Estimation of Exponential AUC using Jeffrey's Prior Information

[1]T. Leo Alexander, [2] M. Vanakumar

[1]Associate Professor, Department of Statistics, Loyola College, Chennai–34, India
[2] Ph.D. Research Scholar, Department of Statistics, Loyola College, Chennai–34, India

Abstract - In our analogy, the upper Area under Curve (AUC) represents the healthier characteristic between the subject/patient with the disease and without the disease. Receiver Operating Characteristic (ROC) curve was used to obtain the superiority of categorization in numerous health disorders. We used Bayesian methodology to the evaluation based on AUC for Exponential Distribution through Jeffrey's Prior Information. Finally, the simulation studies and an illustrative example demonstrate the theoretical results discussed.

Key Words - AUC, Bayesian Method, Jeffrey's method, Lindley's Approach, ROC model.

I. INTRODUCTION

In this paper, the Bayesian inference of the exponential AUC is of main interest. Lavanya et.al. (2016, 2017) discussed the Bayesian Estimation of Constant Shape Bi-Weibull AUC. The key purpose of this paper is to compare the Bayesian estimate of AUC values by the use of Jeffrey's three loss functions.

Assume that there are binary clusters of patients: ill-healthy and healthy. Firstly, the continuous biomarker is denoted as S. The True Positive Probability (TPP) namely $P(S \mid I)$ is the proportion of ill-healthy subjects detected by the diagnostic test. On the other hand, the False Positive Probability (FPP) namely $P(S \mid H)$ is the proportion of healthy subjects detected by a diagnostic test. Let α and β be the experiment scores detected from two populations, namely, ill-healthy individuals and healthy individuals respectively which follow exponential distributions. The density functions of exponential distributions are as follows

$$f\left(\alpha \middle| I\right) = \frac{1}{\eta_I} e^{-\left(\frac{x}{\eta_I}\right)} \tag{1}$$

and

$$g\left(\beta \middle| H\right) = \frac{1}{\eta_H} e^{-\left(\frac{x}{\eta_H}\right)} \quad , \tag{2}$$

where η_I and η_H are the scale parameters of ill-healthy and healthy populations respectively.

The ROC curve be represented as $ROC(x) = \beta(x)^b$, where $x = -\eta_H \ln \beta(x)$ is the threshold and $b = \frac{\eta_H}{\eta_I}$.

The accurateness based on diagnostic trial can be described by AUC. Moreover, the AUC value is always lies between zero and one. If AUC is range from 0.5 to 1 then it is a perfect classifier and if AUC is less than 0.5 then it considered as worst case of classification. Thus, AUC is defined as

$$\text{AUC} = \int_0^1 \beta(x)^b \, d\alpha(x) \quad .$$

Now AUC for exponential distribution is

$$\text{AUC} = \frac{1}{1+b}. \tag{3}$$

This paper is organized as follows: In Section II, the Bayesian estimation of AUC under Jeffrey's Prior Information deliberated. Section III and IV, provide simulation study and an illustrative example for the proposed theoretical results. Conclusions are given in Section V.

II. BAYESIAN ESTIMATION OF AUC

Bayesian estimation method has received a lot of attention for evaluating data. The Bayesian estimation approach is availed when prior knowledge about the parameters as well as the data is available. In case, prior information about the parameter does not exist, it is possible to make use of the non-informative prior in Bayesian analysis which is introduced by C. B. Guure et.al. (2012). Now this case, we introduce a non-informative prior method for estimating the parameters. Let $\alpha_1, \alpha_2, ..., \alpha_m$ is a sample taken randomly from $E(\eta_I)$ and one more sample $\beta_1, \beta_2, ..., \beta_n$ is taken randomly from $E(\eta_H)$.

The likelihood function based on particular sample is given by

$$\mathrm{L}\left(\alpha_i, \beta_j \middle| t\right) = \prod_{i=1}^{m} f_\alpha\left(\alpha_i \middle| \eta_I\right) \prod_{j=1}^{n} g_\beta\left(\beta_j \middle| \eta_H\right).$$

where $t = (\eta_I, \eta_H)'$.

Therefore
$$\mathrm{L}=\prod_{i=1}^{m}\frac{1}{\eta_I}e^{-\left(\frac{\alpha_i}{\eta_I}\right)}\prod_{j=1}^{n}\frac{1}{\eta_H}e^{-\left(\frac{\beta_j}{\eta_H}\right)}. \tag{4}$$

The log-likelihood function is obtained as
$$\ln L=-m\ln\eta_I-n\ln\eta_H-\frac{1}{\eta_I}\sum_{i=1}^{m}\alpha_i-\frac{1}{\eta_H}\sum_{i=1}^{m}\beta_j. \tag{5}$$

2.1 Jeffrey's Prior information

The Jeffrey's prior concept is used for invariance according to Sinha (1986). By presenting Jeffrey's Prior information such that $v_1(\eta)\propto\left(\frac{1}{\eta}\right)$.

Let us, consider the sample $X=\left(x_1,x_2,...,x_n\right)$.

Thus likelihood function from equation (4) becomes, $L=\prod_{i=1}^{n}\frac{1}{\eta}e^{-\left(\frac{x_i}{\eta}\right)}$.

The joint posterior distribution of the parameter η based on Baye's theorem is
$$\pi^*\left(\eta|x\right)\propto L\left(x|\eta\right)v_1\left(\eta\right), \tag{6}$$
where $L\left(x_i|\eta\right)=\frac{k}{\eta}\prod_{i=1}^{n}\frac{1}{\eta}e^{-\left(\frac{x_i}{\eta}\right)}$ and the marginal distribution is $\iint\frac{k}{\eta}\prod_{i=1}^{n}\frac{1}{\eta}e^{-\left(\frac{x_i}{\eta}\right)}\mathrm{d}\eta$.

Here, the posterior density function (pdf) is achieved properly based on normalizing constant k by using equation (6). Also, to estimate the AUC values we availed three Loss functions, namely Linear and Exponential (LINEX), General Entropy (GE) and Squared Error (SE) Loss functions.

2.1.1 AUC for LINEX Loss Function

The LINEX Loss function is one of the Asymmetric Loss functions for which minimal loss attains at $\hat{t}=t$, which is given as
$$L\left(\hat{t}-t\right)\propto\exp\left\{l\left(\hat{t}-t\right)\right\}-l\left(\hat{t}-t\right)-1,$$
where $\hat{t}$ is an estimation of t and l is the shape parameter which is different from 0.

Also, the sign and magnitude of the 'l' represents the direction and degree of symmetry respectively.

If $l>0$ then there is overestimation and if $l<0$ then there is underestimation. Moreover, the LINEX loss function is same as the Squared Error Loss function when $l\cong 0$.

According to **Pandey et.al. (2011)**, the posterior expectation is
$$E_tL\left(\hat{t}-t\right)\propto\exp\left(l\hat{t}\right)E_t\left(\exp(-lt)\right)-l\left(\hat{t}-E_t\left(t\right)\right)-1.$$

The estimate $\hat{t}_{JL}$ denotes the Bayes Estimator of t under LINEX Loss Function and it is formed as $\hat{t}_{JL}=-\frac{1}{l}\ln E_t\left(\exp(-lt)\right)$, in which $E_t\left(\exp(-lt)\right)$ is real and finite. Also, the Bayes estimator $\hat{t}_{JL}$ based on a function $u=u\left(\exp\left(l\eta_I\right),\exp\left(l\eta_H\right)\right)$ is expressed as
$$\hat{u}_{JL}=E\left(\left(\exp\left(l\eta_I\right),\exp\left(l\eta_H\right)\right)|x\right)=\frac{\iint u\left(\exp\left(l\eta_I\right),\exp\left(l\eta_H\right)\right)\pi^*\left(\eta_I,\eta_H\right)d\eta_I d\eta_H}{\iint\pi^*\left(\eta_I,\eta_H\right)d\eta_I d\eta_H}. \tag{7}$$

From equation (7), we detected the fraction of integrals which cannot be obtained systematically. Further to estimate the parameters we considered Lindley's approximation method.

The ratio of integral according to Abdel-Wahid (1987), is
$$E\left[u\left(t\right)/z\right]=\frac{\int w\left(t\right)L\left(t\right)dt}{\int v\left(t\right)L\left(t\right)dt},$$
where $L\left(t\right)$ is the log-likelihood function and the arbitrary functions of t are $w\left(t\right)$ and $v\left(t\right)$. Now, by concerning Lindley's method, it is assumed that $v\left(t\right)$ remains the prior distribution for t and $w\left(t\right)=u\left(t\right)v\left(t\right)$, $u\left(t\right)$ which is the actuality certain function.

Now, based on Lindley's method for estimating parameters we considered the following equation as
$$\hat{t}=u+\frac{1}{2}\left[u_{20}\delta_{20}+u_{02}\delta_{02}\right]+u_{10}\rho_{10}\delta_{20}+u_{01}\rho_{01}\delta_{02}+\frac{1}{2}\left[L_{30}u_{10}\delta_{20}^2+L_{03}u_{01}\delta_{02}^2\right], \tag{8}$$

Here L is the log-likelihood based on (5).

Now based on equation (8), the first and second order derivatives for η_1 and η_H are derived which are shown as follows,

$$u(\eta_I)=\exp(-l\eta_I),\qquad u_{10}=\frac{\partial u}{\partial \eta_I}=-l\exp(-l\eta_I),\qquad u_{20}=\frac{\partial^2 u}{\partial \eta_I^2}=-l^2\exp(-l\eta_I),\qquad u_{01}=u_{02}=0$$

$$u(\eta_H)=\exp(-l\eta_H),\qquad u_{01}=\frac{\partial u}{\partial \eta_H}=-l\exp(-l\eta_H),\qquad u_{02}=\frac{\partial^2 u}{\partial \eta_H^2}=-l^2\exp(-l\eta_H),\qquad u_{10}=u_{20}=0,$$

$$\rho(\eta_I,\eta_H)=-\ln(\eta_I)-\ln(\eta_H),$$

$$\rho_{10}=\frac{\partial \rho}{\partial \eta_I}=-\frac{1}{\eta_I},\qquad \rho_{01}=\frac{\partial \rho}{\partial \eta_H}=-\frac{1}{\eta_H},\qquad \delta_{20}=(-L_{20})^{-1},\quad \delta_{02}=(-L_{02})^{-1},$$

$$L_{20}=\frac{\partial^2 L}{\partial \eta_I^2}=\frac{m}{\eta_I^2}-2\frac{\sum_{i=1}^{m}\alpha_i}{\eta_I^3},\quad L_{02}=\frac{\partial^2 L}{\partial \eta_H^2}=\frac{n}{\eta_H^2}-2\frac{\sum_{j=1}^{n}\beta_j}{\eta_H^3},\quad L_{30}=\frac{\partial^3 L}{\partial \eta_I^3}=-\frac{2m}{\eta_I^3}+6\frac{\sum_{i=1}^{m}\alpha_i}{\eta_I^4}\ \text{and}\ L_{03}=\frac{\partial^3 L}{\partial \eta_H^3}=-\frac{2n}{\eta_H^3}+6\frac{\sum_{j=1}^{n}\beta_j}{\eta_H^4}.$$

Finally, the obtained Bayes estimator for AUC through LINEX Loss function is

$$A\hat{U}C_{JL}=\frac{1}{1+\hat{b}_{JL}}\ \text{where}\ \hat{b}_{JL}=\frac{\hat{\eta}_{HJL}}{\hat{\eta}_{IJL}}.$$

2.1.2 AUC for General Entropy Loss Function

The General Entropy (GE) Loss is another type of asymmetric loss function that clearly explains the Entropy Loss and it is given as

$$L(\hat{t}-t)\propto\left(\frac{\hat{t}}{t}\right)^k-k\ln\left(\frac{\hat{t}}{t}\right)-1.$$

The estimate $\hat{t}_{JG}$ represents the Bayes Estimator of t under General Entropy Loss function and defined as $\hat{t}_{JG}=\left[E_t\left(t^{-k}\right)\right]^{-\frac{1}{k}}$ provided $E_t\left(t^{-k}\right)$ be present and finite. Also, the $\hat{t}_{JG}$ based on this Loss Function is

$$\hat{u}_{JG}=E\left(u\left(\eta_I^{-k},\eta_H^{-k}\right)|x\right)=\frac{\iint u\left(\eta_I^{-k},\eta_H^{-k}\right)\pi^*(\eta_I,\eta_H)\,d\eta_I d\eta_H}{\iint \pi^*(\eta_I,\eta_H)\,d\eta_I d\eta_H}.$$

Similarly based on equation (8), the first and second order derivatives for η_1 and η_H are derived which are given as

$$u=\eta_I^{-k},\qquad u_{10}=\frac{\partial u}{\partial \eta_I}=-k\eta_I^{-(k+1)},\qquad u_{20}=\frac{\partial^2 u}{\partial \eta_I^2}=k(k+1)\eta_I^{-(k+2)},\qquad u_{01}=u_{02}=0,$$

$$u=\eta_H^{-k},\qquad u_{01}=\frac{\partial u}{\partial \eta_H}=-k\eta_H^{-(k+1)},\qquad u_{02}=\frac{\partial^2 u}{\partial \eta_H^2}=k(k+1)\eta_H^{-(k+2)}\quad\text{and}\quad u_{10}=u_{20}=0.$$

Finally, the Bayes estimator for AUC through GE loss is given by

$$A\hat{U}C_{JG}=\frac{1}{1+\hat{b}_{JG}}\ \text{where}\ \hat{b}_{JG}=\frac{\hat{\eta}_{HJG}}{\hat{\eta}_{IJG}}.$$

2.1.3 AUC for Symmetric Loss Function

The symmetric loss function which is known as SE Loss function is $L(\hat{t}-t)\propto(\hat{t}-t)^2$. The nature of this Loss function is symmetric, that is, it gives the same weight age to mutually over and under estimation. In tangible lifetime, we come upon numerous situations somewhere over estimation might be more severe than under estimation and vice versa. The SE which is also called squared error function is the most common loss function in Bayesian estimation.

The squared error loss is specified as $E_t(x/t)=\left(\hat{t}(x)-t\right)^2$, where the expectation is taken over the joint distribution of θ and t.

The estimate $\hat{t}_{JS}$ denotes the Bayes Estimator of t based on SE and is defined as $\hat{t}_{JS}=E_t(t)$, only if $E_t(t)$ be existent. Moreover, the $\hat{t}_{JS}$ based on SE Loss Function is

$$\hat{u}_{JS}=E\left((\eta_I,\eta_H)|x\right)=\frac{\iint u(\eta_I,\eta_H)\pi^*(\eta_I,\eta_H)\,d\eta_I d\eta_H}{\iint \pi^*(\eta_I,\eta_H)\,d\eta_I d\eta_H}.$$

In the same way based on equation (8), the first and second order derivatives for η_1 and η_H are derived which are given as

$$u=\eta_I,\quad u_{10}=\frac{\partial u}{\partial \eta_I}=1,\quad u_{20}=u_{01}=u_{02}=0,\quad u=\eta_H,\quad u_{01}=\frac{\partial u}{\partial \eta_H}=1 \quad \text{and} \quad u_{02}=u_{10}=u_{20}=0.$$

Hence, the necessary estimation for AUC based on SE loss is given as

$$A\hat{U}C_{JS}=\frac{1}{1+\hat{b}_{JS}} \text{ where } \hat{b}_{JS}=\frac{\hat{\eta}_{HJS}}{\hat{\eta}_{IJS}}.$$

III. Simulation Study

The key reason for conducting the simulations is to demonstrate exactly how the exponential AUC values owns different values based on scale parameters of the ill-healthy and healthy distributions change. Moreover, the AUC is obtained based on each parameter mixture and sample size.

Now, we picked a sample of size (n) is equivalent to 25, 50 and 100 that denoting moderately lesser, intermediate and huge data set. Also, based on Bayes estimation using Jeffrey's method the parameters were determined. Also, the parameter values are preferred as $\eta_H=0.8$ and $\eta_I=1.8$. The parameters based on loss function are preferred as $l=k=\pm 0.2$ and $l=k=\pm 1.2$.

In Table 1, we obtained the expected values for AUC based on Bayesian Estimation using Jeffrey's prior information by means of Jeffery's LINEX Loss function(JL), Jeffery's General Entropy Loss function (JG) and Jeffery's Squared Error Loss function(JS).

Table 1: Estimated AUC values based on Jeffrey's prior information

Sample size	η_I	η_H	$\widehat{AUC}_{JS}$	$\widehat{AUC}_{JL}$	$\widehat{AUC}_{JG}$	$\widehat{AUC}_{JL}$	$\widehat{AUC}_{JG}$	$\widehat{AUC}_{JL}$	$\widehat{AUC}_{JG}$	$\widehat{AUC}_{JL}$	$\widehat{AUC}_{JG}$
				$l=k=0.2$		$l=k=-0.2$		$l=k=1.2$		$l=k=-1.2$	
(25,25)	1.8	0.8	0.6980	0.4484	0.4551	0.5543	0.5438	0.2530	0.2401	**0.8087**	0.7298
(50,50)	1.8	0.8	0.7085	0.4436	0.4530	0.5585	0.5461	0.2287	0.2308	**0.8183**	0.7418
(100,100)	1.8	0.8	0.7154	0.4570	0.4660	0.5620	0.5521	0.2349	0.2812	**0.8281**	0.7632

From above Table 1, it is perceived that LINEX and GE Loss functions estimate the least AUC values when $l=k=0.2$ and 1.2. Moreover, the highest AUC values are attained based on LINEX loss function when $l=k=-1.2$. Hence the LINEX loss function turned out to be the ideal loss function for estimating the exponential AUC.

IV. Illustrative Example

We availed the dataset with low birth weight (LBW) extracted from www.umass.edu, which contains 189 observations. From the data set the mother's age variable treated as influential variable for making a diagnosis. Here, 59 babies are LBW and 130 are not. Table 2 illustrates the expected exponential AUC values using LBW sample data.

Table 2: LBW Data AUC values

(I, H)=(59,130)	$\eta_D=1.8$, $\eta_H=0.8$
JS	0.6923
JL(l=k=0.2)	0.4501
JL(l=k=–0.2)	0.5498
JL(l=k=1.2)	0.2310
JL(l=k=–1.2)	**0.7684**
JG(l=k=0.2)	0.4595
JG(l=k=–0.2)	0.5404
JG(l=k=1.2)	0.2743
JG(l=k=–1.2)	0.7258

From Table 2, we detect that highest AUC value is obtained under LINEX Loss function when parameters of loss function values are l = k = –1.2. So that, the LINEX loss function is the best estimation method for exponential AUC using LBW data set. Also, graphical representation is used to validate the proposed approach. Figure 1, depicts the Expected exponential AUC values using LBW sample data.

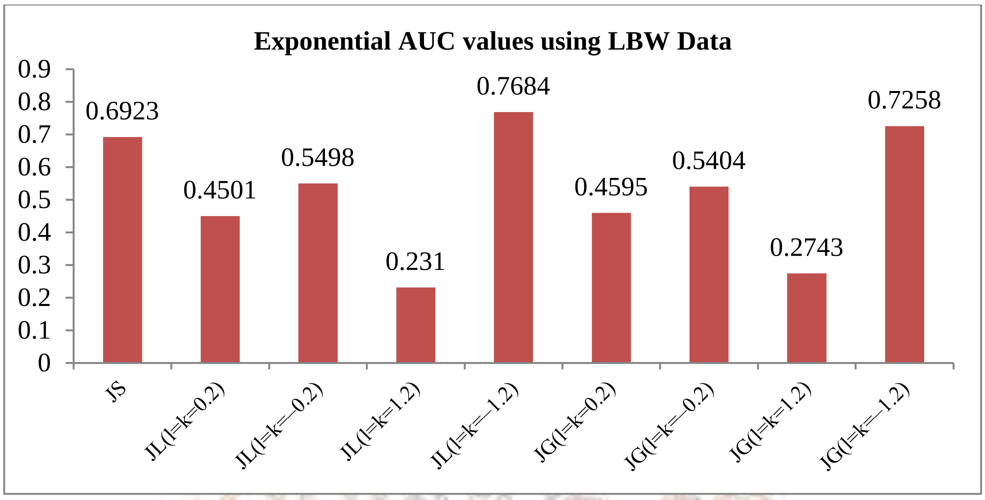

Figure 1: Estimated Exponential AUC values

From the above Figure 1, it is envisaged that LINEX Loss function gives the uppermost AUC value while loss parameters are l = k = –1.2.

V. Conclusion

The principal objective of this paper is to associate the Bayesian estimate AUC values through Jeffrey's three loss functions. Also, the Simulation study helps to observe and associate the performance of the AUC estimators aimed at different sample sizes through different loss parameter values. Moreover, we noted that LINEX loss function gives the maximum AUC values while loss parameters are l = k = –1.2. Hence, the Jeffery's LINEX loss function is the best estimation method for estimating the exponential AUC values.

VI. References

1. Abdel-Wahid, A. A., & Winter bottom, A. (1987). Approximate Bayesian estimates for the Weibull reliability function and hazard rate from censored data. Journal of Statistical Planning and Inference, 16(3), 277–283.
2. Guure, C. B., Ibrahim, N. A. and Al Omari, A. M. (2012). Bayesian estimation of two-parameter weibull distribution using extension of Jeffreys' prior information with three loss functions, Mathematical Problems in Engineering, 13.
3. Lavanya, A. & Leo Alexander, T. (2016). Confidence Intervals Estimation for ROC Curve, AUC and Brier Score under the Constant Shape Bi-Weibull Distribution. International Journal of Science and Research, 5(8), 371-378.
4. Lavanya, A. & Leo Alexander, T. (2016). Asymmetric and Symmetric Properties of Constant Shape Bi-Weibull ROC Curve Described by Kullback-Leibler Divergences. International Journal of Applied Research, 2(8), 713-720.
5. Lavanya, A. & Leo Alexander, T. (2016). Bayesian Estimation of Parameters under the Constant Shape Bi-Weibull Distribution Using Extension of Jeffreys' Prior Information with Three Loss Functions. International Journal of Science and Research, 5(9), 96-103.
6. Lavanya, A. & Leo Alexander, T. (2016). Estimation of Parameters using Lindley's Method. International Journal of advanced Research, 4(9), 1767-1778.
7. Lavanya, A. & Leo Alexander, T. (2017).Bayesian Estimation of AUC of A Constant Shape Bi-Weibull Failure Time Distribution. Journal of Management and Science, (1), 76-91.
8. Leo Alexander, T. & Lavanya, A. (2016). Estimation of AUC for Constant Shape Bi-Weibull Failure Time Distribution. International Journal of Recent Scientific Research, 7(10), 13826-13836.
9. Pandey, B. N., Dwividi N. & Pulastya B. (2011). Comparison between bayesian and maximum likelihood estimation of the scale parameter in Weibull distribution with known shape under linex loss function. Journal of Scientific Research, (55), 163–172.
10. Sinha, S. K. (1986). Bayesian estimation of the reliability function and hazard rate of a Weibull failure time distribution. Tranbajos De Estadistica, 1(2), 47–56.
11. https://www.umass.edu/statdata/statdata/data/lowbwt.txt