

Journal of Statistics Applications & Probability Letters An International Journal

http://dx.doi.org/10.18576/jsapl/070202

Statistical Analysis of the BIESEP ROC Curve

Ahmad Flaih ^{1,*}, Chary Akmyradov², Jose Guardiola³, and Hassan Elsalloukh⁴

Received: 15 Oct. 2019, Revised: 5 Dec. 2019, Accepted: 12 Feb. 2020

Published online: 1 May 2020

Abstract: In this work, we consider a case study for the Bi-Epsilon Skew Exponential Power (BIESEP) ROC curve proposed by Flaih et al. [3]. This model is a generalization of the Epsilon Skew bi-normal ROC curve proposed by Mashtare Jr. and Huston [10]. Elsalloukh [1,2] provided the Epsilon Skew Exponential Power (ESEP) which is less sensitive to outliers. The ESEP family can be adopted to cope with skewness and kurtosis of a data set. The ESEP model provides an appropriate choice to increase the robustness of data analysis. We develop the Epsilon Skew bi-normal ROC curve based on the outcomes of the diagnostic test that is distributed according to the ESEP distribution. More specifically, we derive the BIESEP ROC parameters and the Area Under the Curve (AUC). Also, we consider the parameter estimation of BIESEP ROC curve and the AUC of a diagnostic test. We employ the BIESEP ROC curve to analyze a real dataset.

Keywords: ROC Curve, Epsilon Skew, Exponential Power, Distribution Theory

1 Introduction

Medical diagnostic testing is frequently used in medical practice because it plays an important role in distinguishing between different health states, e.g. healthy and non-healthy. The Receiver Operating Characteristic (ROC) curve is an appropriate and well accepted statistical tool for displaying the performance and accuracy of a medical diagnostic test in situations where there are two possible states, diseased / non-diseased, event/ non-event, or any binary outcome. Some standard methods for estimating the ROC curve and the related measures are parametric, non-parametric, and semi parametric methods. The parametric approach specifies a distribution for the diagnostic test outcomes. The non-parametric methods do not require any assumptions on either the density function of data or the function of the ROC curve. The semi-parametric methods assume the ROC curve as a smooth function, and present fewer assumptions than the parametric methods. In this work, we suppose that the outcome measurement *Y* of a medical diagnostic test result is continuous and distributed according to the Epsilon Skew Exponential Power (ESEP) distribution family proposed by Elsalloukh [1,2]. We employ the BIESEP ROC proposed by Flaih et al. [3] to analyze the ROC curve when the test results are distributed as ESEP. As a standard method, the two by two table for any set of test results is used.

Table 1: Frequencies of Diagnostic Test Outcome

\downarrow Results, Status \rightarrow	event	non-event	Total
Positive	True positives= a	False positives= b	a+b
Negative	False negatives= c	True negatives= d	c+d
Total	$m_1 = a + c$	$m_2 = b + d$	$m_1 + m_2$

^{*} Corresponding author e-mail: ahmad.flaih@qu.edu.iq

¹Department of Statistics, University of Al-qadisiyah, Diwaniyah 58001, Iraq

²Arkansas Children's Research Institute, 13 Children's Way, Slot 842, Little Rock, AR, 72202, U.S.

³Department of Mathematics and Statistics, Texas A&M University-Corpus Christi, 6300 Ocean Drive, CI-109, Corpus Christi, TX 78412, U.S.

⁴Mathematics and Statistics Department, University of Arkansas at Little Rock, 2801 S University Ave., Little Rock, AR, 72204, U.S.



The summary indices of the diagnostic test performance are described in the following table:

Table 2: Frequencies of Diagnostic Test Outcome					
\downarrow Results, Status \rightarrow	event	non-event			
Positive	TPR = a/(a+c)	FPR = b/(b+d)			
Negative	FNR = c/(a+c)	TNR = d/(b+d)			
Total	1	1			

Table 2: Frequencies of Diagnostic Test Outcome

2 The Receiver Operating Characteristic (ROC) Curve and the Area Under Curve (AUC)

In this section, we give an overview of using the ROC curve as a statistical decision tool. Faraggi and Reiser [4] investigated some nonparametric and parametric methods to estimate and compare the area under the ROC curve. Betinec [5] developed the ROC curve based on the exponential distribution as a distribution for the diagnostic test measurements. Zou et al. [6] discussed two parametric models, bi-normal and bi-Weibull models, and developed a goodness of fit test for the ROC curve. The medical diagnostic test with a continuous outcome *Y* distributed according to a normal distribution for detecting the disease assuming *t* as the threshold (cutoff) value of *Y* is called the bi-normal ROC curve [7]. The bi-normal ROC curve with a given threshold *t* is commonly assessed using the probabilities that correctly classify outcomes, which are called True Positive Rate (TPR) and False Positive Rate (FPR), defined respectively as

$$TPR = P(Y_E > t/event) = 1 - \Phi(\frac{t - \theta_E}{\sigma_E}), \tag{1}$$

$$FPR = P(Y_{\ddot{E}} > t/non - event) = 1 - \Phi(\frac{t - \theta_{\ddot{E}}}{\sigma_{\ddot{F}}}), \tag{2}$$

where $Y_E \sim N(\theta_E, \sigma_E^2)$ and $Y_{\bar{E}} \sim N(\theta_{\bar{E}}, \sigma_{\bar{E}}^2)$ are diagnostic test outcomes for detecting whether a subject is event (E) and/or non-event (E), and $\Phi(\cdot)$ is the standard normal cdf. Suppose $\theta_E > \theta_{\bar{E}}$, then we can define the ROC curve as

$$ROC(t) = (FPR(t), TPR(t)); t \in \Re.$$
 (3)

Hence, the bi-normal ROC curve equation is defined by

$$h(x, \theta) = 1 - \Phi[-a + b\Phi^{-1}(1-x)]; \quad 0 < x < 1,$$

where

$$a = \frac{\theta_E - \theta_{\ddot{E}}}{\sigma_E}, \quad b = \frac{\sigma_{\ddot{E}}}{\sigma_E}.$$

The ROC curve is a monotonically increasing function as illustrated in Figure (1).

The most commonly used summary index of the performance of a diagnostic test based on an ROC curve, Krzanowski and Hand [8], is the Area Under the ROC Curve, denoted by (AUC) is

$$AUC = P(Y_E \ge Y_{\stackrel{.}{E}}) = \Phi(\frac{\theta_E - \theta_{\stackrel{.}{E}}}{\sqrt{\sigma_E^2 + \sigma_E^2}}).$$

Moreover, AUC index is a popular summary measure of diagnostic test accuracy based on an ROC curve [9]. Mashtare Jr. and Huston [10] considered the Epsilon Skew bi-normal ROC curve, derived TPR and FPR equations, estimated the area under the ROC curve, and discussed its application in biomedicine.

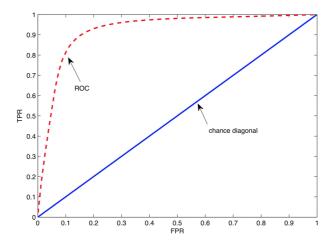


Fig. 1: ROC Curve

3 The Epsilon Skew Exponential Power (ESEP) distribution

Elsalloukh [1,2] introduced the Epsilon-Skew Exponential Power (ESEP) distribution that can accommodate heavy-tailed (Leptokurtic) and skewed data. The ESEP distribution family is attractive and flexible because it allows continuous variation from normality to non-normality. That is, the ESEP includes the normal distribution as a special case, so it is a "robust model". The ESEP density is denoted by $ESEP(\theta, \sigma, \alpha, \varepsilon)$ and defined by

$$f(y) = \frac{\alpha}{2\sigma\sqrt{2}\Gamma(\frac{1}{\alpha})} \begin{cases} \exp[-(\frac{y-\theta}{\sqrt{2}\sigma(1-\varepsilon)})^{\alpha}]; & y \ge \theta \\ \exp[-(\frac{\theta-y}{\sqrt{2}\sigma(1+\varepsilon)})^{\alpha}]; & y < \theta \end{cases}, \tag{4}$$

where $-1 < \varepsilon < 1$ is the skewness parameter, $\theta \in \Re$ is the location parameter, $\sigma > 0$ is the scale parameter, and $\alpha \in \Re$ is the shape parameter. Moreover, the density function (4) is known as the Epsilon Skew Exponential Power of order α . The cumulative distribution function of the $ESEP(\theta, \sigma, \alpha, \varepsilon)$

$$F(y) = \begin{cases} 1 - \frac{1 - \varepsilon}{2\Gamma(\frac{1}{\alpha})} \Gamma(\frac{1}{\alpha}, g(y)); & y \ge \theta \\ \frac{1 + \varepsilon}{2\Gamma(\frac{1}{\alpha})} \Gamma(\frac{1}{\alpha}, h(y)); & y < \theta \end{cases},$$
 (5)

where, $\Gamma(\cdot,\cdot)$ is the incomplete gamma function, $g(y) = [\frac{y-\theta}{2^{1/2}(1-\varepsilon)\sigma}]^{\alpha}$ and $h(y) = [\frac{\theta-y}{2^{1/2}(1+\varepsilon)\sigma}]^{\alpha}$. The quantile function of Y is

$$F_{ESEP}^{-1}(\vartheta/\varepsilon,\alpha) = \begin{cases} \theta - 2^{1/2}\sigma(1+\varepsilon)[G^{-1}(\frac{\vartheta}{1+\varepsilon}2\Gamma(\frac{1}{\alpha});\frac{1}{\alpha})]^{\frac{1}{\alpha}}; & 0 < \vartheta < \frac{1+\varepsilon}{2} \\ \theta + 2^{1/2}\sigma(1-\varepsilon)[G^{-1}(\frac{1-\vartheta}{1-\varepsilon}2\Gamma(\frac{1}{\alpha});\frac{1}{\alpha})]^{\frac{1}{\alpha}}; & \frac{1+\varepsilon}{2} < \vartheta < 1 \end{cases},$$
(6)

where $G^{-1}(\cdot)$ is the inverse function of the gamma cdf $G(\cdot)$, and

$$G(y,\gamma) = \Gamma = (\Gamma(\gamma))^{-1} \int_0^y z^{\gamma-1} \exp(-z) dz.$$

In this research, we employed the special case of the $ESEP(\theta, \sigma, 1, \varepsilon)$, which is also known as Epsilon Skew Laplace (ESL) to analyze the simulation study and the application.

4 The Bi-Epsilon Skew Exponential Power (BIESEP) ROC Curve

Consider the receiver operating characteristic ROC curve as defined in (3), the TPR and FPR based on the ESEP family are

$$TPR = 1 - P(Y_E \le t) = 1 - F(\frac{t - \theta_E}{\sigma_E}),\tag{7}$$



$$FPR = 1 - P(Y_{\ddot{E}} \le t) = 1 - F(\frac{t - \theta_{\ddot{E}}}{\sigma_{\ddot{E}}}), \tag{8}$$

Let $Y_{\ddot{E}} \sim ESEP(\theta_{\ddot{E}}, \sigma_{\ddot{E}}, \alpha_{\ddot{E}}, \varepsilon_{\ddot{E}})$, and $Y_E \sim ESEP(\theta_E, \sigma_E, \alpha_E, \varepsilon_E)$ be the diagnostic test outcome for detecting whether a subject is event (E) or non-event (\ddot{E}) . The area under the BIESEP ROC curve is defined as

$$\widehat{AUC(s, \lambda)} \simeq \int_0^1 ROC(s, \lambda) ds.$$

To estimate this quantity from sample data, we can estimate the ROC curve by fitting a smooth curve assuming each of the two populations (event/non-event) has been distributed by ESEP, and then we obtained the $\widehat{AUC(s, \lambda)}$ numerically using the trapezoidal rule. Moreover, Flaih et al. [3] pointed out that the following propositions are necessary to set up the ESEP ROC function and its properties:

Proposition 1. If $Y_{\ddot{E}} \sim ESEP(\theta_{\ddot{E}}, \sigma_{\ddot{E}}, \alpha_{\ddot{E}}, \varepsilon_{\ddot{E}})$, and $Y_E \sim ESEP(\theta_E, \sigma_E, \alpha_E, \varepsilon_E)$ denote non-event and event test results, assuming $\theta_E > \theta_{\ddot{E}}$, then

$$ROC(s, \lambda) = 1 - F_E[-\alpha + \beta F_{F}^{-1}(1 - s)]; \quad s \in (0, 1),$$
 (9)

where

$$\alpha = \frac{\theta_E - \theta_{\ddot{E}}}{\sigma_E}, \quad \beta = \frac{\sigma_{\ddot{E}}}{\sigma_E}, \quad \boldsymbol{\lambda} = (\theta_E, \sigma_E, \alpha_E, \varepsilon_E, \theta_{\ddot{E}}, \sigma_{\ddot{E}}, \alpha_{\ddot{E}}, \varepsilon_{\ddot{E}})$$

and $F^{-1}(\cdot)$ and $F(\cdot)$ are the quantile function and the distribution function of the ESEP, respectively. The more convenient expression for equation (9) with the parameters α and β is

$$ROC(s) = 1 - F_E[-\alpha + \beta F_{F}^{-1}(1-s)]; \quad s \in (0,1),$$

where $\alpha = \frac{\theta_E - \theta_E}{\sigma_E}$, $\beta = \frac{\sigma_E}{\sigma_E}$ are the parameters of the BIESEP ROC curve.

Proposition 2. The MLE $\hat{\lambda}$ of λ is asymptotically normal,

$$\sqrt{n}(\hat{\boldsymbol{\lambda}} - \hat{\boldsymbol{\lambda}}) \stackrel{d}{\to} N(0, \Sigma),$$
 (10)

where Σ is the variance-covariance matrix,

$$\Sigma = \left[\frac{I_{E}^{-1} \mid 0}{0 \mid I_{E}^{-1}} \right] \tag{11}$$

and I is the Fisher information matrix as defined in Theorem 4.1 in section 4.3 Elsalloukh [1].

Proposition 3. Let Y_E and Y_{E} denote any binary continuous random variables, then

$$\widehat{AUC(s, \lambda)} \simeq \int_0^1 ROC(s, \lambda) ds$$

$$\simeq \frac{1}{m} \left[\frac{1}{2} g(0; \hat{\lambda}) + \frac{1}{2} g(1; \hat{\lambda}) + \sum_{i=2}^m g(\frac{i-1}{m}; \hat{\lambda}) \right]$$

$$\simeq \frac{1}{m} \left[\sum_{i=0}^m g_i(\hat{\lambda}) \right],$$
(12)

where m is the number of intervals, each of size $\frac{1}{m}$ and $ROC(s, \hat{\lambda}) = g(\cdot)$.

Proposition 4. Let $\widehat{AUC(s, \lambda)}$ be as defined in (12) with $\widehat{\lambda} \sim AN(\lambda, \Sigma_m)$, then for large samples $\widehat{AUC(s, \lambda)}$ is $AN(AUC(s, \lambda), Var(\widehat{AUC(s, \lambda)}))$, where

$$Var(\widehat{AUC(s, \boldsymbol{\lambda})}) = \frac{1}{m^2} [\sum_{i=0}^{m} g_i(\hat{\boldsymbol{\lambda}}) + 2 \sum_{i=0}^{m} \sum_{j>i}^{m} \sigma_{ij}(g_i(\hat{\boldsymbol{\lambda}}), g_j(\hat{\boldsymbol{\lambda}}))].$$
(13)



Proposition 5. Let $g(s, \hat{\lambda})$ be $(m-1) \times 1$ vector of the $g(\frac{i-1}{m}, \hat{\lambda})$, i = 1, ..., m-1, then as $m \to \infty$,

$$g(s, \hat{\boldsymbol{\lambda}}) \sim AN(g(s, \boldsymbol{\lambda}), \frac{1}{m} A \Sigma_m A'),$$
 (14)

where A is the 8×1 vector of partial derivatives of the $g_i(\hat{\lambda})$, and Σ_m is as defined in proposition 2.

Proposition 6. As $m \to \infty$,

$$\widehat{AUC(s, \boldsymbol{\lambda})} \sim AN(AUC, \frac{1}{m^2} \mathbf{1'}_{m-1} D\mathbf{1}_{m-1})$$

where 1_{m-1} is a $(m-1) \times 1$ vector of 1's, and $D = \frac{1}{m} A \Sigma_m A'$ is as defined in proposition 5.

In the context of BIESL ROC curve modeling, we discuss the problem of goodness of fit to illustrate how we can compare the underlying symmetric distribution with skewed distribution to fit a given data set using the skewness parameters ε_E and ε_{E} as a measure of distance of the distribution of the underlying variable from symmetry with respect to skewed one. Then, the test

$$H_0: arepsilon_E = 0$$
 $H_0: arepsilon_{\ddot{E}} = 0$ and $H_1: arepsilon_E
eq 0$ $H_1: arepsilon_{\ddot{E}}
eq 0$

can be performed using the methods described above.

5 Simulation Study

In this section we employ the simulation analysis to explore the behavior of the estimated values of AUC based on the BIESEP ROC. We followed the same procedure that was proposed by Mashtare Jr. and Huston [10] to compare the results. Mashtare Jr. and Huston [10] examined different scenarios for estimating the AUC and the ROC equations. In the first simulation, we estimate the value of the AUC assuming that, $Y_{\bar{E}} \sim ESEP(0,1,1,\varepsilon_{\bar{E}})$, and $Y_E \sim ESEP(1,1.5,1,\varepsilon_E)$ with sample size $n_E = n_{\bar{E}} = 100$ and assuming that $\varepsilon_{\bar{E}}$ and ε_E range is (-0.25,0.25), that means we have nine scenarios. AUC calculations are based on the $ESEP(\theta,\sigma,1,\varepsilon)$ distribution. MATLAB® provided the results that are summarized in Table 3. We compared our results to the results of simulation study reported by Mashtare Jr. and Huston [10]. The results for calculating the BIESL revealed the benefits of a skewed parametric model in terms of smoothness compared to other models.

 Table 3: AUC for bi-normal, ESBN, Nonparametric, and BIESL Models

ε_E	$arepsilon_{\ddot{E}}$	Expected AUC	bi-normal	ESBN	Nonparametric	BIESL
-0.25	-0.25	0.679	0.670	0.679	0.679	0.823
-0.25	0	0.779	0.781	0.780	0.779	0.746
-0.25	0.25	0.866	0.863	0.865	0.866	0.660
0	-0.25	0.603	0.588	0.603	0.602	0.864
0	0	0.710	0.710	0.711	0.710	0.797
0	0.25	0.809	0.810	0.808	0.808	0.718
0.25	-0.25	0.522	0.500	0.522	0.531	0.890
0.25	0	0.634	0.630	0.633	0.633	0.832
0.25	0.25	0.741	0.743	0.740	0.739	0.762

A second simulation study is conducted to compare the BIESL ROC curve with bi-normal ROC. Assuming the same assumptions in the first simulation about the distributions of the Y_E and Y_{E} and under the variety of ε_E and ε_{E} with different values of the FPR(t) when the TPR is t. Figure 2 shows BIESL ROC curve plot versus bi-normal ROC curve plot. In each scenario the BIESL ROC curve plot satisfies the condition that the ROC curve is a monotonic increasing function, which means a trade off between TPR and FPR. In most cases, we found that the BIESL ROC curve follows the left-hand border and then the top border of the ROC space, which means that the BIESL ROC curve is a closer curve than bi-normal ROC curve. Moreover, when departures from normality are observed, the BIESL ROC curve sometimes crosses the bi-normal ROC curve.



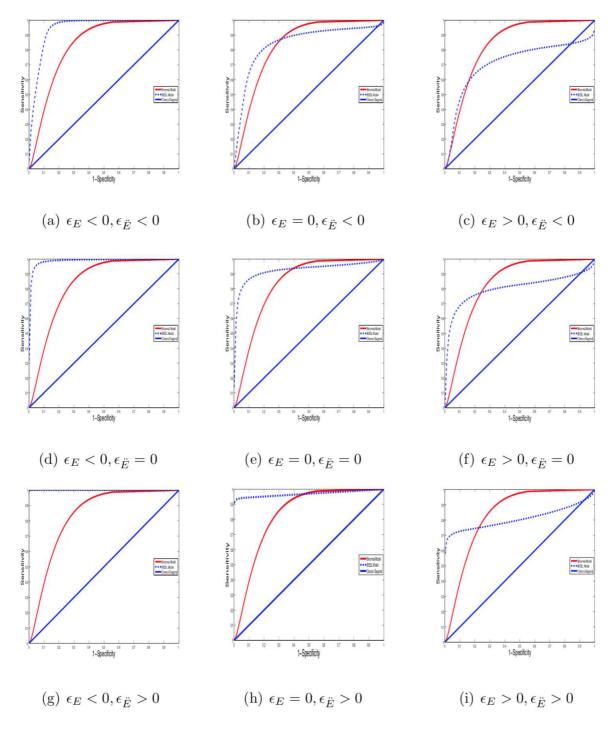


Fig. 2: The BIESL Curve and Normal Curve

6 Data Analysis

In this section, we provide an application that uses the $ESPS(\theta, \sigma, 1, \varepsilon)$ distribution to plot the BIESL ROC curve and compare it with the bi-normal ROC curve. The data are taken from Heinz et al. [11], who examined several skeletal measurements as possible predictors for gender using the bi-normal model. From the dataset consisting of skeletal measurements in 507 physically active individuals, 247 men Y_m and 260 women Y_w , we choose the chest depth

measurements between spine and sternum at nipple level as a predictor for gender. Figure 3 and Figure 4 show the histogram of the chest depth measurements for men and the chest depth measurement for women versus the fitted curves of the ESL function and normal density function. The maximum likelihood estimation of the Y_m and Y_w are $(\theta_m = 19.9, \sigma_m = 0.2704, \varepsilon_m = 1.7412)$ and $(\theta_w = 17.1, \sigma_w = 0.2298, \varepsilon_w = 1.5968)$, respectively. For testing $H_0: \varepsilon_m = 0$ and $H_0: \varepsilon_w = 0$, P-values were 0.0002 and 0.008 respectively indicating that the BIESL model is more appropriate than bi-normal model for fitting the dataset. Also, we test the goodness of fit for our data using Kolmogorov-Smirnov (K-S) test. We have computed the differences between the empirical cdf $\frac{i-0.375}{n+0.25}$ and the theoretical distribution function (5) of the $ESEP(\theta,\sigma,1,\varepsilon)$. We have found that the Maximum Differences (MD) is 0.1033 for women and 0.1037 for men. Comparing these MD-s with the critical values at 10% significance level which is obtained from a table of K-S, 0.10108 and 0.10371, as might be expected from the hypotheses tests, K-S indicate that the ESL model is a better fit than the normal model. The estimated AUC for the BIESL ROC curve was 0.9120 and 0.8820 for the bi-normal ROC curve, as expected from the second simulation study. The BIESL ROC curve is comparable to the bi-normal ROC curve. Figure 5 shows the estimated ROC curve, so we conclude that the diagnostic test with $ESEL(\theta,\sigma,1,\varepsilon)$ outcomes is better than with outcomes distributed according to normal distribution to predict gender.

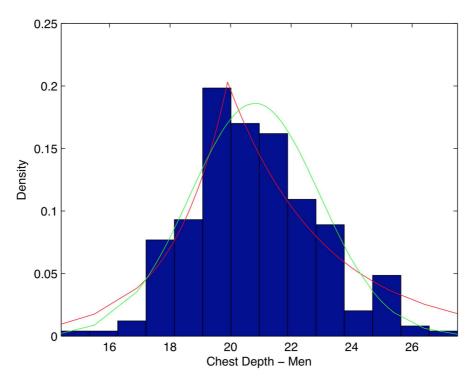


Fig. 3: Histogram and Fitted ESL and Normal Densities of Chest Depth Measurements by Men

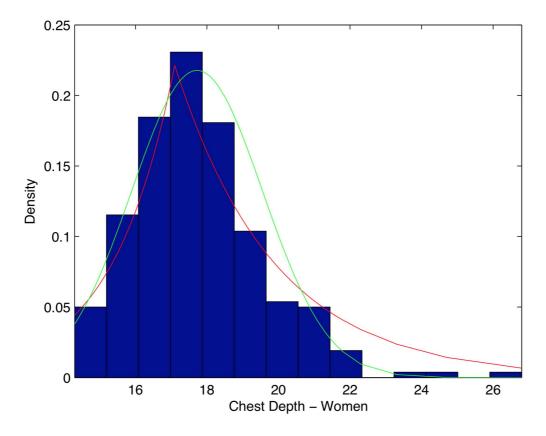


Fig. 4: Histogram and Fitted ESL and Normal Densities of Chest Depth Measurements by Women

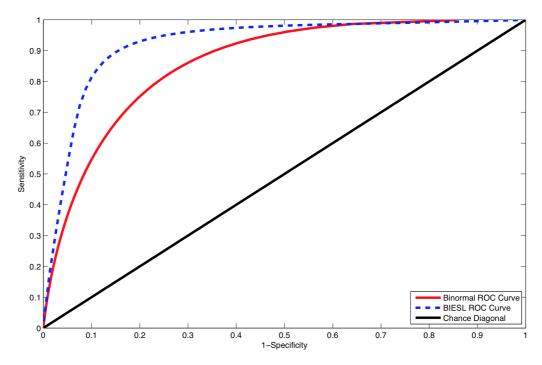


Fig. 5: ROC curves for Chest Depth as a Predictor of Gender Using BIESL and bi-normal Models

References

- [1] H. Elsalloukh. The epsilon skew exponential power distribution. Ph.D. Dissertation, Baylor University, USA, (2004).
- [2] H. Elsalloukh, J. Guardiola, and M. D. Young. The Epsilon-Skew Exponential Power Distribution Family. *Far East Journal of Theoretical Statistics*, **16**, 1, 97–112, (2005).
- [3] A. Flaih, H. Elsalloukh and J. Guardiola. The Bi-Epsilon Skew Exponential Power (BIESEP) ROC Curve. *JSM Statistical Computing Section. American Statistical Association*, in Proc. (2012).
- [4] D. Faraggi, and B. Reiser. Estimation of the area under the ROC curve. Statistics in Medicine, 21, 3093–3106, (2002).
- [5] M. Betinec. Testing the difference of the ROC curves in biexponential model. Tatra Mt. Math. Publ, 39, 215–223, (2008).
- [6] K. H. Zou, J. L. Gastwirth, and B. J. McNeil. A Goodness-of-Fit Test for a Receiver Operating Characteristic Curve from Continuous Diagnostic Test Data. *Institute of Mathematical Statistics*, **43**, 59–68, (2003).
- [7] D. D. Dorfman, and E. Alf. Maximum likelihood estimation of parameters of signal detection theory and determination of confidence intervals rating method data. *Journal of Mathematical Psychology*, **6**, 487–496, (1968).
- [8] W. J. Krzanowski, and D. J. Hand. ROC Curves for Continuous Data. Taylor and Francis Group, (2009).
- [9] D. Bamber. The area above the ordinal dominance graph and the area below the receiver operating graph. *Journal of Mathematical Psychology*, **12**, 387–415, (1975).
- [10] T. Mashtare Jr., and A. Huston. Epsilon-skew-binormal receiver operating characteristic (ROC) curves. (2009) (in press).
- [11] N. Henze. A probabilistic representation of the 'skew-normal' distribution. Scand. J. Statist., 13, (1986).





Ahmad Flaih is an Assistant Professor, Department of Statistics, University of Al-qadisiyah, Diwaniyah 58001, Iraq. (email: ahmad.flaih@qu.edu.iq).



Chary Akmyradov is a Senior Biostatistician, Arkansas Children's Research Institute, 13 Children's Way, Slot 842, Little Rock, Arkansas, 72202 (email: akmyradovc@archildrens.org).



Jose Guardiola is an Associate Professor, Department of Mathematics and Statistics, Texas A&M University- Corpus Christi, 6300 Ocean Drive, CI-109, Corpus Christi, TX 78412 (email: jose.guardiola@tamucc.edu).



Hassan Elsalloukh is a Professor, Department of Mathematics and Statistics, University of Arkansas at Little Rock, 2801 South University Avenue, Little Rock, Arkansas, 72204 (email: hxelsalloukh@ualr.edu).