line explaining less accuracy and possibility that both populations are overlapped. The varying values of AUC and Brier Score not only depend on **b** but also on a priori probability **α**. From the results obtained, it is observed that as the parameter **α** is varied, a one to one relationship between Brier Score and AUC and the Brier Scores are observed and is used to explain the accuracy of the test. Therefore, the Brier

Score is used to assess the performance of a test as similar to AUC in binary classification. The functional relationship is established by the constant shape Bi-Weibull distribution. Here, the definition of Brier score does not require the assumption of underlying distribution. Therefore, this can be applied to any other parametric ROC model. However, it should be noted that the relationship between Brier Score and AUC derived is valid only under the assumption that the calibration in the observer's probability estimate is strong.

References

1. Bernard, C.K. Choi., 1998, Slopes of a Receiver Operating Characteristic Curve and Likelihood Ratios for a Diagnostic Test, *American Journal of Epidemiology* **148 (11)**, 1127-1132.
2. DeGroot, M., Fineberg, S., 1983, the comparison and evaluation of forecasters, *the statistician* **32**, 12-22.
3. Foster, D., Vohra, R.V., 1998, Asymptotic calibration, *Biometrika.* 379-390.
4. Mitsuru Ikeda., Takeo Ishigaki., Kazunobu Yamauchi., 2002, Relationship between brier score and area under the binormal roc curve, *Computer methods and programs in Biomedicine* **67,** 187-194.
5. Poses, R.M., Cebul, R.D., Centor, R.M., 1988, Evaluating physicians' probabilistic judgments, *Med. Decis. Making* **8,** 233-240.
6. Redelmeier, A., Bloch, D.A., Hickam, D.H., 1991, assessing predictive accuracy: how to compare Brier scores, *J. Clin. Epidemiol* **44**, 1141-1146.
7. Spiegelhalter, D.J., 1986, Probabilistic prediction in patient management and clinical trials, *Stat. Med* **5,** 421-433.
8. Sox, H.C., Blatt, M.A., Higgins, M.C., 1988, Medical Decision Making, *MA. Butterworth Publishers.*
9. Vishnu Vardhan, R., Lavanya, A., and Ganesh, T., 2015, Functional relationship between Brier Score and Area Under the ROC Curve in the Context of Skewed Distributions, *IJSIMR* **3**, 935-939.
10. https://vincentarelbundock.github.io/Rdatasets/datasets.html.

****

How to cite this article:

Leo Alexander T and lavanya A.2016, Functional Relationship between Brier Score and Area under the Constant Shape Bi-Weibull Roc Curve. *Int J Recent Sci Res.* 7(7), pp. 12330-12336.