Epsilon-skew-binormal receiver operating characteristic (ROC) curves

Terry L. Mashtare Jr.

Department of Biostatistics

University at Buffalo, 249 Farber Hall, 3435 Main Street, Buffalo, NY 14214-3000, U.S.A.

Roswell Park Cancer Institute, Elm & Carlton Streets, Buffalo, NY 14263, U.S.A.

Alan D. Hutson

Department of Biostatistics

University at Buffalo, 249 Farber Hall, 3435 Main Street, Buffalo, NY 14214-3000, U.S.A.

April 30, 2009

Abstract

In this note we extend the well-known binormal model via implementation of the epsilon-skew-normal (ESN) distribution developed by Mudholkar and Hutson (2000). We derive the equation for the receiver operating characteristic (ROC) curve assuming epsilon-skew-binormal (ESBN) model and examine the behavior of the maximum likelihood estimates for estimating the ESBN parameters. We then summarize the results of a simulation study to examine the asymptotic properties of the maximum likelihood estimates in the ESBN model and compare with the maximum likelihood estimates in the binormal

model. We also summarize the results of a simulation study comparing the two parametric models to the nonparametric ROC model. We then illustrate the maximum likelihood estimation of the ESBN model using data involving skeletal measurements in 507 physically active individuals.

Keywords: AUC, diagnostic testing, prediction.

1 Introduction

It is commonplace in medical studies to dichotomize a continuous predictor at a cutoff c for the purpose of diagnosing disease (yes/no). A well known method for summarizing the choice of c in terms of medical decision making is receiver operating characteristic (ROC) analysis. ROC analysis has a long history and spans parametric and non-parametric estimation methods; see [1] for some of the early references. In the last few decades, ROC analysis has been a mainstay in diagnostic testing across various medical decision making scenarios. Some of the more recent developments include time-dependent ROC analysis [2], multireader ROC analysis [3], and Bayesian bootstrap estimation of the ROC curve [4].

In general, each cutoff c produces a true positive fraction (TPF) and a false positive fraction (FFP). The TPF, or sensitivity, is the probability that an individual with the disease tests positive for the disease. The FPF is the probability that a healthy individual tests positive for the disease. The FPF is 1 − specificity, where specificity is the probability that a healthy individual tests negative for the disease. The ROC curve is a plot over all values of c of TPF (sensitivity) versus FPF (1 − specificity). A popular summary measure of diagnostic accuracy for predicting disease is the area under the ROC curve (AUC). Hanley (1982) has shown shown that the AUC is the probability that a measurement for an individual with the disease

individual exceeds the measurement for an individual without the disease [5].

A classic approach for estimating an ROC curve and the corresponding AUC is the binormal model. There are hundreds of studies that have employed the binormal model; e.g. for some recent medical studies that have utilized the binormal model see [6, 7, 8]. It is well known that the binormal model is robust in terms of unbiased estimates of the AUC. However, there can be substatial bias in the pointwise estimates of the estimated ROC curve, which in turn will create inaccurate thresholds for the final decision rule for diagnosing disease [9].

We propose the epsilon-skew-binormal (ESBN) model to provide a more robust alternative to the binormal model in terms of overcoming its deficiencies and to also provide a comparable alternative to nonparametric methods. The epsilon-skew-normal (ESN) distribution was developed by Mudholkar and Hutson (2000) [10]. Furthermore, the binormal model assumptions may be examined due to the fact that it is a special case of the ESBN model.

In section 2 we review the binormal model for estimating ROC curves as well as the ESN distribution. In section 3 we derive the equation for the ROC curve assuming ESBN model and examine the behavior of the maximum likelihood estimates for estimating the ESBN parameters. In section 4 we then summarize the results of a simulation study to examine the asymptotic properties of the maximum likelihood estimates in the ESBN model and compare with the maximum likelihood estimates in the binormal model. We also summarize the results of a simulation study comparing the two parametric models to the empirical ROC model. In section 5 we then illustrate the maximum likelihood estimation of the ESBN model using data involving skeletal measurements in 507 physically active individuals.

2 Background

2.1 Binormal Model

Let Y be a continuous predictor for diagnosing disease in a given population. Let Y_0 and Y_1 denote the measurements of the predictor Y for healthy and diseased populations, respectively. Suppose $Y_0 \sim N(\mu_0, \sigma_0^2)$ and $Y_1 \sim N(\mu_1, \sigma_1^2)$, and $\mu_1 > \mu_0$. For a given cutoff c, the FPF is given by:

$$FPF = P(Y_0 > c) = 1 - P(Y_0 \leq c) = 1 - \Phi\left(\frac{c - \mu_0}{\sigma_0}\right) = \Phi\left(\frac{\mu_0 - c}{\sigma_0}\right), \tag{2.1}$$

where $\Phi(\cdot)$ denotes the standard normal c.d.f. Solving for c, we get

$$c = \mu_0 - \sigma_0 \Phi^{-1}(FPF). \tag{2.2}$$

Similarly, for a given cutoff c, the TPF is given by:

$$TPF = P(Y_1 > c) = 1 - P(Y_1 \leq c) = \Phi\left(\frac{\mu_1 - c}{\sigma_1}\right). \tag{2.3}$$

Substituting (2.2) into (2.3) we have:

$$TPF = \Phi\left[\frac{(\mu_1 - \mu_0)}{\sigma_1} + \frac{\sigma_0}{\sigma_1}\Phi^{-1}(FPF)\right]. \tag{2.4}$$

Thus, the equation of the binormal ROC curve is given by:

$$g(x; \boldsymbol{\theta}) = \Phi[a + b\Phi^{-1}(x)], \tag{2.5}$$

where $\boldsymbol{\theta} = (\mu_0, \sigma_0, \mu_1, \sigma_1)'$, $a = (\mu_1 - \mu_0)/\sigma_1$, and $b = \sigma_0/\sigma_1$.

For the binormal ROC curve, there is a closed form solution for the AUC [1], given by

$$AUC = \Phi\left(\frac{a}{\sqrt{1 + b^2}}\right), \tag{2.6}$$

where $-\infty < a < \infty$ and $b > 0$. In addition, $t_p = g(p; \boldsymbol{\theta})$, where t_p represent the TPF when the FPF rate is p. Using maximum likelihood estimates for $\boldsymbol{\theta}$, we can

estimate the equation for the ROC curve as well as the AUC and t_p for any chosen $FPF = p$. Confidence intervals for the true AUC and t_p can be found using the delta method. Gönen [11] gives sample code on how this can be done using SAS NLMIXED [12].

2.2 Epsilon-Skew-Normal (ESN) distribution

The ESN distribution, developed by Mudholkar and Hutson (2000) [10] has been used for extending common regression problems [13]. The standardized model for the ESN distribution ESN$(0, 1, \epsilon)$ is defined to be a unimodal distribution with the mode at 0 and probability mass $(1 - \epsilon)/2$ below the mode. The probability density function (p.d.f.), the distribution function (d.f.), and quantile function (q.f.) of its canonical form $ESN(0, 1, \epsilon)$ are respectively:

$$f_0(x) = \begin{cases} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2(1-\epsilon)^2}\right), & \text{if } x < 0, \\ \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2(1+\epsilon)^2}\right), & \text{if } x \geq 0, \end{cases} \tag{2.7}$$

$$F_0(x) = \begin{cases} (1-\epsilon)\Phi\left(\frac{x}{1-\epsilon}\right), & \text{if } x < 0, \\ -\epsilon + (1+\epsilon)\Phi\left(\frac{x}{1+\epsilon}\right), & \text{if } x \geq 0, \end{cases} \tag{2.8}$$

and,

$$Q_0(u) = F_0^{-1}(u) = \begin{cases} (1-\epsilon)\Phi^{-1}\left(\frac{u}{1-\epsilon}\right), & \text{if } 0 < u < (1-\epsilon)/2 \\ (1+\epsilon)\Phi^{-1}\left(\frac{u+\epsilon}{1+\epsilon}\right), & \text{if } (1-\epsilon)/2 \leq u < 1, \end{cases} \tag{2.9}$$

where $-1 < \epsilon < 1$, and $\Phi(x)$ denotes the standard normal c.d.f.

The standard epsilon-skew-normal distribution, $ESN(0, 1, \epsilon)$, is a mixture of two half-normal distributions and reduces to the standard normal distribution when $\epsilon = 0$. The distribution is skewed right for values of $\epsilon > 0$ and skewed left for values of $\epsilon < 0$. The limiting cases of (2.7) as $\epsilon \to \pm 1$ are the well known half-normal distributions. Figure 1 gives some typical ESN probability density functions.

The p.d.f. $f_0(\cdot)$ at (2.7) has derivatives of arbitrary orders. It is differentiable once at the mode.

The general form for the p.d.f., denoted $ESN(\mu, \sigma, \epsilon)$, is $f_0(\frac{x-\mu}{\sigma})/\sigma$, where $f_0(\cdot)$ is given by (2.7). Similarly, the general form for the c.d.f. of $ESN(\mu, \sigma, \epsilon)$ is $F_0(\frac{x-\mu}{\sigma})$, where $F_0(\cdot)$ is given by (2.8). Its quantile function, $Q(u) = \mu + \sigma Q_0(u)$, can be used to generate samples from the $ESN(\mu, \sigma, \epsilon)$ population. Note the relationship $Q(\frac{1-\epsilon}{2}) = \mu$.

The mean of $ESN(\mu, \sigma, \epsilon)$is given by

$$E(X) \;=\; \mu + \frac{4\sigma\epsilon}{\sqrt{2\pi}}, \tag{2.10}$$

and the variance is given by

$$\mathrm{Var}(X) \;=\; \frac{\sigma^2}{\pi}[(3\pi - 8)\epsilon^2 + \pi]. \tag{2.11}$$

Mudholkar and Hutson [10] provide methods for maximum likelihood estimation, which we will utilize in later sections.

3 Epsilon-Skew-Binormal ROC Model

We now extend the binormal ROC model using the ESN distribution as the underlying model for both healthy and diseased populations. As before, let Y be a continuous predictor for diagnosing disease, and let Y_0 and Y_1 denote the measurements of the predictor Y for healthy and diseased populations, respectively. Let $Y_0 \sim ESN(\mu_0, \sigma_0, \epsilon_0)$ and $Y_1 \sim ESN(\mu_1, \sigma_1, \epsilon_1)$ and $\mu_1 > \mu_0$. Let F_{ϵ_0} and F_{ϵ_1} be the c.d.f.'s for the $ESN(0, 1, \epsilon_0)$ and $ESN(0, 1, \epsilon_1)$ distributions, respectively. For a given cutoff c, the FPF is given by

$$FPF = P(Y_0 > c) = 1 - P(Y_0 \leq c) = 1 - F_{\epsilon_0}\left(\frac{c - \mu_0}{\sigma_0}\right). \tag{3.1}$$

Solving for c we have

$$c = \mu_0 + \sigma_0 F_{\epsilon_0}^{-1}(1 - FPF). \tag{3.2}$$

Similarly, for a given cutoff c, the TPF is given by

$$TPF = P(Y_1 > c) = 1 - P(Y_1 \leq c) = 1 - F_{\epsilon_1}\left(\frac{c - \mu_1}{\sigma_1}\right). \tag{3.3}$$

Substituting (3.2) into (3.3) we have

$$TPF = 1 - F_{\epsilon_1}\left[-\frac{(\mu_1 - \mu_0)}{\sigma_1} + \frac{\sigma_0}{\sigma_1} F_{\epsilon_0}^{-1}(1 - FPF)\right] \tag{3.4}$$

Thus the equation for the epsilon-skew-binormal ROC curve is given by

$$h(x; \boldsymbol{\theta}) = 1 - F_{\epsilon_1}\left[-a + bF_{\epsilon_0}^{-1}(1 - x)\right], \tag{3.5}$$

where $\boldsymbol{\theta} = (\mu_0, \sigma_0, \epsilon_0, \mu_1, \sigma_1, \epsilon_1)$, $a = \frac{\mu_1 - \mu_0}{\sigma_1}$, and $b = \frac{\sigma_0}{\sigma_1}$.

To find the area under the ROC curve (AUC), we need to calculate $P(Y_1 > Y_0)$. Unlike in the case of normal distributions, linear combinations of ESN distributions are not ESN. Hence, there is no closed form solution to the AUC in this case. Alternatively, AUC can be expressed as

$$AUC = \int_0^1 h(x; \boldsymbol{\theta}) \mathrm{d}x. \tag{3.6}$$

However, numerical approaches such as the trapezoidal rule [14] provide good approximations for the AUC. We examined the relationship between AUC for the binormal model and the AUC for the ESBN model using Taylor series expansion in Appendix A. In addition, $t_p = h(p; \boldsymbol{\theta})$. Given the maximum likelihood estimates for $\boldsymbol{\theta}$, we can then estimate the equation for the ROC curve as well as the AUC and t_p for any chosen $FPF = p$.

Let $\hat{\boldsymbol{\theta}}$ be the maximum likelihood estimator of $\boldsymbol{\theta}$. For maximum likelihood estimation, Mudholkar and Hutson [10] proved in Theorem 4.7 that as $n \to \infty$, the

maximum likelihood estimator $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ is a centered multivariate normal distribution with variance covariance matrix

$$\boldsymbol{\Sigma} = \left(\begin{array}{c|c} \boldsymbol{I_0} & \mathbf{0} \\ \hline \mathbf{0} & \boldsymbol{I_1} \end{array}\right) \tag{3.7}$$

where

$$\boldsymbol{I_0} = \begin{pmatrix} I^{\mu_0\mu_0} = \frac{3\pi(1-\epsilon_0^2)\sigma_0^2}{3\pi-8} & I^{\mu_0\sigma_0^2} = 0 & I^{\mu_0\epsilon_0} = \frac{-2\sqrt{2\pi}(1-\epsilon_0^2)\sigma_0}{3\pi-8} \\ & I^{\sigma_0^2\sigma_0^2} = 2\sigma_0^4 & I^{\sigma_0^2\epsilon_0} = 0 \\ & & I^{\epsilon_0\epsilon_0} = \frac{\pi(1-\epsilon_0^2)}{3\pi-8} \end{pmatrix}, \tag{3.8}$$

and

$$\boldsymbol{I_1} = \begin{pmatrix} I^{\mu_1\mu_1} = \frac{3\pi(1-\epsilon_1^2)\sigma_1^2}{3\pi-8} & I^{\mu_1\sigma_1^2} = 0 & I^{\mu_1\epsilon_1} = \frac{-2\sqrt{2\pi}(1-\epsilon_1^2)\sigma_1}{3\pi-8} \\ & I^{\sigma_1^2\sigma_1^2} = 2\sigma_1^4 & I^{\sigma_1^2\epsilon_1} = 0 \\ & & I^{\epsilon_1\epsilon_1} = \frac{\pi(1-\epsilon_1^2)}{3\pi-8} \end{pmatrix}. \tag{3.9}$$

In the context of the ESBN modeling the parameters ϵ_0 and ϵ_1 may be interpreted as a measure of distance of the distribution of the continuous predictor from normality with respect to ESN alternative for individuals with negative and positive disease status, respectively. Hence, the tests

$$\begin{array}{ccc} H_0 : \epsilon_0 = 0 & & H_0 : \epsilon_1 = 0 \\ H_1 : \epsilon_0 \neq 0 & \text{and} & H_1 : \epsilon_1 \neq 0 \end{array} \tag{3.10}$$

may be used as a diagnostic tool with respect to the appropriateness of testing the appropriateness of an underlying normal model versus a broad class of skew normal alternatives.

3.1 Limiting Cases as $\epsilon \longrightarrow \pm 1$

Lemma 4.2 in Muholkar and Hutson [10] gives the conditions in which the maximum likelihood estimator for ϵ is ± 1. We need to examine (3.5) in these cases. If $\hat{\epsilon}_0 = \pm 1$

we find that $F^{-1}_{\hat{\epsilon}_0}(1-x)$ is always defined for $0 < x < 1$. However, $-\hat{a}+\hat{b}F^{-1}_{\hat{\epsilon}_0}(1-x)$ is not defined for $x \geq P(Y_0 > \hat{\mu}_1)$ when $\hat{\epsilon}_1 = 1$ and for $x < P(Y_0 > \hat{\mu}_1)$ when $\hat{\epsilon}_1 = -1$. It can be shown that in the case of $\hat{\epsilon}_1 = 1$ that the equation for the ROC curve is given by:

$$h(x;\hat{\boldsymbol{\theta}}) = \begin{cases} 2 - 2\Phi\left(\frac{1}{2}[-\hat{a} + \hat{b}F^{-1}_{\hat{\epsilon}_0}(1-x)]\right), & \text{if } x < P(Y_0 > \hat{\mu}_1), \\ 1, & \text{if } x \geq P(Y_0 > \hat{\mu}_1), \end{cases} \tag{3.11}$$

and in the case of $\hat{\epsilon}_1 = -1$,

$$h(x;\hat{\boldsymbol{\theta}}) = \begin{cases} 0, & \text{if } x < P(Y_0 > \hat{\mu}_1), \\ 1 - 2\Phi\left(\frac{1}{2}[-\hat{a} + \hat{b}F^{-1}_{\hat{\epsilon}_0}(1-x)]\right), & \text{if } x \geq P(Y_0 > \hat{\mu}_1), \end{cases} \tag{3.12}$$

Expressions for $h(x;\hat{\boldsymbol{\theta}})$ in equations (3.11) and (3.12) are based on the c.d.f for the half-normal distributions.

3.2 Estimating AUC of the ESBN ROC Curve

As noted earlier there is not an analytic form for the AUC from the ESBN ROC curve. Here we propose utilizing the trapezoid methods. Let $\hat{\boldsymbol{\theta}}$ be the maximum likelihood estimate for $\boldsymbol{\theta}$. Then

$$\widehat{AUC} = \int_0^1 h(x;\hat{\boldsymbol{\theta}})\mathrm{d}x, \tag{3.13}$$

which can be approximated by

$$\widehat{AUC}_T = \frac{1}{B}\left[\frac{1}{2} + \sum_{i=1}^{B-1} h(i/B;\hat{\boldsymbol{\theta}})\right], \tag{3.14}$$

where B is the number of subintervals of $[0, 1]$.

Theorem 1. *For large samples*

$$\widehat{AUC}_T \sim AN(AUC_T, \frac{1}{B^2}\mathbf{1}'_{B-1}\boldsymbol{V}\mathbf{1}_{B-1}), \tag{3.15}$$

where AUC_T *is* $\widehat{AUC}_T$ *evaluated at* $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}$, $\mathbf{1}_{B-1}$ *is a* $(B-1)\times 1$ *vector of 1's, and* $\boldsymbol{V}$ *is the variance-covariance matrix of the* $h(i/B;\hat{\boldsymbol{\theta}}), i = 1\ldots B-1$.

Proof. Proof. By Theorem 4.7 in Mudholkar and Hutson [10] we have $\hat{\boldsymbol{\theta}} \sim AN(\theta, \frac{1}{n}\boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is given in equation (3.7). By Theorem A, section 3.3 in Serfling [15] we have

$$h(i/B;\hat{\boldsymbol{\theta}}) \sim AN(h(i/B;\boldsymbol{\theta}), \frac{1}{n}[\nabla h_i(\boldsymbol{\theta})]'\boldsymbol{\Sigma}[\nabla h_i(\boldsymbol{\theta})]), \tag{3.16}$$

where $\nabla h_i(\boldsymbol{\theta})$ represents the 6×1 vector of partial derivatives of $h(i/B;\hat{\boldsymbol{\theta}})$ evaluated at $\boldsymbol{\theta}$. Let $\boldsymbol{V}$ is the variance-covariance matrix of the $h(i/B;\hat{\boldsymbol{\theta}}), i = 1\ldots B-1$, and $\boldsymbol{h}_{\hat{\boldsymbol{\theta}}}$ be the $(B-1)\times 1$ vector of the $h(i/B;\hat{\boldsymbol{\theta}}), i = 1\ldots B-1$. Then we have

$$\begin{aligned}\widehat{AUC}_T &= \frac{1}{B}\left[\frac{1}{2} + \sum_{i=1}^{B-1} h(i/B;\hat{\boldsymbol{\theta}})\right], \\ &= \frac{1}{2B} + \frac{1}{B}\mathbf{1}'_{B-1}\boldsymbol{h}_{\hat{\boldsymbol{\theta}}}.\end{aligned}$$

Thus we have that $\widehat{AUC}_T$ is asymptotically normal with

$$E(\widehat{AUC}_T) = AUC_T,$$

and

$$\mathrm{Var}(\widehat{AUC}_T) = \frac{1}{B^2}\mathbf{1}'_{B-1}\boldsymbol{V}\mathbf{1}_{B-1}.$$

□

Theorem 1 can then be used for estimating $100(1-\alpha)\%$ confidence intervals.

The general approach for determining starting values for the maximum likelihood estimation method is to first fit the standard binormal model ($\epsilon_0 = \epsilon_1 = 0$). The parameter estimates from the binormal model are then used as the starting values for μ_0, μ_1, σ_0 and σ_1 with ϵ_0 and ϵ_1 set to zero as the starting values.

4 Simulation

We conducted a simulation study in order to examine the behavior of the maximum likelihood based estimates of the AUC based on the binormal and ESBN ROC curves and two possible prediction thresholds over a variety of model scenarios. Let $Y_0 \sim ESN(0, 1, \epsilon_0)$ and $Y_1 \sim ESN(\mu_1, 1.5, \epsilon_1)$ with sample size $n_0 = n_1 = 100$, respectively. We set $\mu_1 = 1$ and $\mu_1 = 1.5$ and let ϵ_0 and ϵ_1 range from -0.75 to 0.75 by 0.25 over all positive, negative, and zero combinations. In total there were 98 model combinations. We utilized 1000 simulations under each scenario to compute the expected AUC, $t_{0.05}$, and $t_{0.10}$, where t_p represent the true positive fraction when the false positive fraction is p.

For the binormal model we calculated $\widehat{AUC} = \Phi\left(\frac{\hat{a}}{\sqrt{1+\hat{b}}}\right)$, $\hat{t}_{0.05} = \Phi\left(\hat{a} + \hat{b}\Phi^{-1}(0.05)\right)$, and $\hat{t}_{0.1} = \Phi\left(\hat{a} + \hat{b}\Phi^{-1}(0.1)\right)$. When fitting the ESBN model, we calculated (3.14), $\hat{t}_{0.05} = 1 - F_{\hat{\epsilon}_1}[-\hat{a} + \hat{b}F^{-1}_{\hat{\epsilon}_0}(0.95)]$, and $\hat{t}_{0.05} = 1 - F_{\hat{\epsilon}_1}[-\hat{a} + \hat{b}F^{-1}_{\hat{\epsilon}_0}(0.90)]$. Estimates for the AUC under binormal model and $t_{0.05}$, and $t_{0.10}$ under both models were easily calculated using delta method from SAS PROC NLMIXED [12]. To estimate AUC under ESBN model, we used the approach outlined in section 3.2, equation (3.14) with $B = 10000$.

Figure 2 shows plots of the observed versus expected AUCs, $t_{0.05}$, and $t_{0.10}$ for $\mu_1 = 1$ under the variety of ϵ_0 and ϵ_1 combinations. We see that while both models appear to be similar in terms of providing relatively unbiased estimates of the AUC, the binormal model is clearly biased when estimating pointwise TPF and FPF. While not shown, plots were similar for $\mu_1 = 1.5$.

Figure 3 illustrates typical ROC plots under a variety of scenarios. In each case, the epsilon-skew-binormal ROC curve follows the true ROC curve. However, when departures from normality are observed, we find that the binormal ROC curve sometimes crosses the true curve, part of the curve below the true ROC curve and

part of the curve above the ROC curve. In effect, AUCs are similar due to lost area from curve below canceling out the gained area from the curve above. However, the bias across the set of potential decision thresholds is obviously inherent in the binormal model. In practice this would yield either an inflated number of true positives or false positives relative to the true population distribution.

We conducted a second simulation study to compare the AUC from the nonparametric ROC method to the binormal and epsilon-skew-binormal models for a select number of model scenarios. As in the previous simulation, we utilized 1000 simulations with $Y_0 \sim ESN(0, 1, \epsilon_0)$ and $Y_1 \sim ESN(\mu_1, 1.5, \epsilon_1)$ and sample size $n_0 = n_1 = 100$, respectively. We set $\mu_1 = 1$ and let ϵ_0 and ϵ_1 range from -0.25 to 0.25 by 0.25 over all positive, negative, and zero combination. In total there were nine model combinations. Table 1 summarizes our results. We find that results for calculating the empirical AUC and the epsilon-skew-binormal models are similar in terms of bias and variance. This provides evidence that the ESBN model provides the benefits of a parametric model in terms of smoothness, yet is relatively robust and efficient as compared to the nonparametric model. Our example in the next section also provides a good illustration in terms of the relative similarities of ESBN and nonparametric approaches, particularly in terms of the pointwise estimates.

We illustrate the maximum likelihood estimation of the ESBN model using skeletal measurements in 507 physically active individuals [16] as compared to binormal and nonparametric models. Heinz et al. examined several skeletal measurements as possible predictors for sex using the binormal model. The relevance of these types of biometric measures in terms of prediction includes archeological and forensics applications. We chose chest depth between spine and sternum at nipple level as a predictor for sex. There were 247 men and 260 women in the original study.

Figure 5 gives histograms of the chest depth measurements for males and females with overlays of the fitted normal and ESN distributions. When fitting the ESN

distributions to the data, $\hat{\epsilon}_0 = 0.47$ and $\hat{\epsilon}_1 = 0.24$. For testing $H_0 : \epsilon_0 = 0$ and $H_1 : \epsilon_1 = 0$, p-values from the large sample approximation at (3.10) were < 0.0001 and 0.007, respectively, indicating that the binormal distribution is not a good fit.

We fit the nonparametric, binormal, and ESBN models to the chest depth measurements to predict sex. Using the nonparametric model, the estimated AUC was 0.8734 with standard error 0.0154. Under the binormal model, estimated AUC was 0.8633 with standard error of 0.0158. With the ESBN model, estimated AUC was 0.8649 with standard error of 0.0155. As might be expected from our simulations estimating AUC is comparable for all three methods. However, when estimating possible cut-offs, we find that the estimated $t_{0.05}$ is 0.5138 (s.e. 0.0412) for the binormal model and 0.4204 (s.e. 0.0494) for the ESBN model. For $t_{0.10}$ is 0.6350 (s.e. 0.0365) for the binormal model and 0.5751 (s.e. 0.0464) for the ESBN model. We see that while the estimated AUC is similar for all three models, there is likely bias present when estimating pointwise TPF and FPF. Figure 4 shows the estimated ROC curves for the binormal, ESBN, and nonparametric models. We see from figure 4 a crossing of the binormal and ESBN. In addition the ESBN model follows the empirical model closely, indicating the ESBN model is a reasonable model choice and illustrating potential bias of the binormal model.

5 Conclusions

In this note we have extended the binormal model via utilization of the epsilon-skew-normal family of distributions. We illustrated the following:

1. If the binormal assumptions are correct, the epsilon-skew-binormal model verifies the model fit.

2. AUC estimates are similar between binormal, ESBN and nonparametric mod-

els.

3. If the ESBN model is correct ($\epsilon_0, \epsilon_1 \neq 0$), then the binormal is biased in terms of estimating TPF and FPF.

4. The epsilon-skew-binormal model is comparable to nonparametric model in terms of the unbiased AUC estimates and relative efficiency.

Overall, the epsilon-skew-binormal ROC model provides a robust, easily implemented, and smooth parametric alternative to the binormal and nonparametric models.

Appendix A

We first note that

$$\frac{\partial F_{\epsilon_1}(x)}{\partial \epsilon_1} = \begin{cases} \frac{x}{1-\epsilon_1}\ \phi\left(\frac{x}{1-\epsilon_1}\right) - \Phi\left(\frac{x}{1-\epsilon_1}\right), & \text{if } x < 0 \\ \frac{-x}{1+\epsilon_1}\ \phi\left(\frac{x}{1+\epsilon_1}\right) - \Phi\left(\frac{-x}{1+\epsilon_1}\right), & \text{if } x \geq 0, \end{cases} \tag{5.1}$$

and

$$\frac{\partial F_{\epsilon_0}^{-1}(1-u)}{\partial \epsilon_0} = \begin{cases} \frac{1-u}{(1-\epsilon_0)\phi\left[\Phi^{-1}\left(\frac{1-u}{1-\epsilon_0}\right)\right]} - \Phi^{-1}\left(\frac{1-u}{1-\epsilon_0}\right), & \text{if } 0 < u \leq \frac{1+\epsilon_0}{2} \\ \frac{u}{(1+\epsilon_0)\phi\left[\Phi^{-1}\left(1-\frac{u}{1+\epsilon_0}\right)\right]} + \Phi^{-1}\left(1-\frac{u}{1+\epsilon_0}\right), & \text{if } \frac{1+\epsilon_0}{2} < u < 1, \end{cases} \tag{5.2}$$

In section 3, the equation of the ROC curve is given by

$$h(x) = 1 - F_{\epsilon_1}[-a + bF_{\epsilon_0}^{-1}(1-x)], \tag{5.3}$$

where $a = \frac{\mu_1 - \mu_0}{\sigma_1}$ and $b = \frac{\sigma_0}{\sigma_1}$.

Let $\boldsymbol{\epsilon} = (\epsilon_0, \epsilon_1)'$, and let $\zeta = a + b\Phi^{-1}(x)$.

Taylor's series of (3.5) expanded around $\epsilon_0 = 0$ and $\epsilon_1 = 0$ is given by

$$h_{\boldsymbol{\theta}}(x) \approx \Phi(\zeta) + A_0\epsilon_0 + A_1\epsilon_1 \tag{5.4}$$

where

$$A_0 = \begin{cases} -\phi(\zeta) \cdot \left[\frac{1-x}{\phi[\Phi^{-1}(x)]} + \Phi^{-1}(x)\right], & \text{if } 0 < x \leq \frac{1}{2} \\ -\phi(\zeta) \cdot \left[\frac{x}{\phi[\Phi^{-1}(x)]} - \Phi^{-1}(x)\right], & \text{if } \frac{1}{2} < x < 1, \end{cases} \tag{5.5}$$

and

$$A_1 = \begin{cases} \zeta\phi(\zeta) - \Phi(\zeta), & \text{if } 0 < x \leq \Phi(\frac{-a}{b}) \\ -\zeta\phi(\zeta) + \Phi(\zeta), & \text{if } \Phi(\frac{-a}{b}) < x < 1, \end{cases} \tag{5.6}$$

Thus

$$AUC_{ESN} \approx AUC_{NOR} + \epsilon_0 \int_0^1 A_0 \mathrm{d}\, x + \epsilon_1 \int_0^1 A_1 \mathrm{d}\, x. \tag{5.7}$$

References

[1] Metz CE, Herman BA, Shen J. Maximum liklihood estimation of receiver operating characteristic (ROC) curves from continuously-distributed data. *Statistics in Medicine* 1998; **17**:1033–1053.

[2] Pepe MS, Zheng Y, Jin Y, Huang Y, Parikh CR, Levy WC. Evaluating the ROC performance of markers for future events. *Lifetime Data Analysis* 2008; **14**(1):86–113, DOI: 10.1007/s10985-007-9073-x

[3] Hillis SL, Berbaum KS, Metz CE. Recent developments in the Dorfman-Berbaum-Metz procedure for multireader ROC study analysis *Academic Radiology* 2008; **15**(5):647–661, DOI: 10.1016/j.acra.2007.12.015

[4] Gu JZ, Ghosal S, Roy A. Bayesian bootstrap estimation of ROC curve. *Statistics in Medicine* 2009; **27**(26):5407–5420, DOI: 10.1002/sim.3366

[5] Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; **143**:29–36.

[6] Shibata Y, Yamamoto T, Takano S, Katayama W, Takeda T, Matsumura A. Direct comparison of thallium-201 and technetium-99m MIBI SPECT of a glioma by receiver operating characteristic analysis. *Journal of Clinical Neurosciences* 2009; **16**:264–269, DOI: 10.1016/j.jocn.2008.04.010.

[7] Wei L, Yang Y, Nishikawa RM. Microcalcification classification assisted by content-based image retrieval for breast cancer diagnosis. *Pattern Recognition* 2009; **42**:1126–1132, DOI: 10.1016/j.patcog.2008.08.028.

[8] Xi LF, Kiviat NB, Galloway DA, Zhou X, Ho J, Koutsky LA. Effect of cervical cytologic status on the association between human papillomavirus type 16

DNA load and the risk of cervical intraepithelial neoplasia grade 3. *Journal of Infectious Diseases* 2008; **198**:324–331, DOI: 10.1086/589715.

[9] Walsh SJ. Limitations to the robustness of binormal ROC curves: effects of model misspecification and location of decision thresholds on bias, precision, size, and power. *Statistics in Medicine* 1997; **16**:669-679.

[10] Mudholkar GS, Hutson AD. The epsilon-skew-normal distribution for analyzing near-normal data. *Journal of Statistical Planning and Inference* 2000; **83**(2):291–309.

[11] Gönen M. Analyzing Receiver Operating Characteristic Curves with SAS. Cary, NC: SAS Institute Inc., 2007.

[12] SAS (v9.1) SAS Institute Inc., Cary, NC, USA.

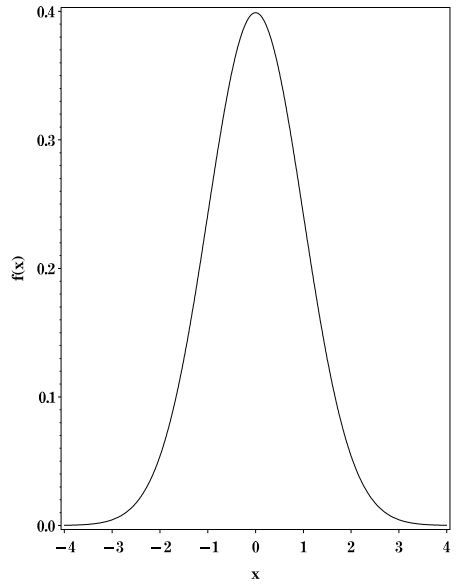
[13] Hutson AD. Utilizing the flexibility of the epsilon-skew-normal normal distribution for common regression problems. *Journal of Applied Statistical Sciences* 2004; **31**:673–683. DOI: 10.1080/1478881042000214659

[14] Margolis DJ, Bilker W, Boston R, Localio R, Berlin JA. Statistical characteristics of area under the receiver operating characteristic curve for a simple prognostic model using traditional and bootstrapped approaches. *Journal of Clinical Epidemiology* 2002; **55**:518–524.

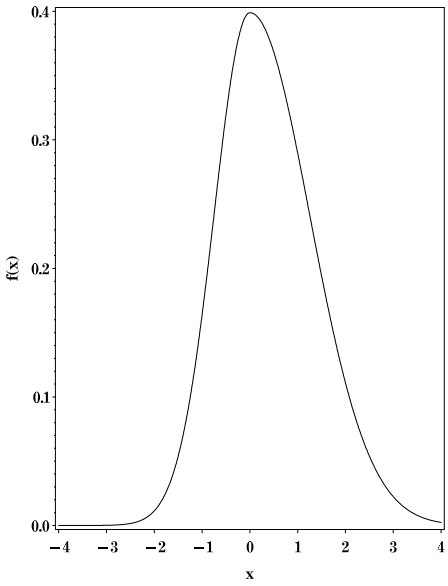
[15] Serfling RJ. *Approximation Theorems of Mathematical Statistics* Wiley: New York, 1980.

[16] Heinz G, Peterson LJ, Johnson RW, Kerk CJ. Exploring relationships in body dimensions. *Journal of Statistics Education* 2003; **11**(2): www.amstat.org/publications/jse/v11n2/datasets.heinz.html

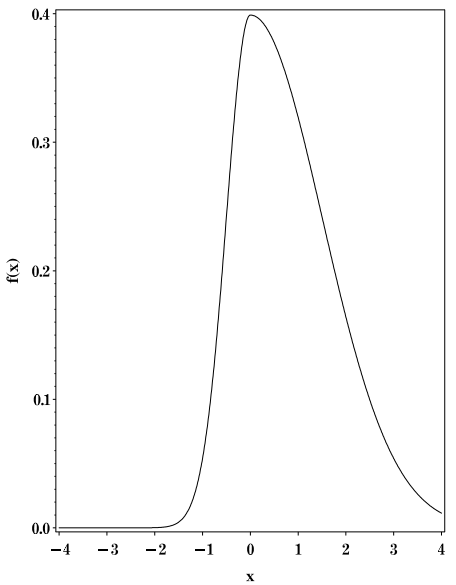
Figure 1: Some typical $ESN(0, 1, \epsilon)$ probability density functions.



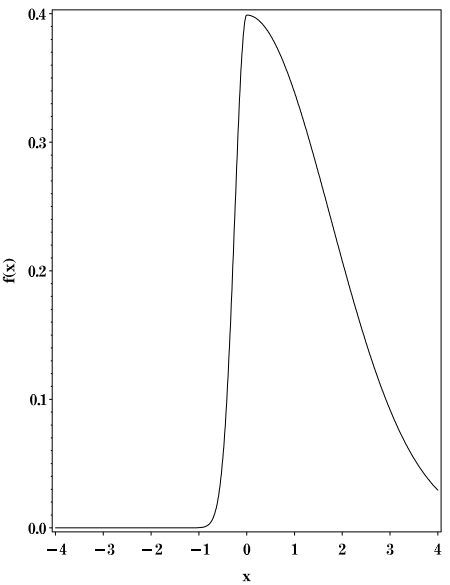
(a) $\epsilon = 0$



(b) $\epsilon = 0.25$



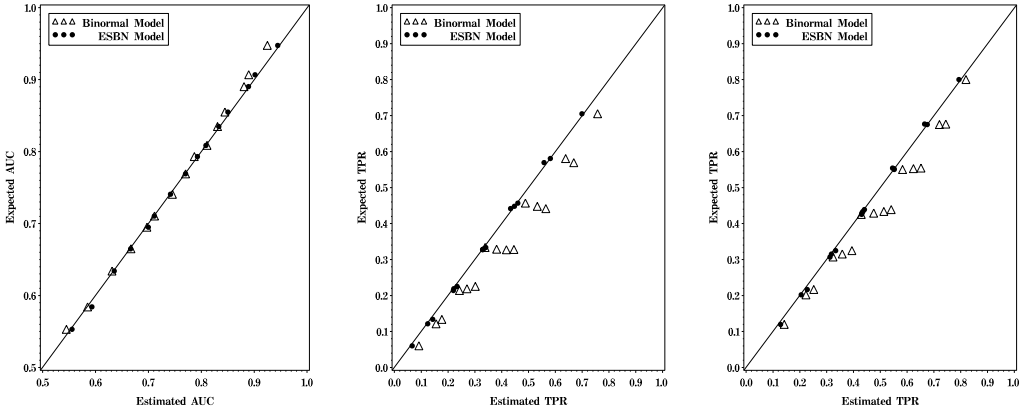
(c) $\epsilon = 0.50$



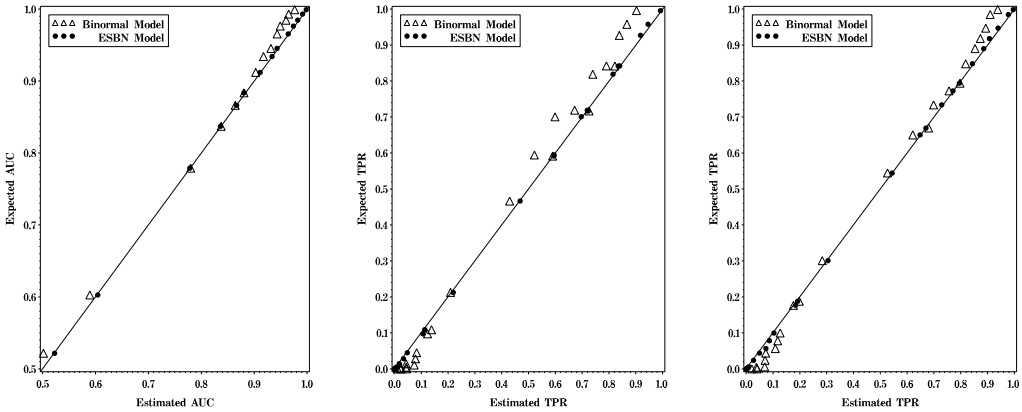
(d) $\epsilon = 0.75$

Figure 2: Plots of expected versus estimated AUC, $t_{0.05}$, and $t_{0.10}$ for binormal versus ESBN models

$\epsilon_0, \epsilon_1 \in [0, 0.75]$



$\epsilon_0 \in [-0.75, 0), \epsilon_1 \in (0, 0.75]$ or $\epsilon_0 \in (0, 0.75], \epsilon_1 \in [-0.75, 0)$



$\epsilon_0, \epsilon_1 \in [-0.75, 0)$

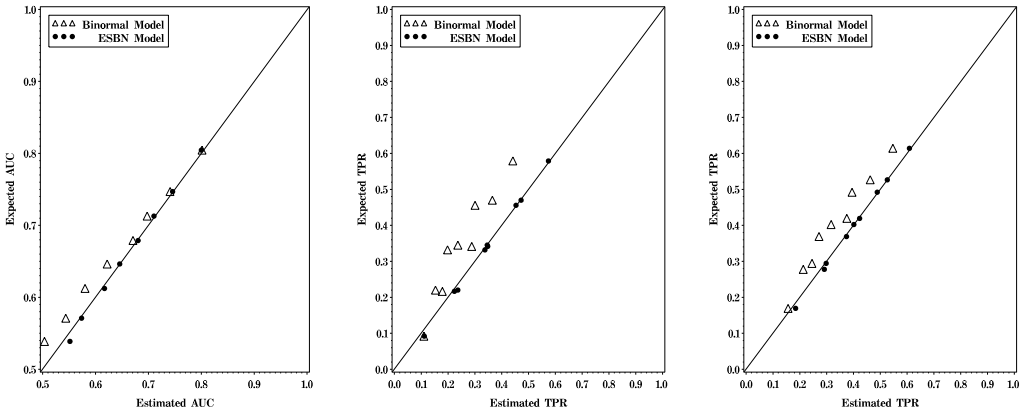
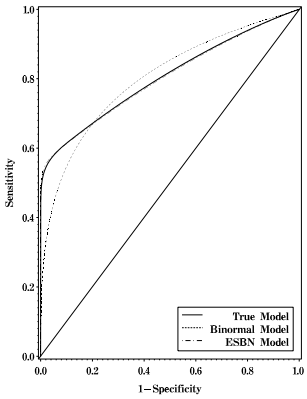
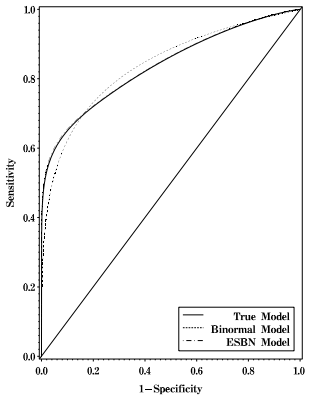


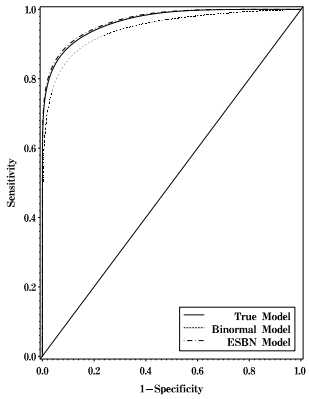
Figure 3: Some typical ROC curves (binormal versus ESBN)



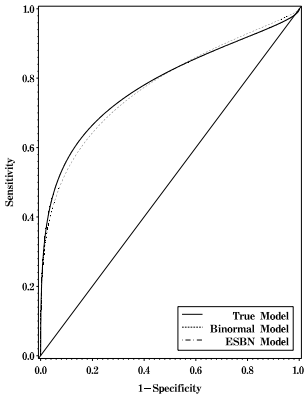
(a) $\epsilon_0 < 0, \epsilon_1 < 0, \delta = 1$



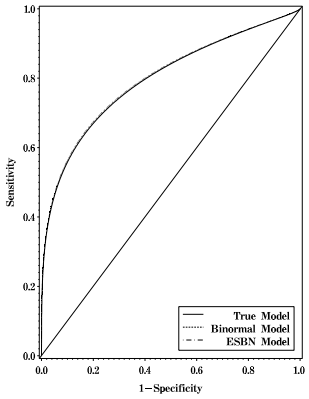
(b) $\epsilon_0 < 0, \epsilon_1 = 0, \delta = 1$



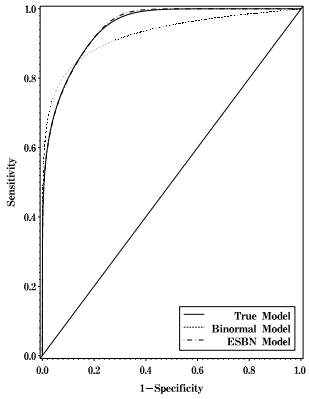
(c) $\epsilon_0 < 0, \epsilon_1 > 0, \delta = 1$



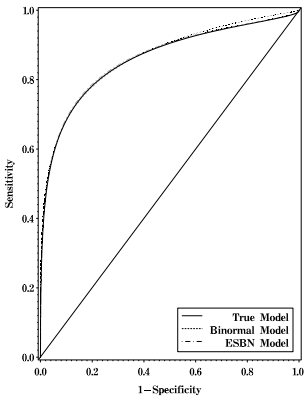
(d) $\epsilon_0 = 0, \epsilon_1 < 0, \delta = 2$



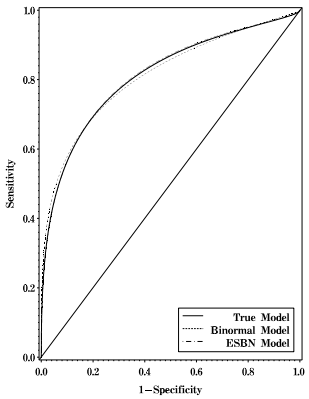
(e) $\epsilon_0 = 0, \epsilon_1 = 0, \delta = 1.5$



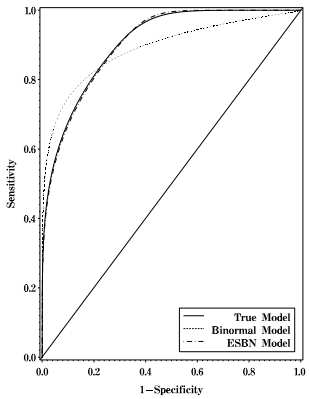
(f) $\epsilon_0 = 0, \epsilon_1 > 0, \delta = 1$



(g) $\epsilon_0 > 0, \epsilon_1 < 0, \delta = 3$



(h) $\epsilon_0 > 0, \epsilon_1 = 0, \delta = 2$



(i) $\epsilon_0 > 0, \epsilon_1 > 0, \delta = 1$

Figure 4: Chest depth as a predictor of gender using ESBN, binormal and nonparametric models

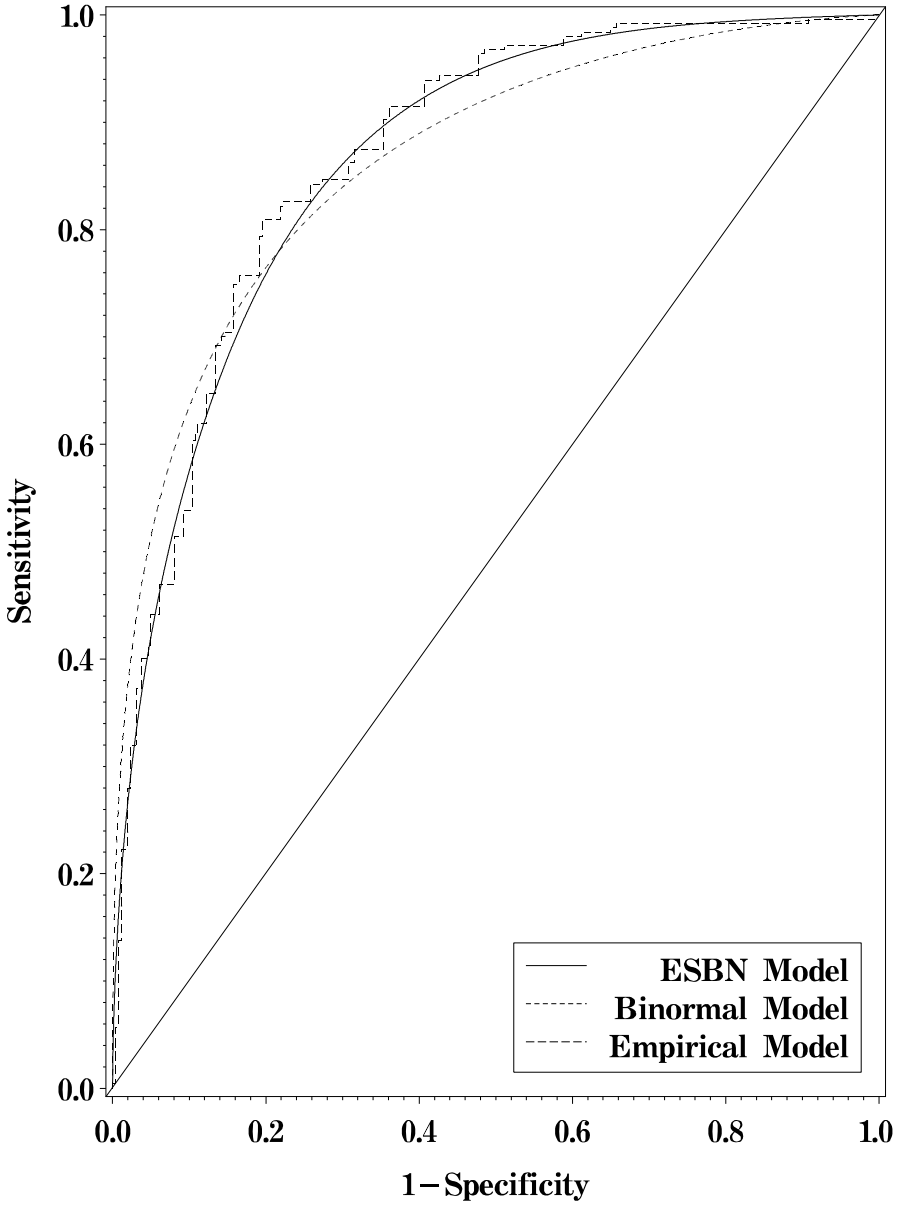
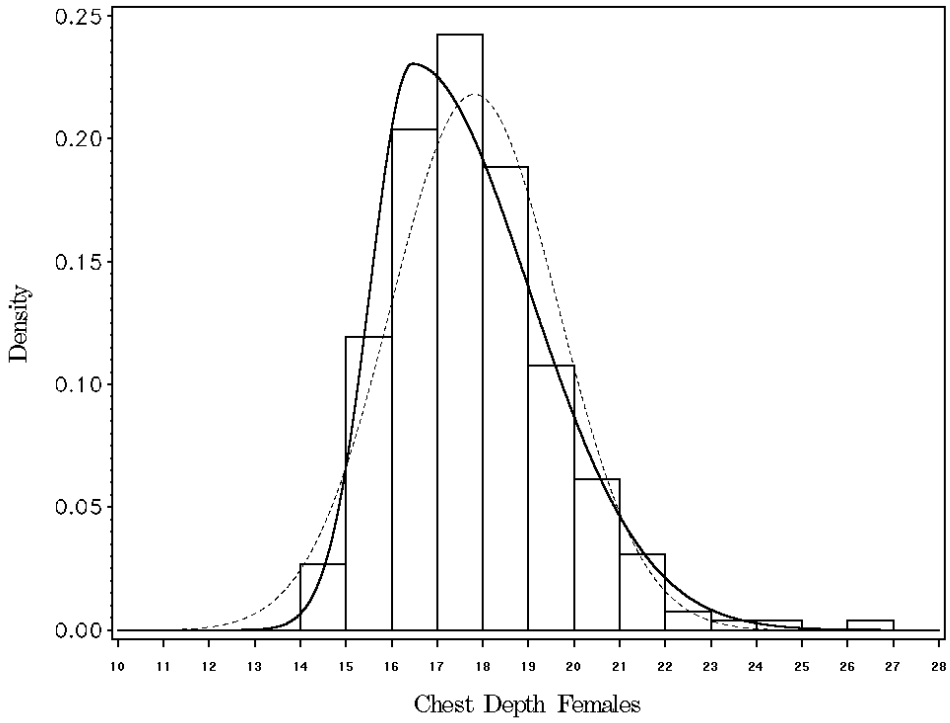


Figure 5: Histogram of chest depth measurements by sex



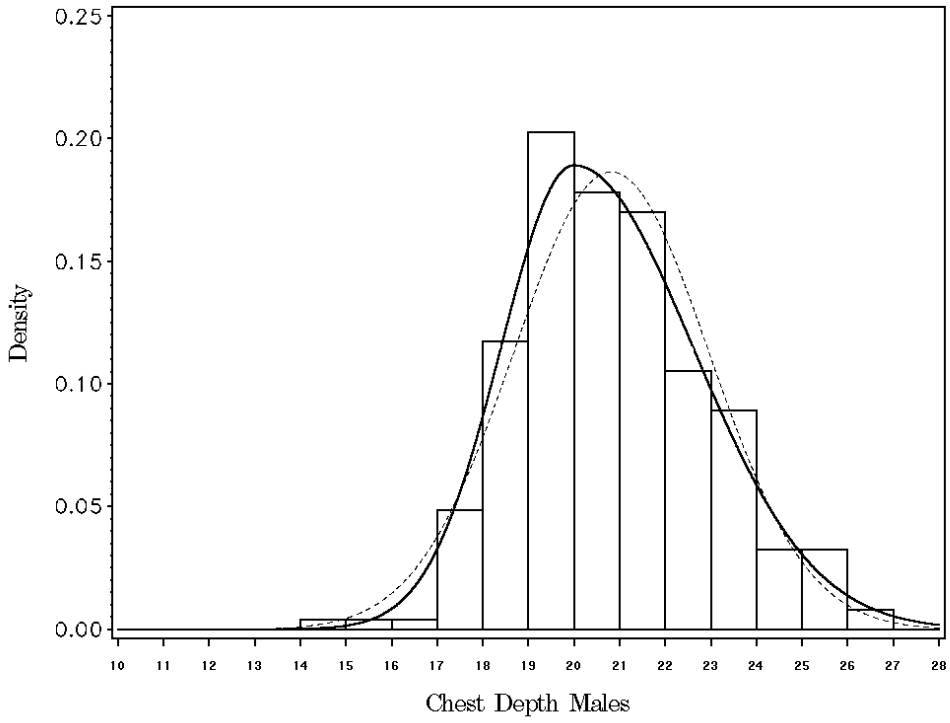


Table 1: Comparison of AUC for binormal, ESBN, and nonparametric models

ϵ_0	ϵ_1	Expected AUC	Binormal	ESBN	Nonparametric
-0.25	-0.25	0.679	0.670 ± 0.040	0.679 ± 0.039	0.679 ± 0.040
-0.25	0	0.779	0.781 ± 0.032	0.780 ± 0.033	0.779 ± 0.034
-0.25	0.25	0.866	0.863 ± 0.022	0.865 ± 0.025	0.866 ± 0.024
0	-0.25	0.603	0.588 ± 0.041	0.603 ± 0.040	0.602 ± 0.041
0	0	0.710	0.710 ± 0.038	0.711 ± 0.038	0.710 ± 0.038
0	0.25	0.809	0.810 ± 0.028	0.808 ± 0.032	0.808 ± 0.031
0.25	-0.25	0.522	0.500 ± 0.040	0.522 ± 0.041	0.531 ± 0.035
0.25	0	0.634	0.630 ± 0.039	0.633 ± 0.041	0.633 ± 0.040
0.25	0.25	0.741	0.743 ± 0.033	0.740 ± 0.035	0.739 ± 0.035