AUC ESTIMATION
UNDER VARIOUS SURVIVAL MODELS

by

Fazhe Chang

A Dissertation Submitted to the Faculty of

The Charles E. Schmidt College of Science

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

Florida Atlantic University

Boca Raton, FL

August 2012

AUC ESTIMATION

UNDER VARIOUS SURVIVAL MODELS

by

Fazhe Chang

This dissertation was prepared under the direction of the candidate's dissertation advisor, Dr. Lianfen Qian, Department of Mathematical Sciences, and has been approved by the members of his supervisory committee. It was submitted to the faculty of the Charles E. Schmidt College of Science and was accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

SUPERVISORY COMMITTEE:

Lianfen Qian, Ph.D.
Dissertation Advisor

Heinrich Niederhausen, Ph.D.

Lee Klingler, Ph.D.

Hongwei Long, Ph.D.

Lee Klingler, Ph.D.
Chair, Department of Mathematical Sciences

Gary W. Perry, Ph.D.
Dean, The Charles E. Schmidt College of Science

Barry T. Rosson, Ph.D.
Dean, Graduate College

July 12, 2012
Date

ACKNOWLEDGEMENTS

I would first like to thank my disserattion advisor Dr. Lianfen Qian. I can not imagine how I can finish this dissertation without her untiring patience and encouragement.

I would like to acknowledge my gratitude to my committee members: Dr. Niederhausen, Dr. Klingler and Dr. Long.

At last, I have to admit that my wife Kangli Ren endless love gives me braveness and power to overcome many unexpected difficulties.

ABSTRACT

Author:	Fazhe Chang
Title:	AUC Estimation Under Various Survival Models
Institution:	Florida Atlantic University
Dissertation Advisor:	Dr. Lianfen Qian
Degree:	Doctor of Philosophy
Year:	2012

In the medical science, the receiving operating characteristic (ROC) curve is a graphical representation to evaluate the accuracy of a medical diagnostic test for an any cut-off point. The area under the ROC curve (AUC) is an overall performance measure for a diagnostic test. There are two parts in this dissertation. In the first part, we study the properties of bi-Exponentiated Weibull models. First, we derive a general moment formula for single Exponentiated Weibull models. Then we move on to derive the precise formula of AUC and study the maximum likelihood estimation (MLE) of the AUC. Finally, we obtain the asymptotic distribution of the estimated AUC. Simulation studies are used to check the performance of MLE of AUC under the moderate sample sizes. The second part of the dissertation is to study the estimation of AUC under the crossing model, which extends the AUC formula in Gonen and Heller (2007).

AUC ESTIMATION

UNDER VARIOUS SURVIVAL MODELS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

1.1 BACKGROUND

In medical science, the accuracy of a medical diagnostic test is typically evaluated by sensitivity and specificity. A test result is called a biomarker (Hanley, 1989) as an indicator of disease status for subjects. The Receiver Operating Characteristic (ROC) curve is a graphical representation of the relationship between sensitivity and specificity of a biomarker. The area under the ROC curve (AUC) is an overall performance measure for a biomarker. Non-parametric and semi-parametric approaches have been proposed to estimate the ROC curve and AUC in many papers. Fushing and Bruce (1996) summarized the estimators and their asymptotic properties of ROC curve and AUC using the two fore-mentioned approaches. Suppose there are two study groups: diseased and non-diseased, then AUC can be computed as a probability: $P(X > Y)$, where X is a biomarker from the diseased group and Y is a biomarker from the non-diseased group.

In ROC analysis, one of the most popular models is called a bi-normal model. It assumes the existence of a monotonic increasing transformation such that the biomarker can be transformed into a normal random variable for both diseased and non-diseased subjects. ROC curves of a biomarker under a monotonic increasing transformation keep invariant. A good estimator of ROC curves should satisfy this invariance property. Many results on ROC curve analysis are based on above bi-normal assumption.

The literature for ROC analysis under the bi-normal model is extensive: Green and Swets (1966) used the ROC method in the signal detection theory. Hanley (1989) gave a comprehensive review on ROC methodology. Metz (1998) grouped continuous test data into ordinal-scale categorical data and used the maximum likelihood method to estimate parameters by assuming the continuous variable can be approximated by a discrete variable in finite support points. Zou and Hall (2000) developed the MLE algorithms to estimate ROC curve under an unspecific or a specific monotonic transformation. Gu, Ghosal and Roy (2005) introduced a non-parametric method based on the Bayesian bootstrap technique and compared it with several different methods (semi-parametric and non-parametric) to estimate ROC curves. Pepe and Cai (2004) considered the ROC curve as the distribution of placement values and then estimated the ROC curve by maximizing the pseudo-likelihood function of the estimated placement values. All the above methods require complicated computations. Zhou and Lin (2008) proposed a semi-parametric maximum likelihood (ML) estimate of ROC which satisfies the property of invariance. Their method largely reduced the number of nuisance parameters, therefore computational complexity. Simulation studies showed that the proposed estimator is efficient and robust.

However, in some circumstances, the distribution of the biomarker may be skewed and the normality assumption can be not reasonable. Under this situation, other models such as Gamma, Weibull and Exponentiated Weibull provide a reasonable alternative. For instance, see Faraggi, Reiser and Schisterman (2003).

The estimation of $P(X > Y)$ (where (X, Y) is a bi-model) is also of great importance in engineering reliability, typically, in stress-strength model. For example, if X is a strength of the system which is subject to a stress Y, then $P(X > Y)$ is a measure of the system performance. The system fails if and only if at any time the applied stress is greater than its strength. During the past twenty years, many

results have been obtained on the estimation of probability under the different parametric models. The most recent results were obtained by Kundu and Gupta (2005 and 2006) who studied the estimation of the probability under the bi-Weibull (WE) and bi-Generalized Exponential models (GE). Our research partially extends their results.

Pepe (2003) summarized ROC analysis by three different approaches: a complete non-parametric approach which uses the empirical method to estimate survival functions and therefore ROC and AUC; directly modeling the distribution of biomarkers; the semi-parametric approach which models ROC curve rather than the distribution of biomarkers. For example, the bi-normal model belongs to this approach.

Gonen and Heller (2007) presented an alternative to the bi-normal model based on the Lehmann family, also known as the proportional hazards specification. The resulting ROC curve and its index AUC have simple analytic forms. Closed-form expressions for the functional estimates and their corresponding asymptotic variances are derived.

Yang and Prentice (2005) proposed a semi-parametric model (crossing model) to handle the case when two survival curves cross. This model includes the Cox proportional hazard model and the proportional odds model as submodels.

In this dissertation, we explore a parametric approach to estimate AUC under the bi-Exponentiated Weibull (EW) models. We give a general formula of AUC under bi-EW. The maximum likelihood estimation (MLE) method is used to estimate the parameters and hence the AUC. Their asymptotic distributions are derived. Simulation studies are performed to compare the accuracy of estimated parameters and AUC under different sample sizes and parameter settings in terms of absolute relative biases (ARB) and the square root of mean square error (RMSE).

Our second part of this dissertation is to study the ROC analysis plus its index

AUC under the crossing model. We extend the results obtained by Gonen and Heller (2007).

1.2 ROC AND AUC

Let T be a continuous biomarker, assume for larger T, a subject is more likely to be diseased. A subject is classified as positive or disease status, if $T > c$ and as negative or non-disease status otherwise, where c is a cut-point. Let D be the disease status: $D = 1$ represents diseased population while $D = 0$ represents non-diseased population. Sensitivity of T is defined as the probability of correctly classifying as disease status and specificity as the probability of correctly classifying as non-disease status. That is

$$Sensitivity(c) = P(T > c|D = 1)$$

and

$$Specificity(c) = P(T < c|D = 0).$$

Definition 1.2.1. *The ROC curve is* $\{1 - Specificity(c), Sensitivity(c)\}, c$ *ranges over all possible values of* T.

Let $X = \{T|D = 1\}$ and $Y = \{T|D = 0\}$. The distribution functions (survival functions) of X and Y are F_1 (S_1) and F_0 (S_0) and the density functions are f_1 and f_0, respectively. Then

$$Sensitivity(c) = 1 - F_1(c)$$

and

$$Specificity(c) = F_0(c).$$

Using a change of variable: $t = 1-Specificity(c)$, then $t = 1-F_0(c)$, so $c = F_0^{-1}(1-t)$. Therefore

$$ROC(t) = 1 - F_1(F_0^{-1}(1 - t)) = S_1(S_0^{-1}(t)),$$

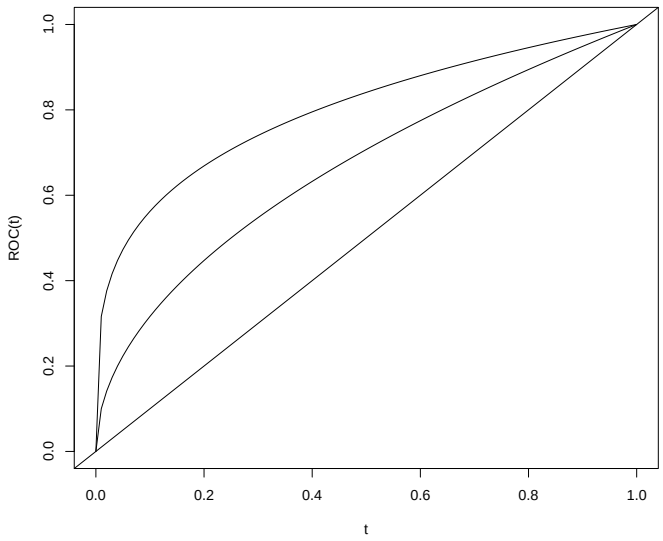

Figure 1.2.1: ROC curve for a good and an uninformative test

where $0 < t < 1$.

Notice that ROC curve is a monotonic increasing curve in a unit square starting from $(0, 0)$ and ending at $(1, 1)$. A good biomarker has a very concave down ROC curve as shown in Figure 1.2.1.

Now we state several main properties of ROC curves quoted from Pepe (2003), for completeness purpose. Detail proofs can be seen in Pepe's book.

Property 1.2.1. *(Invariance Property, Pepe (2003), result 4.1) Let T be a continuous random variable and h be any monotonic increasing function. Then T and $h(T)$ have the same ROC curve.*

Property 1.2.2. *(Pepe (2003), result 4.3) Suppose F_0 is strictly increasing distribution function. Then the ROC curve has slope*

$$\frac{dROC(t)}{dt} = \frac{f_1(F_0^{-1}(1-t))}{f_0(F_0^{-1}(1-t))}.$$

Property 1.2.3. *(Pepe, (2003), result 4.4) If the densities f_1 and f_0 have a monotonic likelihood ratio, then the ROC curve is concave down.*

When we have two different biomarkers, the one with uniformly higher ROC curve is better. However, when two ROC curves cross, we can not simply compare and find which one is better. Many different numerical indices have been proposed to summarize and compare ROC curves of different biomarkers (Simpson and Fitter, 1973; Lee and Hsiao, 1996; Gail and Green, 1976; Cammpbell, 1994). Here we list three popular summary indices.

- Kolmogorov-Smirnov measure of distance (Gail and Green, 1976; Cammpbell, 1994):

$$KS = \max_{0<t<1} |ROC(t) - t| = \max_{0<t<1} |S_1(S_0^{-1}(t)) - t| = \sup_c |S_1(c) - S_0(c)|, \quad (1.2.1)$$

  where $c = S_0^{-1}(t)$. A better biomarker has larger KS, shown in Figure 1.2.2 (a).

- The point $(t, ROC(t))$ is called a symmetry point if $ROC(t) = 1 - t$. Therefore, a better biomarker has a symmetry point closer to $(0, 1)$, shown in Figure 1.2.2 (b).

- The area under the ROC curve (AUC): $AUC = \int_0^1 ROC(t)dt$, shown in Figure 1.2.3.

In this dissertation, we adopt AUC as a main summary index of ROC. The partial area under the curve, $pAUC(t_0)$, is a summary index that restricts $t(1 - specificity)$ at and below t_0. It is defined as

$$pAUC(t_0) = \int_0^{t_0} ROC(t)dt.$$

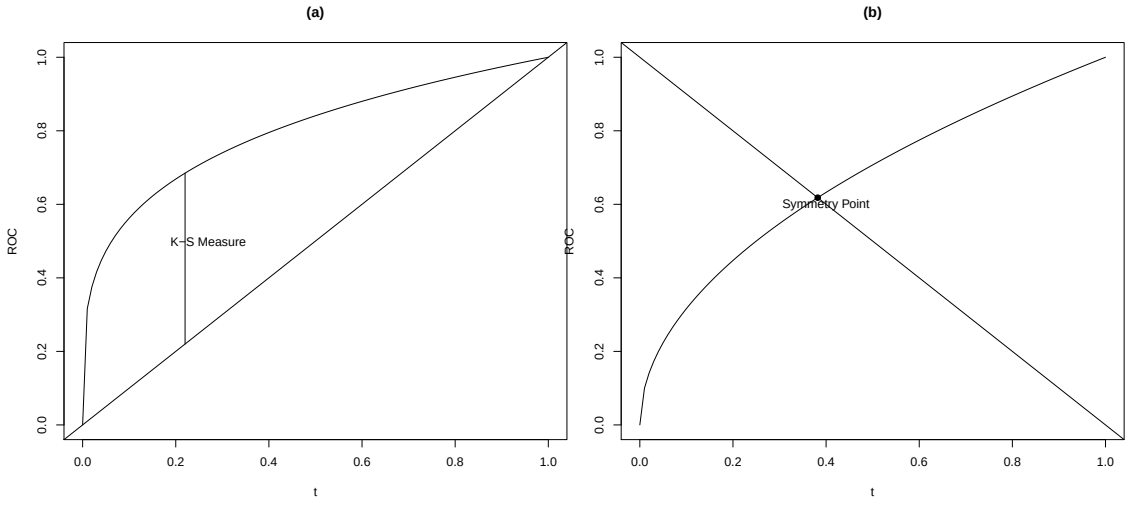

Figure 1.2.2: (a): K-S measure of distance and (b): Symmetry point

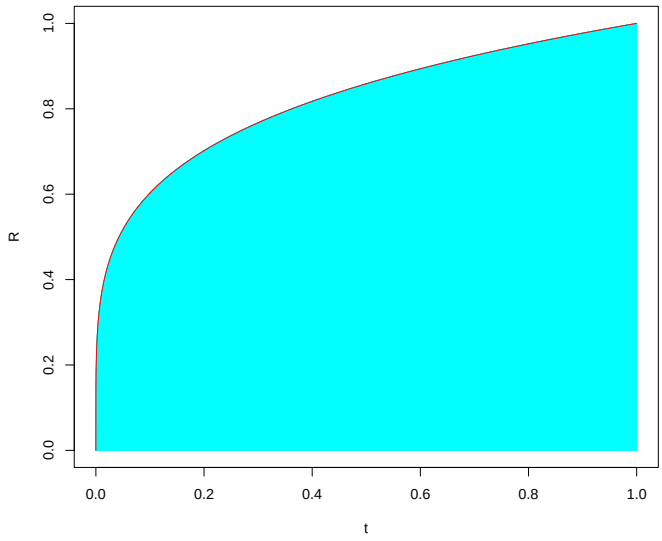

Figure 1.2.3: Area Under the Curve (AUC)

Its values range from $\frac{t_0^2}{2}$ for a completely uninformative test $ROC(t) = t$ to t_0 for a perfect test.

We say (X, Y) is a bi-model if X and Y are independent and are from the same distribution family but with different parameters. Under the bi-model (X, Y), we have:

Proposition 1.2.1. *$AUC = P(X > Y)$.*

Proof. Recall that $X = \{T|D = 1\}$ and $Y = \{T|D = 0\}$. From the definition of AUC, we have:

$$\begin{aligned} AUC &= \int_0^1 ROC(t)dt \\ &= \int_0^1 S_1(S_0^{-1}(t))dt \\ &= \int_{-\infty}^{\infty} S_1(x)dS_0(x) \\ &= \int_{-\infty}^{\infty} P(X > x)f_0(x)dx, \end{aligned}$$

where $x = S_0^{-1}(t)$. Since X and Y are independent, we have:

$$\begin{aligned} AUC &= \int_{-\infty}^{\infty} E_X I(X > x)f_0(x)dx \\ &= E_Y[E_X I(X > Y)] \\ &= P(X > Y). \end{aligned}$$

■

Example: Let (X, Y) be a bi-Weibull model with common shape parameter β and different scale parameters λ_1 and λ_2 respectively. That is: $X \sim \text{WE}(\beta, \lambda_1)$ and $Y \sim \text{WE}(\beta, \lambda_2)$, then

$$AUC = \frac{\lambda_1}{\lambda_1 + \lambda_2}. \tag{1.2.2}$$

Similarly, if (X, Y) is a bi-Generalized Exponential model with common scale parameter λ but different shape parameters α_1 and α_2. That is: $X \sim \text{GE}(\alpha_1, \lambda)$ and $Y \sim \text{GE}(\alpha_2, \lambda)$, then

$$AUC = \frac{\alpha_2}{\alpha_1 + \alpha_2}. \tag{1.2.3}$$

For details, see Gupta and Kundu (2005, 2006).

1.3 EXISTING MODELING METHODS OF ROC CURVE

Throughout this dissertation, we always assume X and Y are from the diseased and non-diseased population with distribution function F_1 and F_0 respectively, and $X_1, X_2, ...X_m$ is a random sample from F_1, $Y_1, Y_2, ...Y_n$ is a random sample from F_0. In ROC analysis, Pepe (2003) summarized three approaches to estimate ROC curve and its summary indices:

The first one is to apply non-parametric empirical methods to obtain the empirical ROC curve.

Recall that

$$ROC(t) = S_1(S_0^{-1}(t))$$

and S_1 and S_0 can be estimated by empirical survival functions $\widehat{S_{1m}}$ and $\widehat{S_{0n}}$. That is $\widehat{S_{1m}}(x) = \frac{1}{m}\sum_{i=1}^{m} I[X_i > x]$ and $\widehat{S_{0n}}(y) = \frac{1}{n}\sum_{j=1}^{n} I[Y_j > y]$. $S_0^{-1}(t)$ is estimated by $\widehat{S_{0n}}^{-1}(t) = inf\{y, \widehat{S_{0n}}(y) > t\}$. So we obtain an empirical estimator $\widehat{ROC}(t) = \widehat{S_{1m}}(\widehat{S_{0n}}^{-1}(t))$. The following theorem states the strong consistency of the empirical estimator.

Theorem 1.3.1. *(Hsieh and Turnbull (1996)) Suppose the samples from $F_1(t)$ and $F_0(t)$ have size m and n respectively, and $\frac{m}{n} \to p > 0$ as $m, n \to \infty$, and the slope of the curve $S_1(S_0^{-1}(t))$, that is $\frac{f_1(F_0^{-1}(1-t))}{f_0(F_0^{-1}(1-t))}$ is bounded on any subinterval (a, b) of*

$(0,1)$, *then we have*

$$\sup_{0<t<1} |\widehat{ROC}(t) - ROC(t)| \to 0 \quad a.s.as \quad n \to \infty.$$

The second approach is to model the test result distributions. Since $ROC(t) = S_1(S_0^{-1}(t))$, suppose we model each distribution as a parametric distribution with parameters α and β for the diseased and non-diseased populations: $S_1(t) = S_1(t, \alpha)$ and $S_0(t) = S_0(t, \beta)$. We estimate α from the diseased test result data and β from the non-diseased test result data. Then the resultant ROC estimate is

$$\widehat{ROC}(t, \widehat{\alpha}, \widehat{\beta}) = S_1(S_0^{-1}(t, \widehat{\beta}), \widehat{\alpha}).$$

If the parameters are estimated by MLE methods, then the ROC estimator is unbiased and fully efficient.

Another way in this approach is semi-parametric location-scale models. Consider the semi-parametric location-scale model for the test results:

$$X = \mu_1 + \sigma_1 \epsilon,$$

$$Y = \mu_0 + \sigma_0 \epsilon,$$

where ϵ are independent random variable with mean 0 and variance 1 and survival function G, and G is unspecified. Under this location-scale model, ROC curve can be written as

$$ROC(t) = G\left(\frac{(\mu_0 - \mu_1)}{\sigma_1} + (\frac{\sigma_0}{\sigma_1})G^{-1}(t)\right).$$

The estimates of the location-scale parameters can be calculated using sample means and sample standard deviations. The empirical survival function of ϵ is defined as

$$\widehat{G}(y) = \frac{1}{m+n}\left(\sum_i I\left[\frac{X_i - \widehat{\mu_1}}{\widehat{\sigma_1}} \geq y\right] + \sum_j I\left[\frac{Y_j - \widehat{\mu_0}}{\widehat{\sigma_0}} \geq y\right]\right).$$

Therefore we get an estimator of $ROC(t)$:

$$\widehat{ROC}(t) = \widehat{G}\left(\frac{(\widehat{\mu_0} - \widehat{\mu_1})}{\widehat{\sigma_1}} + (\frac{\widehat{\sigma_0}}{\widehat{\sigma_1}})\widehat{G}^{-1}(t)\right).$$

The third approach is called parametric distribution-free methods: we assume the form of the ROC curve but we do not make assumptions for the distribution of test results. Metz, et. al (1998) dealed with continuous data by categorizing them into a finite number of pre-defined categories, therefore ordinal data.

Among the above three approaches, the bi-normal model has been the most popular model which attracts most attention. The bi-normal ROC curve plays a central role in ROC analysis just as the normal distribution is a classic model for distribution function. We have seen that the ROC curve is invariant under monotonic increasing transformation. If $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_0, \sigma_0^2)$, then $ROC(t) = \Phi(a+b\Phi^{-1}(t))$, where $a = \frac{\mu_1 - \mu_0}{\sigma_1}, b = \frac{\sigma_0}{\sigma_1}$. and Φ denotes the standard normal distribution function. The proof is shown as follows:

For any threshold c,

$$1 - Specificity(c) = P[Y > c] = \Phi\left(\frac{\mu_0 - c}{\sigma_0}\right)$$

and

$$Sensitivity(c) = P[X > c] = \Phi\left(\frac{\mu_1 - c}{\sigma_1}\right).$$

Let $t = 1 - Specificity(c)$, then $c = \mu_0 - \sigma_0\Phi^{-1}(t)$. Hence

$$\begin{aligned} ROC(t) &= Sensitivity(c) \\ &= \Phi\left(\frac{\mu_1 - c}{\sigma_1}\right) \\ &= \Phi\left(\frac{\mu_1 - \mu_0 + \sigma_0\Phi^{-1}(t)}{\sigma_1}\right) \\ &= \Phi\left(a + b\Phi^{-1}(t)\right). \end{aligned}$$

Conversely, if (X, Y) has the above ROC function, we can prove that it must be a bi-normal model.

Proposition 1.3.1. *If a bi-model* (X, Y) *has ROC function* $ROC(t) = \Phi(a+b\Phi^{-1}(t))$ *and* S_0 *is a strictly decreasing function, then there exists a monotonic increasing function* h*, such that* $h(X) \sim N\left(\frac{a}{b}, \frac{1}{b^2}\right)$, $h(Y) \sim N(0,1)$.

Proof: We define $h(y) = -\Phi^{-1}(S_0(y))$. If $ROC(t) = \Phi(a + b\Phi^{-1}(t))$ and S_0 is a strictly decreasing function, notice that $1 - \Phi(-y) = \Phi(y)$.

We have

$$\begin{aligned} P(h(Y) < y) &= P\left(\Phi^{-1}(S_0(Y)) > -y\right) \\ &= P\left(1 - S_0(Y) < 1 - \Phi(-y)\right) \\ &= P(F_0(Y) < \Phi(y)) = \Phi(y) \end{aligned}$$

since $F_0(Y)$ has a uniform distribution. On the other hand, we have

$$\begin{aligned} P(h(X) > y) &= P\left(\Phi^{-1}(S_0(X)) < -y\right) \\ &= P(S_0(X) < \Phi(-y)) \\ &= P(X > S_0^{-1}(\Phi(-y))) \\ &= S_1\left(S_0^{-1}(\Phi(-y))\right) \\ &= ROC(\Phi(-y)) \\ &= \Phi(a - by) \\ &= 1 - \Phi\left(\frac{y - \frac{a}{b}}{\frac{1}{b}}\right). \end{aligned}$$

This completes the proof.

Corollary 1.3.2. *Suppose that the test results* X *and* Y *are from diseased and non-diseased populations respectively. The ROC function of* (X, Y) *is* $ROC(t) = \Phi(a + b\Phi^{-1}(t))$ *if and only if* (X, Y) *is a bi-normal model.*

Most ROC analyses are based on this bi-normal assumption. Under this assumption, we require the estimation of the unspecific monotonic increasing function in addition to the two parameters, which will bring many nuisance parameters, and therefore a lot of computations.

Another developmental direction in ROC analysis is to consider the use of multiple diagnostic tests (biomarkers) to maximize the diagnostic accuracy (such as AUC). Su, J. Q. and Liu, J. S. (2003) used the optimal linear combination of multiple diagnostic markers to maximize AUC. Debashis Ghosh (2004) used the copula function which is used to set up relationship between distribution function of multiple biomarkers and those of two single biomarkers and derived a multivariate extension of the bi-normal model. There are not many research results in the literature about the case of multiple biomarkers.

1.4 EXISTING ESTIMATING METHODS OF AUC

AUC Estimation Via Estimated ROC. From the definition of ROC, we have

$$ROC(t) = 1 - F_1(F_0^{-1}(1-t))$$

for the general setting, where $0 \leq t \leq 1$. Especially under the bi-normal model, we have

$$ROC(t) = \Phi(a + b\Phi^{-1}(t)),$$

where Φ is a standard normal distribution function. The main problem is to estimate the two parameters a and b. Fushing and Turnbull (1996) carefully analyzed the main properties of ROC curves, especially the invariance under monotone increasing transformations of test results, which is also a general requirement for various estimated ROC curves. Metz, C. R., Herman, B. A. and Shen, J. H.(1998) developed two methods to fit ROC curve called LABROC4 and LABROC5. The first one (LABROC4)

is a true maximum likelihood (ML) method, but it needs a lot of computation. The second substantially reduces the computation process, but it is not a rigorous ML method. The basic idea is to transform continuous raw data into ranked data. It also needs good initial value estimates. Pepe and Cai (2004) proposed the concept of placement value and therefore ROC can be regarded as its distribution function. The likelihood of the placement value provides a new approach to the parameters estimation. Cai and Moskowitz (2004) used the profile likelihood method to estimate parameters which can simplify the computation process. Zhou and Lin (2008) used the semi-parametric maximum likelihood to estimate the ROC curve and also reduced the number of estimated nuisance parameters. Its basic idea is similar to Metz (1998). Gu, Ghosal and Roy (2005) proposed a non-parametric method based on the Bayesian bootstrap to estimate ROC. Furthermore, they compared the proposed method with other three existing semi-parametric and non-parametric methods. A simulation study shows that the proposed method compares favorably in terms of accuracy and robustness. Under the bi-normal model, for the estimated parameter $\widehat{a}$ and $\widehat{b}$, we obtain the estimated AUC:

$$\widehat{AUC} = \int_0^1 \Phi(\widehat{a} + \widehat{b}\Phi^{-1}(t))dt.$$

AUC Estimation Via Non-Parametric Method. The non-parametric estimation of AUC, which skips the estimation of ROC, can be directly obtained. There are three non-parametric methods in the rigorous sense and a recently developed empirical likelihood method.

The first one is the Wilcoxon statistic. The non-parametric estimate and corresponding standard error (SE) are calculated directly from the raw data. Let $\widehat{\theta}$ be the estimated AUC from the Wilcoxon statistic, namely $\widehat{\theta} = \dfrac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} I(X_i \geq Y_j)$. The SE of $\widehat{\theta}$ is calculated from Hanley and Mcneil (1982),

$$SE(\widehat{\theta}) = \sqrt{\frac{\widehat{\theta}(1-\widehat{\theta}) + (m-1)(Q_1 - \widehat{\theta}^2) + (n-1)(Q_2 - \widehat{\theta}^2)}{mn}},$$

where $Q_1 = \dfrac{\widehat{\theta}}{2-\widehat{\theta}}$ and $Q_2 = \dfrac{2\widehat{\theta}^2}{1+\widehat{\theta}}$.

The second method is by Delong, et. al (1988). That is

$$\widehat{AUC} = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} r(X_i, Y_j),$$

where $r(X,Y) = I\{X < Y\} + \frac{1}{2}I\{X = Y\}$. This is actually the Mann-Whitney U-statistic. In other words, let $V_1(X_i) = \dfrac{1}{n}\sum_{j=1}^{n} I(X_i \geq Y_j)$ and $V_0(Y_j) = \dfrac{1}{m}\sum_{i=1}^{m} I(X_i \geq Y_j)$. It is easy to get that $\overline{V_1} = \overline{V_0} = \widehat{AUC}$. Thus the averages of $V_1(X_i)$ and $V_0(Y_j)$ are both equivalent to the U statistic, and the variance of $\widehat{AUC}$ follows that $Var(\widehat{AUC}) = \dfrac{Var(V_1)}{m} + \dfrac{Var(V_0)}{n}$.

The third one is the kernel method in Zou, et. al (1998), which chooses the Gaussian kernel and estimates the density function by

$$\widehat{f_m(t)} = \frac{1}{m}\sum_{i=1}^{m}\frac{1}{h_x}\phi\left(\frac{t-x_i}{h_x}\right),$$

where ϕ is the density of standard normal distribution. Zou, et. al (1998) gave a bandwidth

$$h_x = 0.9min(s_x, iqr_x/1.34)m^{-1.5},$$

where s_x and iqr_x are the standard deviation and inter quartile range of the m test results on the diseased sample. Lloyd (1998) showed that the resulting kernel estimate of the AUC can be written as

$$\widehat{AUC} = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\Phi\left(\frac{X_i - Y_j}{\sqrt{h_x^2 + h_y^2}}\right), \text{where } h_y = 0.9min(s_y, iqr_y/1.34)n^{-1.5}.$$

Lloyd and Yong (1999) used a more complex selection of the bandwidth procedure in smoothing the ROC curve. Their method is an extension of the two-stage plug-in

procedure of Wand and Jones (1995). One similar to the above but the different method developed by Qiu and Le (2001) is to partially use the kernel estimation. Since $ROC(t) = 1 - F_1(F_0^{-1}(1-t))$ for $0 < t < 1$, the ROC function is determined by the distribution function F_1 and the quartile function F_0^{-1}. Here they estimate F_1 by the kernel method. Leger and Altman (1995) suggested a formula to compute the asymptotically optimal bandwidth. About the estimation of the quartile F_0^{-1}, they use the estimator proposed by Harrell-Davis (1982) and Dielman (1994).

A Parametric Method Via $P(X > Y)$. Since AUC can be regarded as a probability of the biomarker of disease greater than that of non-disease, the remaining question is to estimate the probability (AUC). In this way, we do not need to estimate the ROC. Instead, we can directly obtain the estimation of AUC. The estimation of the probability is also of importance in engineering reliability, therefore attracts much attention.

1.5 APPLICATION OF $P(X > Y)$ IN ENGINEERING RELIABILITY

In engineering reliability, if X is the strength of a system which is subjected to a stress Y , then $P(X > Y)$ is a measure of the system performance. The system fails if and only if at any time the applied stress is greater than its strength. Notice that $P(X > Y)$ is AUC in medical application and reliability in engineering. Throughout rest of this dissertation, we simply use AUC to represent the probability $P(X > Y)$. Kotz, Lumelskii and Pensky (2002) gave a comprehensive analysis and summarization about stress-strength model. Many research results have been obtained for the estimation of AUC under popular survival models. Awad (1981) estimated AUC when (X, Y) has the bi-Exponential distribution. Woodward and Kelley (1977) estimated

AUC when (X, Y) has the bi-normal distribution. Constantine and Kerson (1986) considered AUC when (X, Y) has the bi-Gamma distribution. Ahmad, Fakhry and Jaheen (1997) and Surles and Padgett (1998, 2001) considered the estimation of AUC when (X, Y) has the bi-Burr type X distribution. Raqab and Kundu (2005) considered the AUC when X and Y have the Generalized Exponential distribution and Burr type X distribution respectively.

The newest results were obtained by Kundu and Gupta (2005, 2006). They studied the estimation under bi-GE and bi-WE: for the bi-GE model, they considered the case when X and Y have a common scale parameter but different shape parameters and derived the MLE of AUC and its asymptotic distribution. Furthermore, for the common scale parameter, they derived a uniformly minimum variance unbiased MLE. For unknown common scale parameter, they gave a Bayes estimator. They proposed other different estimations of confidence intervals by bootstrap methods for small sample sizes. Simulation studies were performed to compare the different methods. For the bi-WE model, they considered the case when X and Y have common shape parameter but different scale parameters, derived a MLE, an approximate MLE of AUC and a Bayes estimator. They constructed two bootstrap confidence intervals and corresponding credible intervals of the Bayes estimator. The simulation studies were performed to compare the different estimators.

However, all popular survival models such as GE, WE and Gamma do not allow non-monotonic failure rates which often occur in practice. The most common non-monotonic failure-rate phenomenon in engineering and biological science involves bathtub shapes.

To overcome the above drawback, many models such as generalized gamma (Stacy, 1962), generalized F distribution (Prentice, 1975), a two parameter family (Slymen and Lachenbruch, 1984), a four-parameter family (Graver and Acar, 1979) and a three-parameter (IDB) family (Hjorth, 1980) have been proposed for modeling non-monotone failure-rate data. A recent tendency is to extend the Weibull distribution by adding more parameters: Carrasco, Ortega and Corderio (2008) proposed a four-parameter modified Weibull distribution and used this new model to fit Serum-reversal data and Radiotherapy data. Wahed, Luong and Jeong (2009) proposed a new generalization of the Weibull distribution called Beta-Weibull distribution and used it to fit breast cancer data and compared it with other extended Weibull distributions.

Our interest is on Exponentiated Weibull model (EW) proposed by Mudholkar and Srivastava (1995). This is an extension of both GE and WE. The EW model not only includes distributions with unimodel and bathtub failure rates but also allows for a broader class of monotonic hazard rates. Another advantage is that we find the EW model may fit better than GE for some actual data. For instance, Khedhairi, Sarhan and Tadj (2007) analyzed real data (a set of cracking data). They used three different models to fit it: Exponential distribution, Generalized Exponential distribution and Generalized Rayleigh (a special case of EW) distribution. Both MLE value and Kolmogorov-Smirnov test show that the generalized Rayleigh model fits better than the other two models.

This dissertation contains two parts. The first part (chapters 2-3) studies the estimation of AUC under Bi-EW. Our results generalize the ones in Kundu and Gupta (2005, 2006). We derive the MLE of AUC and its asymptotic distribution. Then

we compare it with the results under bi-GE and bi-WE. Simulation studies are performed to compare MLE performance of different parameters under different moderate sample sizes and parameter settings and investigate the MLE performance of the corresponding AUC. The second part studies the estimation of AUC under the crossing model (chapter 4). The remaining chapters are organized as follows: in Chapter 2, we introduce the Exponentiated-Weibull (EW) distribution and analyze its property; in Chapter 3, we study the estimation of AUC under bi-EW models, derive the asymptotic normality of estimated AUC and do simulations to compare and investigate the performance of every parameter and AUC; in Chapter 4, we propose a new semi-parametric model which extends the Lehmann Family and study AUC under the new semi-parametric model; in Chapter 5, we discuss several questions for future research.

CHAPTER 2
EXPONENTIATED WEIBULL DISTRIBUTION AND ITS PROPERTIES

In this chapter, we first review two popular survival models: Weibull (WE) and Generalized Exponential (GE) distribution. Then we introduce an extension of them: Exponentiated-Weibull (EW) distribution. As a current popular survival model to fit bathtub shape data, EW was proposed by Mudholkar and Srivastava (1995). We summarize its statistical properties and give general moment formulas.

2.1 EXPONENTIATED WEIBULL DISTRIBUTION

In this section, we first review two popular survival models: Weibull distribution and Generalized Exponential distribution, then introduce an extension of both: Exponentiated-Weibull distribution.

Definition 2.1.1. *A random variable X is said to have Weibull distribution if X has the following distribution function:*

$$F(x) = 1 - e^{-\lambda x^{\beta}}, \qquad x > 0, \beta > 0, \lambda > 0$$

and density function

$$f(x) = \lambda\beta x^{\beta-1} e^{-\lambda x^{\beta}}, \qquad x > 0, \beta > 0, \lambda > 0.$$

Then, its hazard function is:

$$h(x) = \frac{f(x)}{S(x)} = \frac{\lambda\beta x^{\beta-1} e^{-\lambda x^{\beta}}}{e^{-\lambda x^{\beta}}} = \lambda\beta x^{\beta-1}, \qquad x > 0, \beta > 0, \lambda > 0.$$

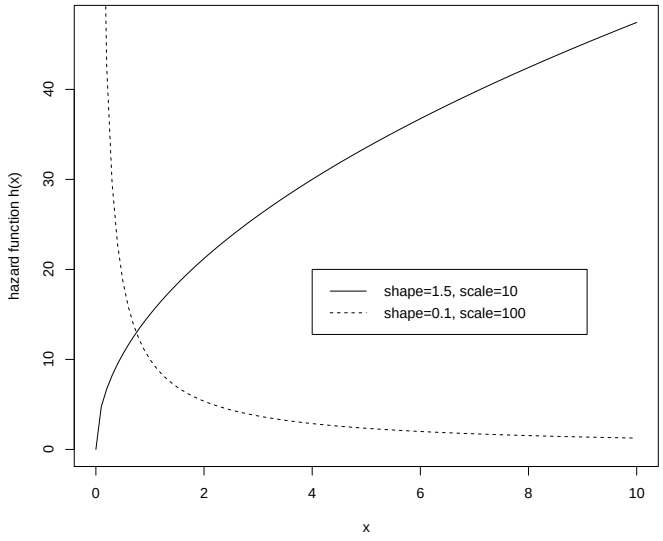

Figure 2.1.1: Hazard function shapes of Weibull distribution

We denote $X \sim WE(\beta, \lambda)$. Obviously, the monotonicity of the hazard function only depends on the shape parameter β. It is monotonic increasing when $\beta > 1$ and decreasing when $\beta < 1$,

Definition 2.1.2. *A random variable X is said to have Generalized Exponential (GE) distribution if X has the following distribution function:*

$$F(x) = (1 - e^{-\lambda x})^{\alpha} \qquad x > 0, \alpha > 0, \lambda > 0$$

and density function

$$f(x) = \alpha\lambda e^{-\lambda x}(1 - e^{-\lambda x})^{\alpha-1} \qquad x > 0, \alpha > 0, \lambda > 0.$$

Then, its hazard function is

$$h(x) = \frac{f(x)}{S(x)} = \frac{\alpha\lambda e^{-\lambda x}(1 - e^{-\lambda x})^{\alpha-1}}{1 - (1 - e^{-\lambda x})^{\alpha}} \qquad x > 0, \alpha > 0, \lambda > 0.$$

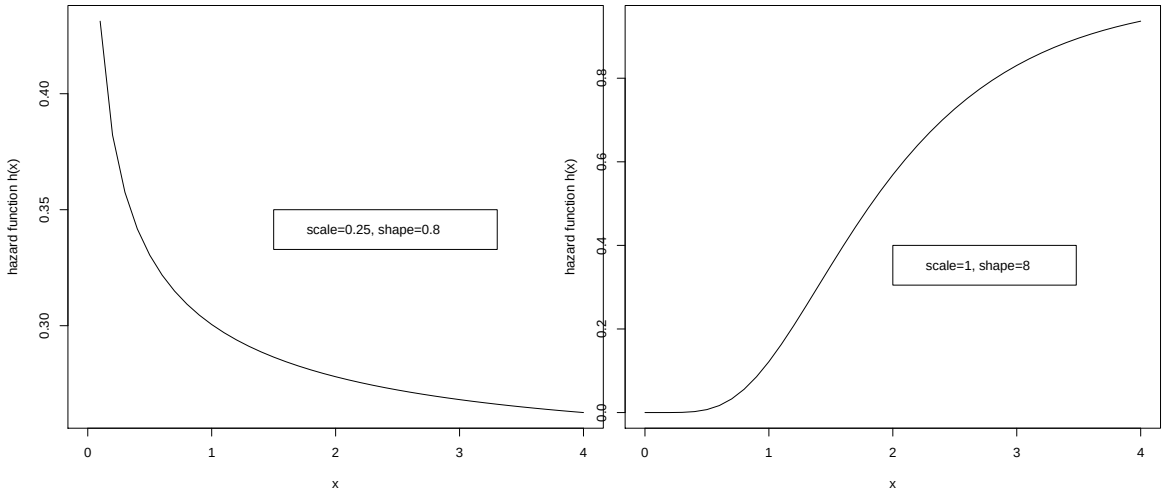

Figure 2.1.2: (a): Decreasing hazard function and (b): Increasing hazard function

We denote $X \sim GE(\alpha, \lambda)$.

For the monotonicity of hazard function of GE, we have the following results.

Theorem 2.1.3. *(Barlow and Proschan, 1975). For a GE distribution, its hazard function is increasing if the shape parameter $\alpha > 1$ and decreasing if $\alpha < 1$, it is a constant for $\alpha = 1$.*

The proof follows by observing the second partial derivative of the logarithm of the density function

$$\frac{\partial^2}{\partial x^2} \ln f(x, \alpha, \lambda) = -(\alpha - 1)\frac{e^{-\lambda x}}{(1 - e^{-\lambda x})^2}.$$

Use the log convexity of the density function. The shapes of hazard function of GE are shown in Figure 2.1.2.

Definition 2.1.4. *A random variable X is said to have Exponentiated-Weibull (EW) distribution if X has the following distribution function:*

$$F(x) = (1 - e^{-\lambda x^{\beta}})^{\alpha}, \qquad x > 0, \alpha > 0, \beta > 0, \lambda > 0$$

and density function

$$f(x) = \alpha\beta\lambda x^{\beta-1}e^{-\lambda x^{\beta}}(1-e^{-\lambda x^{\beta}})^{\alpha-1}, \qquad x>0, \alpha>0, \beta>0, \lambda>0.$$

Then, its hazard function is

$$h(x) = \frac{f(x)}{S(x)} = \frac{\alpha\beta\lambda x^{\beta-1}e^{-\lambda x^{\beta}}(1-e^{-\lambda x^{\beta}})^{\alpha-1}}{1-(1-e^{-\lambda x^{\beta}})^{\alpha}}, \qquad x>0,\ \alpha>0, \beta>0, \lambda>0.$$

We say that $X \sim EW(\alpha, \beta, \lambda)$. Note that EW is a generalization of WE and GE. That is, when $\alpha = 1$, EW reduces to be WE, and when $\beta = 1$, it reduces to be GE.

Wahed, Luong and Jeong (2009) presented a general way to extend existing distributions. It has been demonstrated by Wahed, Ferreira and Steel (2006) that any parametric family of distribution can be incorporated into larger families through an application of the probability intergal transform. In this way, we can derive the EW distribution: given two valid pdf f_1 and f_2, with the latter having support on the unit interval, the new pdf f may be obtained by applying the equation

$$f(x) = f_2(F_1(x))f_1(x),$$

where F_1 is cdf corresponding to f_1. Here we choose f_1 as a two-parameter Weibull distribution and f_2 as a two parameter Beta distribution with pdf

$$f_2(\mu, \alpha_1, \alpha_2) = C(\mu)^{\alpha_1-1}(1-\mu)^{\alpha_2-1}, 0<\mu<1, 0<\alpha_1, \alpha_2<\infty.$$

Therefore we obtain a Beta-Weibull distribution with pdf:

$$f(x, \alpha_1, \alpha_2, \beta, \gamma) = \frac{1}{B(\alpha_1, \alpha_2)}\frac{\gamma}{\beta}(\frac{x}{\beta})^{\gamma-1}e^{-\alpha_2(\frac{x}{\beta})^{\gamma}}(1-e^{-(\frac{x}{\gamma})^{\gamma}})^{\alpha_1-1},$$

where $0<x, \alpha_1, \alpha_2, \beta, \gamma<\infty$ and $B(\alpha_1, \alpha_2)$ is a Beta function. The drawback is the survival function of Beta-Weibull distribution can not be expressed in a closed-form. We only use the incomplete Beta function to express the hazard function as

$$h(x; \alpha_1, \alpha_2, \beta, \gamma) = \frac{\gamma}{\beta}\frac{(1-e^{-(\frac{x}{\beta})^{\gamma}})^{\alpha_1-1}e^{-\alpha_2(\frac{x}{\beta})^{\gamma}}(\frac{x}{\beta})^{\gamma-1}}{1-I(1-e^{-(\frac{x}{\beta})^{\gamma}}; \alpha_1, \alpha_2)},$$

where

$$I(v, \alpha_1, \alpha_2) = \frac{\int_0^v u^{\alpha_1 - 1}(1-u)^{\alpha_2 - 1} du}{\int_0^1 u^{\alpha_1 - 1}(1-u)^{\alpha_2 - 1} du}$$

is the incomplete Beta function. When $\alpha_2 = 1$, the Beta-Weibull distribution reduces to be an Exponentiated -Weibull distribution.

The following theorem states the property of the hazard function of EW. It is also the main motivation that we study the model.

Theorem 2.1.5. *(Mudholkar, Srivastava and Freimer, 1995). For EW$(x, \alpha, \beta, \lambda)$ distribution, its hazard function is (a) monotonic increasing for $\beta \geq 1$ and $\alpha\beta \geq 1$; (b) monotonic decreasing for $\beta \leq 1$ and $\alpha\beta \leq 1$; (c) bathtub-shaped for $\beta > 1$ and $\alpha\beta < 1$; (d) uni-model for $\beta < 1$ and $\alpha\beta > 1$, The monotonicities are strict except for $\alpha = \beta = 1$.*

The proof of above theorem involves the technical details of derivatives. Figure 2.1.3 shows that the hazard function of EW allows the non-monotonic case and presents the bath-tub shape under some parameter setting. Qian (2012) shows the properties through a graphical approach.

2.2 GENERAL MOMENT FORMULAS

Moment formulas for relatively complicated survival models are a topic to attract much attention. Many papers gave discussion and obtained some results (see Choudhury (2005), Mudholkar, Srivasttava and Freimer (1995), Nadarajah (2005)). Some papers only derived the moments with strong restricted condition. Mudhdkar, Srivastava and Freimer (1995) derived an expression for moments of Exponentiated Weibull by the integration of a quantile function for an integer valued α. The k-th moment

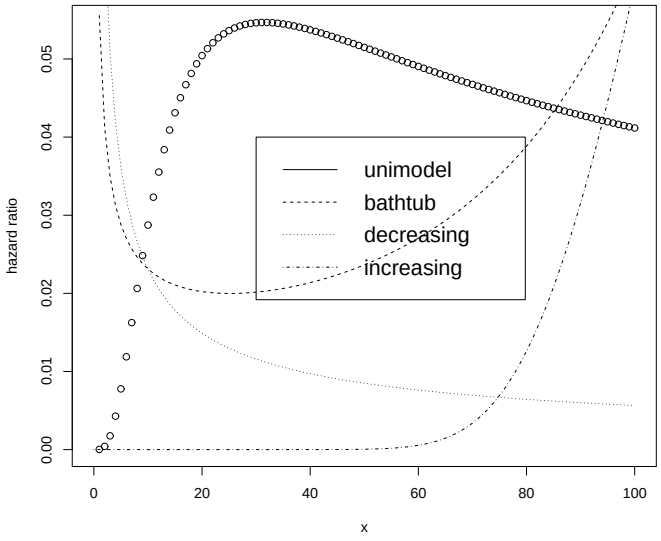

Figure 2.1.3: Hazard function shapes of EW distribution

formula is

$$\mu_k = \alpha\lambda^{-\frac{k}{\beta}}\Gamma\left(\frac{k}{\beta}+1\right)\sum_{i=0}^{\alpha-1}(-1)^i\binom{\alpha-1}{i}\frac{1}{(i+1)^{\frac{k}{\beta+1}}}. \tag{2.2.1}$$

Moreover, if $k\beta^{-1} = r$ is a positive integer, then

$$\mu_k = \alpha(-\lambda)^{-r}\left[\frac{d^r}{ds^r}B(s,\alpha)\right]\Bigg|_{s=1}, \tag{2.2.2}$$

where $B(s,\alpha) = \int_0^1 x^{s-1}(1-x)^{\alpha-1}dx$.

Here we derive more general moment formulas of the EW family. Our formulas include all the results above as special cases. Suppose X follows the EW distribution with distribution function $F_X(x;\alpha,\lambda,\beta) = (1-e^{-\lambda x^\beta})^\alpha$. Then we have the following formulas.

Theorem 2.2.1.

$$\begin{aligned} & E[X^{i\beta}(\ln X)^j] \\ = \; & (-1)^j \frac{\alpha}{\lambda^i \beta^j} \sum_{l=0}^{j} \binom{j}{l} \Gamma^{(l)}(i+1) \sum_{k=0}^{\infty} (-1)^k \binom{\alpha-1}{k} \frac{[\ln \lambda(k+1)]^{j-l}}{(k+1)^{i+1}}, \end{aligned}$$

where $i = 1, 2, 3 \cdot \cdot, j = 0, 1, 2, 3, \cdots, \Gamma(i) = \int_0^\infty e^{-t} t^{i-1} dt$.

Proof:

$$\begin{aligned} E[X^{i\beta}(\ln X)^j] &= \alpha\lambda\beta \int_0^\infty x^{i\beta} (\ln x)^j (1 - e^{-\lambda x^\beta})^{\alpha-1} e^{-\lambda x^\beta} x^{\beta-1} dx \\ &= \alpha\lambda\beta \sum_{k=0}^{\infty} (-1)^k \binom{\alpha-1}{k} \int_0^\infty x^{(i+1)\beta-1} e^{-\lambda(k+1)x^\beta} (\ln x)^j dx. \end{aligned}$$

Let $t = \lambda(k+1)x^\beta$, then

$$dx = \frac{1}{\beta} \frac{t^{\frac{1}{\beta}-1}}{[\lambda(k+1)]^{\frac{1}{\beta}}} dt, \quad x^{(i+1)\beta-1} = \frac{t^{i+1-\frac{1}{\beta}}}{[\lambda(k+1)]^{i+1-\frac{1}{\beta}}}.$$

The integration part is

$$\begin{aligned} & \int_0^\infty e^{-t} \left(\frac{\ln t - \ln \lambda(k+1)}{\beta} \right)^j \frac{1}{\beta} \frac{t^i}{[\lambda(k+1)]^{i+1}} dt \\ = \; & \frac{1}{\beta^{j+1} [\lambda(k+1)]^{i+1}} \int_0^\infty e^{-t} t^i [\ln t - \ln \lambda(k+1)]^j dt \\ = \; & \frac{1}{\beta^{j+1} [\lambda(k+1)]^{i+1}} \sum_{l=0}^{j} \binom{j}{l} [-\ln \lambda(k+1)]^{j-l} \int_0^\infty e^{-t} t^i (\ln t)^l dt \\ = \; & \frac{1}{\beta^{j+1} [\lambda(k+1)]^{i+1}} \sum_{l=0}^{j} \binom{j}{l} [-\ln \lambda(k+1)]^{j-l} \Gamma^{(l)}(i+1). \end{aligned}$$

So we have

$$\begin{aligned} & E[X^{i\beta}(\ln X)^j] \\ = \; & \alpha\lambda\beta \sum_{k=0}^{\infty} (-1)^k \binom{\alpha-1}{k} \frac{1}{\beta^{j+1} [\lambda(k+1)]^{i+1}} \sum_{l=0}^{j} \binom{j}{l} [-\ln \lambda(k+1)]^{j-l} \Gamma^{(l)}(i+1) \\ = \; & (-1)^j \frac{\alpha}{\lambda^i \beta^j} \sum_{l=0}^{j} (-1)^l \binom{j}{l} \Gamma^{(l)}(i+1) \sum_{k=0}^{\infty} \binom{\alpha-1}{k} \frac{[\ln \lambda(k+1)]^{j-l}}{(k+1)^{i+1}}. \end{aligned}$$

The general moment formulas extend the following results:

- Case 1: Let $j = 0, i\beta = n$, then they simplify to be the formula in Choudlhury (2005), the moments of the EW family.

- Case 2: Let $j = 0, \beta = 1$, then they are the moments of the GE family, the formulas simplify to be the ones in Gupta and Kunda (2005).

- Case 3: Let $j = 0, \alpha = 1, i\beta = n$, then they are the moments of the WE family, the formulas simplify to be the ones in Gupta and Kunda (2006).

In the following, we will present some existing results on Exponentiated Weibull distribution.

2.3 SOME EXISTING RESULTS ON EXPONENTIATED WEIBULL DISTRIBUTION

The Exponentiated Weibull (EW) distribution as a simple and useful model which allows a bathtub shape hazard function is attracting more and more attention although several other models were also proposed to model bathtub shape data. Up to now, there is not much literature on the EW model. Here we present some existing results: Mudholkar and Houston (1996), Jiang and Murthy (1999) and Gupta et al. (2005) studied the properties of EW and showed useful applications of EW in reliability. Singh et al (2006) obtained the Bayes estimations of the distribution parameters and the reliability function and hazard function with a type II censored sample under squared error. Ashour and Afifiy (2007) considered the analysis of EW family distributed lifetime data observed under the type I progressive interval censoring with random removals. The maximum likelihood estimators of the parameters and their asymptotic variance are derived. Ashour and Afifiy (2008) derived the maximum likelihood estimators of EW with the type II progressive interval censoring with random

removals and their asymptotic variances. Kim et al (2009) derived the maximum likelihood and Bayes estimators for EW lifetime models using the symmetric and asymmetric loss function.

CHAPTER 3
THE ESTIMATION OF AUC UNDER THE BI-EXPONENTIATED-WEIBULL MODEL

In this chapter, we derive a general formula of AUC under bi-EW. It extends the ones under the bi-GE and bi-WE. Then we study the estimation of AUC. We also derive the asymptotic distribution of MLE of AUC. Simulation studies are used to check the performance of MLE of parameters and AUC under different moderate sample sizes and parameter settings.

3.1 FORMULA OF AUC UNDER BI-EW MODEL (X, Y) AND ITS MLE

For comparison, we simply review the results in Gupta and Kundu (2005, 2006). If (X, Y) is a bi-WE: $X \sim WE(\beta, \lambda_1)$ and $Y \sim WE(\beta, \lambda_2)$, then

$$AUC = \frac{\lambda_1}{\lambda_1 + \lambda_2}. \tag{3.1.1}$$

If (X, Y) is a bi-GE: $X \sim WE(\alpha_1, \lambda)$ and $Y \sim WE(\alpha_2, \lambda)$, then

$$AUC = \frac{\alpha_1}{\alpha_1 + \alpha_2}. \tag{3.1.2}$$

We consider a bi-EW model (X, Y). That is X and Y are independent and EW distributed with distribution function $F_1(x; \alpha_1, \beta, \lambda_1)$ and $F_0(y; \alpha_2, \beta, \lambda_2)$, where the two distribution functions share a common shape parameter β.

$$F_1(x; \alpha_1, \beta, \lambda_1) = (1 - e^{-\lambda_1 x^\beta})^{\alpha_1}, \quad x > 0, \alpha_1 > 0, \beta > 0, \lambda_1 > 0$$

and

$$F_0(y;\alpha_2,\beta,\lambda_2)=(1-e^{-\lambda_2 y^\beta})^{\alpha_2},\quad y>0,\alpha_2>0,\beta>0,\lambda_2>0.$$

Then we have the following general expression for AUC.

Theorem 3.1.1. *Suppose that (X,Y) follows the bi-EW model with the above distribution functions. Then*

$$AUC=\sum_{i=0}^{\infty}\sum_{j=1}^{\infty}\binom{\alpha_2}{i}\binom{\alpha_1}{j}(-1)^{i+j-1}\frac{\lambda_1 j}{\lambda_1 j+\lambda_2 i}. \tag{3.1.3}$$

Proof: By Taylor expansion, we have:

$$(1-e^{-\lambda_2 x^\beta})^{\alpha_2}=\sum_{i=0}^{\infty}\binom{\alpha_2}{i}(-e^{-\lambda_2 x^\beta})^i$$

and

$$(1-e^{-\lambda_1 x^\beta})^{\alpha_1}=\sum_{j=0}^{\infty}\binom{\alpha_1}{j}(-e^{-\lambda_1 x^\beta})^j.$$

Thus,

$$\begin{aligned}
AUC &= \int_0^\infty F_0(x;\alpha_2,\beta,\lambda_2)dF_1(x;\alpha_1,\beta,\lambda_1)\\
&= \int_0^\infty \sum_{i=0}^{\infty}\binom{\alpha_2}{i}(-e^{-\lambda_2 x^\beta})^i\sum_{j=1}^{\infty}\binom{\alpha_1}{j}j(-e^{-\lambda_1 x^\beta})^{j-1}e^{-\lambda_1 x^\beta}\lambda_1\beta x^{\beta-1}dx\\
&= \sum_{i=0}^{\infty}\sum_{j=1}^{\infty}\int_0^\infty\binom{\alpha_2}{i}\binom{\alpha_1}{j}j(-1)^{i+j-1}\lambda_1 e^{-(\lambda_2 i+\lambda_1 j)x^\beta}dx^\beta\\
&= \sum_{i=0}^{\infty}\sum_{j=1}^{\infty}\binom{\alpha_2}{i}\binom{\alpha_1}{j}(-1)^{i+j-1}\lambda_1 j\int_0^\infty e^{-(\lambda_2 i+\lambda_1 j)x^\beta}dx^\beta\\
&= \sum_{i=0}^{\infty}\sum_{j=1}^{\infty}\binom{\alpha_2}{i}\binom{\alpha_1}{j}(-1)^{i+j-1}\frac{\lambda_1 j}{\lambda_1 j+\lambda_2 i}.
\end{aligned}$$

Remark 1: Note that the exact expression of AUC is independent of the common parameter β. This is similar to the cases in bi-WE and bi-GE models. See Kundu and Gupta (2005, 2006).

Remark 2: (i) when $\alpha_1 = \alpha_2 = 1$, (3.1.3) reduces to be (3.1.1). (ii) when $\lambda_1 = \lambda_2$ (easy to verify), (3.1.3) reduces to be (3.1.2). To see (ii), we actually prove the following corollary.

Corollary 3.1.2. *For $\alpha_1 > 0$ and $\alpha_2 > 0$, we have*

$$\sum_{i=0}^{\infty}\sum_{j=1}^{\infty}\binom{\alpha_2}{i}\binom{\alpha_1}{j}(-1)^{i+j-1}\frac{j}{j+i} = \frac{\alpha_1}{\alpha_1+\alpha_2}. \tag{3.1.4}$$

Proof: Denote

$$I = \frac{\partial}{\partial s}\int_0^1 \sum_{i=0}^{\infty}\sum_{j=1}^{\infty}\binom{\alpha_2}{i}\binom{\alpha_1}{j}(-1)^{i+j-1}t^{i+j-1}s^j dt.$$

Then, on one hand we have

$$\begin{aligned} I &= \frac{\partial}{\partial s}\int_0^1 \sum_{i=0}^{\infty}\binom{\alpha_2}{i}(-1)^i t^{i-1}\sum_{j=1}^{\infty}\binom{\alpha_1}{j}(-1)^{j-1}(ts)^j dt \qquad (3.1.5)\\ &= \frac{\partial}{\partial s}\int_0^1 \frac{1}{t(1-t)^{\alpha_2}\left(1-(1-ts)^{\alpha_1}\right)}dt \\ &= \int_0^1 \frac{1}{t}(1-t)^{\alpha_2}(1-ts)^{\alpha_1-1}\alpha_1 t dt = \alpha_1\int_0^1 (1-t)^{\alpha_2}(1-ts)^{\alpha_1-1}dt. \end{aligned}$$

On the other hand, we have

$$\begin{aligned} I &= \frac{\partial}{\partial s}\sum_{i=0}^{\infty}\sum_{j=1}^{\infty}\binom{\alpha_2}{i}\binom{\alpha_1}{j}(-1)^{i+j-1}s^j\frac{t^{i+j}}{i+j} \qquad (3.1.6)\\ &= \sum_{i=0}^{\infty}\sum_{j=1}^{\infty}\binom{\alpha_2}{i}\binom{\alpha_1}{j}(-1)^{i+j-1}\frac{j}{i+j}s^{j-1}t^{i+j}. \end{aligned}$$

(3.1.7)

Let $s = 1$, then (3.1.5) implies

$$I = \frac{\alpha_1}{\alpha_1+\alpha_2}$$

while (3.1.6) implies

$$I = \sum_{i=0}^{\infty}\sum_{j=1}^{\infty}\binom{\alpha_2}{i}\binom{\alpha_1}{j}(-1)^{i+j-1}\frac{j}{i+j}.$$

This completes the proof of the corollary.

Now we study the MLE of AUC. First we need to obtain the MLE of the parameters. Then plug into 3.1.3 to get the MLE of AUC. Let $X_1, X_2, \cdots X_m$ and $Y_1, Y_2 \cdots Y_n$ be independent random samples from $X \sim F_1$ and $Y \sim F_0$ respectively. Denote the model parameter vector by $\theta = (\theta^{(i)}) = (\alpha_1, \alpha_2, \beta, \lambda_1, \lambda_2)$ $(i = 1, 2, 3, 4, 5)$.
Then the log-likelihood function is

$$\begin{aligned} l(\theta) &= m\ln\alpha_1 + m\ln\lambda_1 + m\ln\beta + n\ln\alpha_2 + n\ln\lambda_2 + n\ln\beta + \\ &\quad + (\alpha_1 - 1)\sum_{i=1}^{m}\ln(1 - e^{-\lambda_1 x_i^\beta}) + (\alpha_2 - 1)\sum_{j=1}^{n}\ln(1 - e^{-\lambda_2 y_i^\beta}) \\ &\quad - \sum_{i=1}^{m}\lambda_1 x_i^\beta + (\beta - 1)\sum_{i=1}^{m}\ln x_i - \sum_{j=1}^{n}\lambda_2 y_j^\beta + (\beta - 1)\sum_{j=1}^{n}\ln y_i \end{aligned}$$

Take the derivative with respect to the parameters to get the following score equations.

$$\begin{aligned} \frac{\partial l}{\partial \alpha_1} &= \frac{m}{\alpha_1} + \sum_{i=1}^{m}\ln(1 - e^{-\lambda_1 x_i^\beta}) = 0, \\ \frac{\partial l}{\partial \alpha_2} &= \frac{n}{\alpha_2} + \sum_{j=1}^{n}\ln(1 - e^{-\lambda_2 y_j^\beta}) = 0, \\ \frac{\partial l}{\partial \lambda_1} &= \frac{m}{\lambda_1} - \sum_{i=1}^{m} x_i^\beta + (\alpha_1 - 1)\sum_{i=1}^{m}\frac{e^{-\lambda_1 x_i^\beta} x_i^\beta}{1 - e^{-\lambda_1 x_i^\beta}} = 0, \\ \frac{\partial l}{\partial \lambda_2} &= \frac{n}{\lambda_2} - \sum_{j=1}^{n} y_j^\beta + (\alpha_2 - 1)\sum_{j=1}^{n}\frac{e^{-\lambda_2 y_j^\beta} y_j^\beta}{1 - e^{-\lambda_2 y_j^\beta}} = 0, \\ \frac{\partial l}{\partial \beta} &= \frac{m+n}{\beta} - \lambda_1\sum_{i=1}^{m} x_i^\beta \ln x_i - \lambda_2\sum_{j=1}^{n} y_j^\beta \ln y_j + \sum_{i=1}^{m}\ln x_i + \sum_{j=1}^{n}\ln y_j \\ &\quad + (\alpha_1 - 1)\sum_{i=1}^{m}\frac{e^{-\lambda_1 x_i^\beta}\lambda_1 x_i^\beta \ln x_i}{1 - e^{-\lambda_1 x_i^\beta}} + (\alpha_2 - 1)\sum_{j=1}^{n}\frac{e^{-\lambda_2 y_j^\beta}\lambda_2 y_j^\beta \ln y_j}{1 - e^{-\lambda_2 y_j^\beta}} = 0. \end{aligned}$$

The MLE of $\widehat{\theta}$ can be obtained as the solutions of the above score equations by

numerical methods. Plugging it into 3.1.3, we can obtain an estimator of AUC:

$$\widehat{AUC} = \sum_{i=0}^{\infty}\sum_{j=1}^{\infty}\binom{\widehat{\alpha_2}}{i}\binom{\widehat{\alpha_1}}{j}(-1)^{i+j-1}\frac{\widehat{\lambda_1}j}{\widehat{\lambda_1}j+\widehat{\lambda_2}i}.$$

3.2 ASYMPTOTIC NORMALITY OF THE ESTIMATED AUC

In this section, we will first obtain the asymptotic distribution of $\widehat{\theta} = (\widehat{\alpha_1}, \widehat{\alpha_2}, \widehat{\beta}, \widehat{\lambda_1}, \widehat{\lambda_2})$. Then we derive the asymptotic distribution of AUC. We denote the Fisher information matrix of $\theta = (\alpha_1, \alpha_2, \beta, \lambda_1, \lambda_2)$ as $\mathbf{I}(\theta) = (I_{ij}(\theta)), i, j = 1, 2, 3, 4, 5$. Therefore,

$$\mathbf{I}(\theta) = -\begin{pmatrix} E(\frac{\partial^2 l}{\partial\alpha_1^2}) & E(\frac{\partial^2 l}{\partial\alpha_1\partial\alpha_2}) & E(\frac{\partial^2 l}{\partial\alpha_1\partial\beta}) & E(\frac{\partial^2 l}{\partial\alpha_1\partial\lambda_1}) & E(\frac{\partial^2 l}{\partial\alpha_1\partial\lambda_2}) \\ E(\frac{\partial^2 l}{\partial\alpha_1\partial\alpha_2}) & E(\frac{\partial^2 l}{\partial\alpha_2^2}) & E(\frac{\partial^2 l}{\partial\alpha_2\partial\beta}) & E(\frac{\partial^2 l}{\partial\alpha_2\partial\lambda_1}) & E(\frac{\partial^2 l}{\partial\alpha_2\partial\lambda_2}) \\ E(\frac{\partial^2 l}{\partial\beta\partial\alpha_1}) & E(\frac{\partial^2 l}{\partial\beta_1\partial\alpha_2}) & E(\frac{\partial^2 l}{\partial\beta^2}) & E(\frac{\partial^2 l}{\partial\beta\partial\lambda_1}) & E(\frac{\partial^2 l}{\partial\beta\partial\lambda_2}) \\ E(\frac{\partial^2 l}{\partial\lambda_1\partial\alpha_1}) & E(\frac{\partial^2 l}{\partial\lambda_1\partial\alpha_2}) & E(\frac{\partial^2 l}{\partial\lambda_1\partial\beta}) & E(\frac{\partial^2 l}{\partial\lambda_1^2}) & E(\frac{\partial^2 l}{\partial\lambda_1\partial\lambda_2}) \\ E(\frac{\partial^2 l}{\partial\lambda_2\partial\alpha_1}) & E(\frac{\partial^2 l}{\partial\lambda_2\partial\alpha_2}) & E(\frac{\partial^2 l}{\partial\lambda_2\partial\beta}) & E(\frac{\partial^2 l}{\partial\lambda_2\partial\lambda_1}) & E(\frac{\partial^2 l}{\partial\lambda_2^2}) \end{pmatrix}.$$

Moreover,

$$I_{11} = -E(\frac{\partial^2 l}{\partial\alpha_1^2}) = \frac{m}{\alpha_1^2},$$

$$I_{12} = I_{21} = -E(\frac{\partial^2 l}{\partial\alpha_1\partial\alpha_2}) = 0,$$

$$I_{13} = I_{31} = -E(\frac{\partial^2 l}{\partial\alpha_1\partial\beta}) = -m\lambda_1\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{(-\lambda_1(1+i))^j}{j!}E(X^{\beta(1+j)}\ln X),$$

$$I_{14} = I_{41} = -E(\frac{\partial^2 l}{\partial\alpha_1\partial\lambda_1}) = -m\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{(-\lambda_1(1+i))^j}{j!}E(X^{\beta(j+1)}),$$

$$I_{15} = I_{51} = -E(\frac{\partial^2 l}{\partial\alpha_1\partial\lambda_2}) = 0,$$

$$I_{22} = -E(\frac{\partial^2 l}{\partial\alpha_2^2}) = \frac{n}{\alpha_2^2},$$

$$I_{23} = I_{32} = -E(\frac{\partial^2 l}{\partial\alpha_2\partial\beta}) = -n\lambda_2 \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{(-\lambda_2(1+i))^j}{j!} E(Y^{\beta(1+j)} \ln Y),$$

$$I_{24} = I_{42} = -E(\frac{\partial^2 l}{\partial\alpha_2\partial\lambda_1}) = 0,$$

$$I_{25} = I_{52} = -E(\frac{\partial^2 l}{\partial\alpha_2\partial\lambda_2}) = -n \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{(-\lambda_2(1+i))^j}{j!} E(Y^{\beta(j+1)}),$$

$$\begin{aligned}
I_{33} &= -E(\frac{\partial^2 l}{\partial\beta^2}) = m\lambda_1 E(X^\beta(\ln X)^2) + n\lambda_2 E(Y^\beta(\ln Y)^2) + \frac{m+n}{\beta^2} \\
&\quad -m\lambda_1(\alpha_1 - 1)\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{(1-\lambda_1(1+i))}{j!}(-\lambda_1(1+i))^j E(X^{\beta(j+1)}(\ln X)^2) \\
&\quad -n\lambda_2(\alpha_2 - 1)\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{(1-\lambda_2(1+i))}{j!}(-\lambda_2(1+i))^j E(Y^{\beta(j+1)}(\ln Y)^2), \\
I_{34} &= I_{43} = -E(\frac{\partial^2 l}{\partial\lambda_1\partial\beta}) \\
&= mE(X^\beta \ln X) - m(\alpha_1 - 1)\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{(1-\lambda_1(1+i))}{j!}(-\lambda_1(1+i))^j E(X^{\beta(j+1)} \ln X), \\
I_{35} &= I_{53} = -E(\frac{\partial^2 l}{\partial\lambda_2\partial\beta}) \\
&= nE(Y^\beta \ln Y) - n(\alpha_2 - 1)\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{(1-\lambda_2(1+i))}{j!}(-\lambda_2(1+i))^j E(Y^{\beta(j+1)} \ln Y),
\end{aligned}$$

$$I_{44} = -E(\frac{\partial^2 l}{\partial\lambda_1^2}) = \frac{m}{\lambda_1^2} - m(\alpha_1 - 1)\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{(-\lambda_1)^j(i+1)^{j+1}}{j!} E(X^{\beta(j+2)}),$$

$$I_{45} = I_{54} = E(\frac{\partial^2 l}{\partial\lambda_1\partial\lambda_2}) = 0,$$

$$I_{55} = -E(\frac{\partial^2 l}{\partial\lambda_2^2}) = \frac{n}{\lambda_2^2} + n(\alpha_2 - 1)\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{(-\lambda_2)^j(i+1)^{j+1}}{j!} E(Y^{\beta(j+2)}).$$

Since the EW family almost everywhere satisfies all the regularity conditions, we have the following result:

Theorem 3.2.1. *Suppose (X, Y) follows the bi-EW model with the distribution functions $X \sim F_1(x; \alpha_1, \beta, \lambda_1) = (1 - e^{-\lambda_1 x^\beta})^{\alpha_1}$ and $Y \sim F_0(y; \alpha_2, \beta, \lambda_2) = (1 - e^{-\lambda_2 y^\beta})^{\alpha_2}$,*

respectively. Let $m \to \infty$ *and* $n \to \infty$*, but* $\frac{m}{n} \to p$*, where* $0 < p < \infty$*. Then there exists a consistent sequence* $\widehat{\theta_{m,n}}$ *of the roots of the score equation such that*

$$\left(\sqrt{m}(\widehat{\alpha}_1 - \alpha_1), \sqrt{n}(\widehat{\alpha}_2 - \alpha_2), \sqrt{m}(\widehat{\beta} - \beta), \sqrt{m}(\widehat{\lambda_1} - \lambda_1), \sqrt{n}(\widehat{\lambda_2} - \lambda_2)\right) \to N_5\left(0, \mathbf{A}^{-1}(\theta)\right), \tag{3.2.8}$$

where

$$\mathbf{A}(\theta) = \begin{pmatrix} a_{11} & 0 & a_{13} & a_{14} & 0 \\ 0 & a_{22} & a_{23} & 0 & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & 0 & a_{43} & a_{44} & 0 \\ 0 & a_{52} & a_{53} & 0 & a_{55} \end{pmatrix}. \quad \textit{(see below)}$$

The proof of the Theorem. It follows by expanding the derivative of the log-lokelihood function using Taylor series and the central limit theorem.

Remark: We denote $D = diag(\sqrt{m}, \sqrt{n}, \sqrt{m}, \sqrt{m}, \sqrt{n})$. Then $\mathbf{A}(\theta) = \lim_{m,n\to\infty} D\mathbf{I}(\theta)D$.

Let $\mu_{jk}(x) = E(X^{\beta(j+1)}(\ln X)^k), (k = 0, 1, 2, j = 0, 1, 2, ..)$, then we have:

$$a_{11} = \frac{1}{\alpha_1^2},$$

$$a_{12} = a_{21} = 0,$$

$$a_{13} = a_{31} = -\lambda_1 \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{(-\lambda_1(1+i))^j}{j!}\mu_{j,1}(y),$$

$$a_{14} = a_{41} = -\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{(-\lambda_1(1+i))^j}{j!}\mu_{j,0}(x),$$

$$a_{15} = a_{51} = 0,$$

$$a_{22} = \frac{1}{\alpha_2^2},$$

,

$$a_{23} = a_{32} = -\frac{1}{\sqrt{p}}\lambda_2 \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{(-\lambda_2(1+i))^j}{j!}\mu_{j,1}(y),$$

$$a_{24} = a_{42} = 0,$$

$$a_{25} = a_{52} = -\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{(-\lambda_2(1+i))^j}{j!}\mu_{j,0}(y),$$

$$\begin{aligned} a_{33} &= \lambda_1\mu_{0,2}(x) + \frac{1}{\sqrt{p}}\lambda_2\mu_{0,2}(y) + \frac{1+p}{\sqrt{p}\beta^2} \\ &\quad - \lambda_1(\alpha_1 - 1)\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{(1-\lambda_1(1+i))}{j!}(-\lambda_1(1+i))^j\mu_{j,2}(x) \\ &\quad - \frac{1}{\sqrt{p}}\lambda_2(\alpha_2 - 1)\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{(-\lambda_1(1+i))}{j!}(1-\lambda_2(1+i))^j\mu_{j,2}(y), \end{aligned}$$

$$a_{34} = a_{43} = \mu_{0,1}(x) - (\alpha_1 - 1)\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{(1-\lambda_1(1+i))}{j!}(-\lambda_1(1+i))^j\mu_{j,1}(x),$$

$$a_{35} = a_{53} = \frac{1}{\sqrt{p}}\mu_{0,1}(y) - \frac{1}{\sqrt{p}}(\alpha_2 - 1)\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{(1-\lambda_2(1+i))}{j!}(-\lambda_2(1+i))^j\mu_{j,1}(y),$$

$$a_{44} = \frac{1}{\lambda_1^2} + (\alpha_1 - 1)\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{(i+1)^{j+1}}{j!}(-\lambda_1)^j\mu_{j+1,0}(x),$$

$$a_{45} = a_{54} = 0,$$

$$a_{55} = \frac{1}{\lambda_2^2} + (\alpha_2 - 1)\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{(i+1)^{j+1}}{j!}(-\lambda_2)^j\mu_{j+1,0}(y).$$

Theorem 3.2.2. *Suppose (X, Y) follows a bi-EW model with the distribution functions $X \sim F_1(x;\alpha_1,\beta,\lambda_1) = (1-e^{-\lambda_1 x^\beta})^{\alpha_1}$ and $Y \sim F_0(y;\alpha_2,\beta,\lambda_2) = (1-e^{-\lambda_2 y^\beta})^{\alpha_2}$, respectively. Let $m \to \infty$ and $n \to \infty$, but $\frac{m}{n} \to p$, where $0 < p < \infty$, then we have*

$$\sqrt{m}(\widehat{AUC} - AUC) \to N(0, B),$$

where $B = \sum_{i=1}^{5}(\frac{\partial f}{\partial\theta^{(i)}})^2 Var(\widehat{\theta}^{(i)}) + 2\sum_{i\neq j}\frac{\partial^2 f}{\partial\theta^{(i)}\partial\theta^{(j)}}Cov(\widehat{\theta}^{(i)},\widehat{\theta}^{(j)}),$

$$f(\theta) = AUC = \int_0^{\infty} F_0(x;\alpha_2,\beta,\lambda_2)dF_1(x;\alpha_1,\beta,\lambda_1), \tag{3.2.9}$$

$\theta = (\theta^{(i)}) = (\alpha_1,\alpha_2,\beta,\lambda_1,\lambda_2)$ $(i = 1,2,3,4,5)$.

Proof of Theorem. It follows from Theorem 3.2.1.

3.3 NUMERICAL EXPERIMENTS

In this section we present some results based on Monte Caro simulations to check the performance of the MLE methods to estimate parameters and AUC under the moderate sample sizes. We first consider the following moderate sample sizes: $(m, n) = (50, 50), (80, 80), (100, 100), (50, 80), (80, 50), (50, 100), (100, 50), (80, 100), (100, 80)$ plus two large sample sizes: $(150, 150), (200, 200)$. We set two scale parameters $\lambda_1 = \lambda_2 = 1$. The common shape parameter $\beta = 2$, one shape parameter $\alpha_1 = 2$ and $\alpha_2 = 1, 2, 3, 4$ respectively. The simulations are based on 1000 replications. We report absolute relative biases (ARB) and the square root of mean squared errors (RMSE) for the estimators of the five parameters and AUC.

$\text{ARB}(\widehat{\theta^{(j)}}) = \frac{1}{1000}\sum_{i=1}^{1000}\left|\frac{\widehat{\theta}_i^{(j)} - \theta^{(j)}}{\theta^{(j)}}\right|$ and $\text{RMSE}(\widehat{\theta^{(j)}}) = \sqrt{\frac{1}{1000}\sum_{i=1}^{1000}(\widehat{\theta}_i^{(j)} - \theta^{(j)})^2}$ $(j = 1, 2, 3, 4, 5)$, where $\widehat{\theta}_i^{(j)}$ is the estimate of $\theta^{(j)}$ for the $i-$th replicate.

From (3.2.9), we have

$$AUC = \int_0^\infty F_0(x; \alpha_2, \beta, \lambda_2) dF_1(x; \alpha_1, \beta, \lambda_1).$$

The estimator $\widehat{AUC}$ and true value AUC can be computed by the summation

$$\sum_{x=0.0001}^{b} F_0(x; \widehat{\alpha_2}, \widehat{\beta}, \widehat{\lambda_2}) f_1(x; \widehat{\alpha_1}, \widehat{\beta}, \widehat{\lambda_1}) \Delta x,$$

and

$$\sum_{x=0.0001}^{b} F_0(x; \alpha_2, \beta, \lambda_2) f_1(x; \alpha_1, \beta, \lambda_1) \Delta x$$

respectively.

We choose $b = 4$ and $\Delta x = 0.0001$. The interval $[0, 4]$ is divided evenly and each subinterval length is 0.0001. x takes value of right end of each subinterval. We have verified that the summation keeps stable for any larger value b and smaller value Δx. Similarly,

$$\text{ARB}(\widehat{AUC}) = \frac{1}{1000}\sum_{i=1}^{1000}\left|\frac{\widehat{AUC}_i - AUC}{AUC}\right|, \text{RMSE}(\widehat{AUC}) = \sqrt{\frac{1}{1000}\sum_{i=1}^{1000}(\widehat{AUC}_i - AUC)^2},$$

where $\widehat{AUC}_i$ is the estimate of AUC for the $i-$th replicate.

Tables 3.3.1 and 3.3.2 report the ARB and RMSE for the estimators of $\alpha_1, \alpha_2, \beta, \lambda_1$, λ_2 and the AUC in row order.

The first simulation results are reported in Table 3.3.1. A obvious fact is the performance of the MLE for AUC is quite satisfactory in terms of ARB and RMSE under all settings of sample sizes and true parameters. In all cases, the two scale parameters have much smaller ARB and RMSE compared with shape parameters. For two different shape parameters, it is always the common parameter β which behaves better than the other two shape parameters α_1 and α_2. For two shape parameters α_1 and α_2, when $\alpha_2 > \alpha_1, \alpha_1$ always behaves better than α_2. It is reasonable that the ARB and RMSE decrease as the sample size increases.

Our second simulations are to check the performance of MLE as the common shape parameter β increases. We choose the following sample sizes $(m, n) = (50, 50), (80, 80)$, $(50, 80), (80, 50)$. The two scale parameters are still set to 1, and $\alpha_1 = 2, \alpha_2 = 3$. β takes the value: $2, 2.5, 3$. The results are reported in Table 3.3.2. We still find that AUC has very small ARB and RMSE. Under all settings, α_2 has larger ARB and RMSE than that of α_1 when $\alpha_2 > \alpha_1$. When β increases, we do not find a noticeable change of ARB and RMSE of all parameters.

We present the histogram of estimated AUC under the sample sizes $(m, n) = (50, 50), (100, 100), (50, 100), (100, 200)$ for the true parameters: $\alpha_1 = 2, \alpha_2 = 4, \beta = 2, \lambda_1 = \lambda_2 = 1$. Obviously, with the sample increasing, the histogram tends to be more symmetric, bell shaped. Therefore, the MLE method performs very well to estimate AUC under the moderate sample sizes.

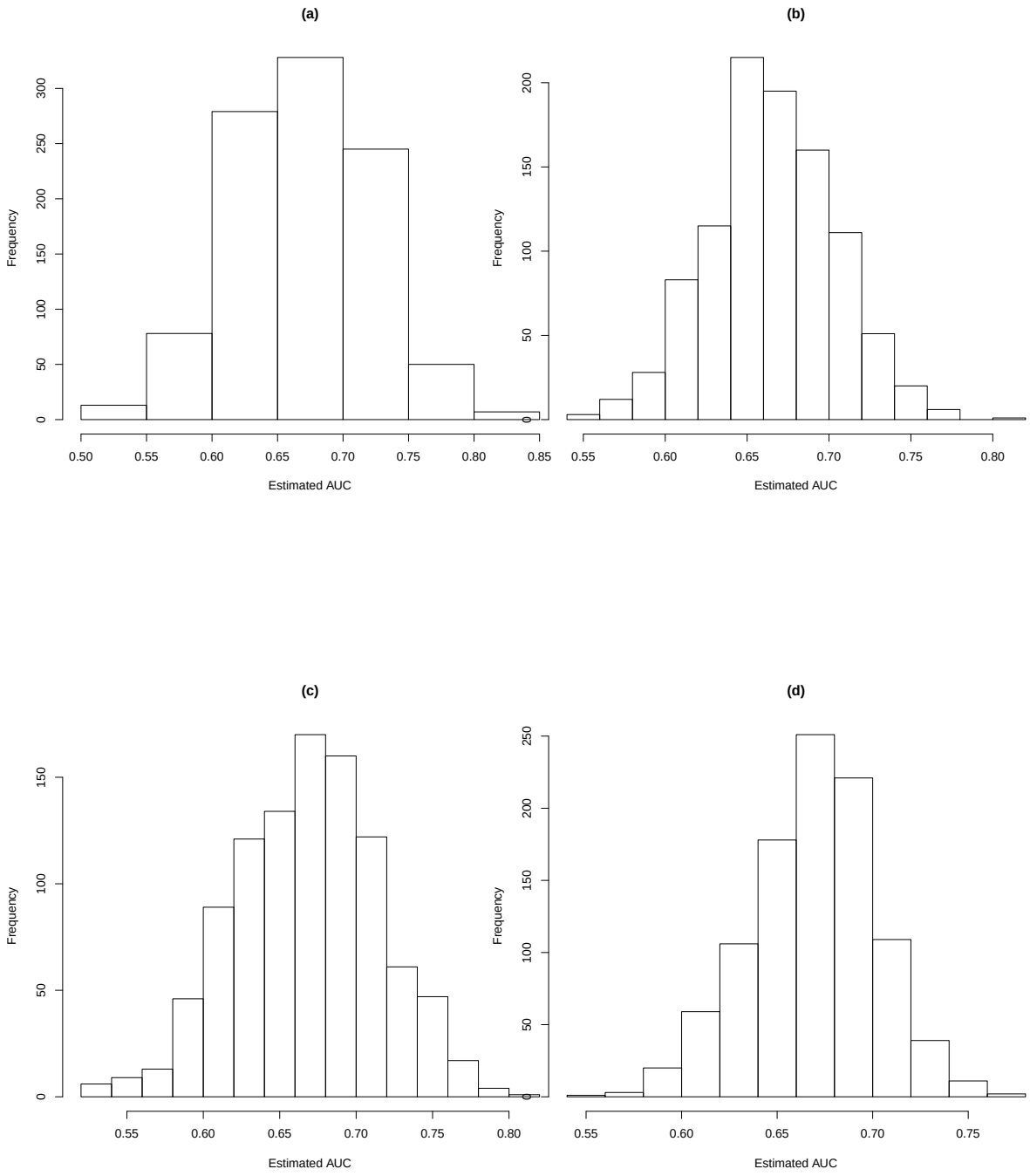

Figure 3.3.1: Histogram of estimated AUC for (a): (m, n)=(50, 50), (b): (m, n)=(100, 100), (c): (m, n)=(50, 100), (d): (m,n)=(100, 200).

Table 3.3.1: ARB and RMSE of five estimated parameters $(\widehat{\alpha_1}, \widehat{\alpha_2}, \widehat{\beta}, \widehat{\lambda_1}, \widehat{\lambda_2})$ and $\widehat{AUC}$, when $\alpha_1 = 2, \beta = 2, \lambda_1 = 1, \lambda_2 = 1$ and different values of α_2

(m,n)		$\alpha_2 = 1$	$\alpha_2 = 2$	$\alpha_2 = 3$	$\alpha_2 = 4$
	$\widehat{\alpha_1}$	0.3721(1.0569)	0.35630(1.111)	0.3607(1.1414)	0.3995(1.1.350)
	$\widehat{\alpha_2}$	0.2556(1.1068)	0.3645(1.2601)	0.4204(1.5814)	0.5547(2.1276)
(50,50)	$\widehat{\beta}$	0.0801(0.7525)	0.0675(0.6655)	0.0615(0.6679)	0.0525(0.6243)
	$\widehat{\lambda_1}$	0.0842(0.5661)	0.0756(0.5421)	0.961(0.5442)	0.0979(0.5683)
	$\widehat{\lambda_2}$	0.0913(0.5235)	0.0805(0.5493)	0.0916(0.560)	0.119(0.6189)
	$\widehat{AUC}$	0.0309(0.0648)	0.0187(0.0734)	0.0074(0.08)	0.0078(0.0648)
	$\widehat{\alpha_1}$	0.1669(1.500)	0.2038(1.603)	0.2215(1.566)	0.1776(1.4947)
	$\widehat{\alpha_2}$	0.1026(0.5253)	0.1941(1.6007)	0.2508(2.2141)	0.2545(2.0528)
(80,80)	$\widehat{\beta}$	0.0517(0.5167)	0.0396(0.4933)	0.0280(0.4878)	0.0338(2.0453)
	$\widehat{\lambda_1}$	0.0318(0.3521)	0.0354(0.3687)	0.0469(0.3720)	0.0369(0.3644)
	$\widehat{\lambda_2}$	0.0340(0.3714)	0.0535(0.4252)	0.0733(0.4549)	0.0712(0.4411)
	$\widehat{AUC}$	0.0123(0.3755)	0.01061(0.064)	0.0035(0.064)	0.0076(0.06)
	$\widehat{\alpha_1}$	0.1205(1.1180)	0.1322(1.2528)	0.1396(1.2052)	0.1276(1.1109)
	$\widehat{\alpha_2}$	0.0726(0.4195)	0.1291(1.2499)	0.1842(2.1241)	0.1295(2.0528)
(100,100)	$\widehat{\beta}$	0.0367(0.4324)	0.0346(0.4317)	0.0290(0.4219)	0.0323(0.4281)
	$\widehat{\lambda_1}$	0.0318(0.3521)	0.0354(0.3687)	0.0469(0.3720)	0.0369(0.3644)
	$\widehat{\lambda_2}$	0.029(0.3286)	0.0323(0.3739)	0.0593(0.3962)	0.0452(0.3954)
	$\widehat{AUC}$	0.0036(0.0412)	0.003(0.0458)	0.0018(0.0331)	0.0031(0.040)
	$\widehat{\alpha_1}$	0.2792(2.2676)	0.2703(2.067)	0.3279(1.3403)	0.2876(1.4635)
	$\widehat{\alpha_2}$	0.1368(0.6440)	0.2144(1.8031)	0.3663(1.4655)	0.3589(1.8229)
(50,80)	$\widehat{\beta}$	0.0581(0.5743)	0.0435(0.5290)	0.0342(0.5362)	0.0473(0.5618)
	$\widehat{\lambda_1}$	0.0665(0.4802)	0.0811(0.4819)	0.0979(0.4999)	0.0789(0.4889)
	$\widehat{\lambda_2}$	0.0487(0.4154)	0.0544(0.4590)	0.0939(0.5176)	0.0737(0.5265)
	$\widehat{AUC}$	0.0105(0.0468)	0.0068(0.0529)	0.0053(0.0519)	0.0054(0.0424)

	$\widehat{\alpha_1}$	0.1955(2.2229)	0.1875(1.5019)	0.2273(1.7404)	0.1815(1.5632)
	$\widehat{\alpha_2}$	0.1319(0.6757)	0.2355(1.6962)	0.3375(1.7291)	0.2930(1.9740)
(80,50)	$\widehat{\beta}$	0.0706(0.5891)	0.0454(0.5334)	0.0473(0.5547)	0.0606(0.5592)
	$\widehat{\lambda_1}$	0.0309(0.2158)	0.0524(0.1971)	0.0635(0.2172)	0.0376(0.1913)
	$\widehat{\lambda_2}$	0.0488(0.4342)	0.0743(0.4556)	0.0845(0.5102)	0.0481(0.4881)
	$\widehat{AUC}$	0.0108(0.0458)	0.0074(0.0509)	0.0058(0.050)	0.0059(0.0435)
	$\widehat{\alpha_1}$	0.1872(1.6205)	0.3035(2.1506)	0.2333(1.5885)	0.2205(1.9536)
	$\widehat{\alpha_2}$	0.0823(0.4995)	0.2404(1.8192)	0.2507(2.6231)	0.2896(2.1170)
(50,100)	$\widehat{\beta}$	0.0636(0.5464)	0.0293(0.5054)	0.0271(0.4935)	0.0467(0.5228)
	$\widehat{\lambda_1}$	0.0348(0.4380)	0.0985(0.4777)	0.0859(0.4421)	0.0653(0.4614)
	$\widehat{\lambda_2}$	0.0217(0.3694)	0.0751(0.4532)	0.0762(0.4510)	0.0612(0.4861)
	$\widehat{AUC}$	0.0072(0.060)	0.005(0.0591)	0.0037(0.0565)	0.0058(0.0424)
	$\widehat{\alpha_1}$	0.1609(1.3638)	0.1689(1.5535)	0.1774(1.3754)	0.1731(1.4751)
	$\widehat{\alpha_2}$	0.1379(0.5709)	0.2031(1.6370)	0.2773(1.9342)	0.3209(1.7746)
(100,50)	$\widehat{\beta}$	0.0467(0.5281)	0.0455(0.4998)	0.0320(0.4814)	0.0425(0.5079)
	$\widehat{\lambda_1}$	0.0399(0.4085)	0.0430(0.4129)	0.0602(0.3962)	0.0461(0.4191)
	$\widehat{\lambda_2}$	0.0709(0.4078)	0.0524(0.4300)	0.0824(0.4498)	0.0668(0.4889)
	$\widehat{AUC}$	0.0048(0.0519)	0.0031(0.0489)	0.0017(0.0469)	0.0056(0.0435)
	$\widehat{\alpha_1}$	0.1306(1.1523)	0.1332(1.2649)	0.1282(1.2685)	0.1989(1.4240)
	$\widehat{\alpha_2}$	0.0731(0.4401)	0.1305(1.1954)	0.0846(1.1941)	0.2563(2.209)
(80,100)	$\widehat{\beta}$	0.0396(0.4595)	0.0315(0.4313)	0.0596(0.6750)	0.0308(0.6548)
	$\widehat{\lambda_1}$	0.0383(0.3716)	0.0407(0.3686)	0.0301(0.3671)	0.0772(0.3977)
	$\widehat{\lambda_2}$	0.0190(0.3325)	0.0361(0.3649)	0.0331(0.3657)	0.0739(0.4317)
	$\widehat{AUC}$	0.00216(0.037)	0.0018(0.0424)	0.001(0.0412)	0.0048(0.0316)

	$\widehat{\alpha_1}$	0.1412(1.2018)	0.1587(1.4156)	0.1137(1.1747)	0.1924(1.3352)
	$\widehat{\alpha_2}$	0.0904(0.4694)	0.1703(1.3251)	0.1662(2.140)	0.2690(2.204)
(100,80)	$\widehat{\beta}$	0.0648(0.6946)	0.0534(0.6871)	0.0557(0.6549)	0.0270(0.6740)
	$\widehat{\lambda_1}$	0.0398(0.3809)	0.0425(0.3907)	0.0272(0.3616)	0.0755(0.3981)
	$\widehat{\lambda_2}$	0.0326(0.3587)	0.0515(0.3976)	0.0361(0.3926)	0.0828(0.4393)
	$\widehat{AUC}$	0.0024(0.0374)	0.001(0.0412)	0.0023(0.04)	0.0019(0.0387)
	$\widehat{\alpha_1}$	0.0672(0.7867)	0.1004(0.8619)	0.1012(0.8744)	0.0959(0.7899)
	$\widehat{\alpha_2}$	0.0474(0.3051)	0.1004(0.8683)	0.1269(1.5003)	0.1234(1.9684)
(150,150)	$\widehat{\beta}$	0.0252(0.3380)	0.0132(0.3341)	0.0120(0.3322)	0.0118(0.3226)
	$\widehat{\lambda_1}$	0.0206(0.2857)	0.0392(0.2993)	0.0409(0.2985)	0.0415(0.2813)
	$\widehat{\lambda_2}$	0.0186(0.2543)	0.0384(0.2939)	0.0478(0.3140)	0.0406(0.3096)
	$\widehat{AUC}$	0.0013(0.0294)	0.0009(0.0328)	0.0010(0.0306)	0.0011(0.0301)
	$\widehat{\alpha_1}$	0.0668(0.6871)	0.0569(0.6666)	0.0700(0.6368)	0.0602(0.6553)
	$\widehat{\alpha_2}$	0.0404(0.2707)	0.0637(0.6581)	0.0833(1.1019)	0.0965(0.7324)
(200,200)	$\widehat{\beta}$	0.0152(0.2879)	0.0135(0.2853)	0.0090(0.2791)	0.0102(0.2756)
	$\widehat{\lambda_1}$	0.0214(0.2487)	0.0215(0.2457)	0.0324(0.2399)	0.0217(0.2430)
	$\widehat{\lambda_2}$	0.0162(0.2274)	0.0222(0.2419)	0.0292(0.2562)	0.0280(0.2715)
	$\widehat{AUC}$	0.0021(0.0256)	0.0033(0.0267)	0.0031(0.0270)	0.0013(0.0264)

Table 3.3.2: ARB and RMSE of five estimated parameters $(\widehat{\alpha_1}, \widehat{\alpha_2}, \widehat{\beta}, \widehat{\lambda_1}, \widehat{\lambda_2})$ and $\widehat{AUC}$, when $\alpha_1 = 2, \alpha_2 = 3, \lambda_1 = 1, \lambda_2 = 1$ and different values of β

(m,n)		$\beta = 2$	$\beta = 2.5$	$\beta = 3$
	$\widehat{\alpha_1}$	0.1925(1.3453)	0.1646(1.4159)	0.1863(1.4226)
	$\widehat{\alpha_2}$	0.4685(2.245)	0.3917(1.9021)	0.4641(2.348)
(50,50)	$\widehat{\beta}$	0.0554(0.6402)	0.0912(0.7801)	0.0558(0.9117)
	$\widehat{\lambda_1}$	0.1044(0.5624)	0.0700(0.5442)	0.0800(0.5353)
	$\widehat{\lambda_2}$	0.1048(0.5779)	0.0746(0.5624)	0.0939(0.5719)
	$\widehat{AUC}$	0.0067(0.0556)	0.0026(0.0557)	0.0036(0.0529)
	$\widehat{\alpha_1}$	0.162(1.3518)	0.1531(1.3652)	0.1594(1.3561)
	$\widehat{\alpha_2}$	0.2096(1.8697)	0.2047(3.9241)	0.1655(2.183)
(80,80)	$\widehat{\beta}$	0.0372(0.4715)	0.0438(0.5934)	0.0458(0.7095)
	$\widehat{\lambda_1}$	0.0464(0.4011)	0.0404(0.3996)	0.0433(0.4048)
	$\widehat{\lambda_2}$	0.0488(0.4211)	0.0417(0.4334)	0.0291(0.4084)
	$\widehat{AUC}$	0.002(0.0412)	0.0023(0.0424)	0.003(0.0435)
	$\widehat{\alpha_1}$	0.2538(1.8319)	0.2472(2.0245)	0.2775(2.2910)
	$\widehat{\alpha_2}$	0.2818(2.0819)	0.2691(2.245)	0.3046(2.4255)
(50,80)	$\widehat{\beta}$	0.0401(0.5415)	0.0557(0.6907)	0.0475(0.8629)
	$\widehat{\lambda_1}$	0.0825(0.4706)	0.0635(0.4821)	0.0815(0.4803)
	$\widehat{\lambda_2}$	0.0724(0.4790)	0.0489(0.4931)	0.0741(0.4952)
	$\widehat{AUC}$	0.0045(0.0490)	0.0015(0.0510)	0.0008(0.050)
	$\widehat{\alpha_1}$	0.1846(1.6341)	0.2311(1.7011)	0.2874(1.7425)
	$\widehat{\alpha_2}$	0.2515(2.2201)	0.3733(2.2412)	0.3969(1.956)
(80,50)	$\widehat{\beta}$	0.0501(0.4921)	0.0414(0.701)	0.0339(0.8101)
	$\widehat{\lambda_1}$	0.0380(0.4365)	0.0665(0.4577)	0.0905(0.4761)
	$\widehat{\lambda_2}$	0.0443(0.4681)	0.0952(0.5278)	0.0962(0.5221)
	$\widehat{AUC}$	0.0003(0.0479)	0.0013(0.0479)	0.0055(0.0489)

CHAPTER 4

THE ESTIMATION OF AUC UNDER THE CROSSING MODEL

In ROC analysis, the most popular semi-parametric model is the bi-normal model. Here we present an alternative to the bi-normal model based on the crossing model. The crossing model was proposed by Yang and Prentice (2005) to accommodate the case when two survival curves cross. Under the crossing model, the ROC curve and its functionals (such as AUC) have a simple analytic form. Closed-form expressions for estimated AUC and its corresponding asymptotic variances are derived. The crossing model is an extension of the Lehmann family. Gonen and Heller (2007) studied the AUC estimation under the Lehmann family. Therefore, our work is a natural extension of that of Gonen and Heller (2007).

4.1 CROSSING MODELS

In survival analysis, the most widely used model is the Cox proportional hazard model. The Lehmann family is a special case of the Cox proportional hazard model. It assumes the hazard ratio is a constant. However, when the assumption of constant hazard ratio is suspected from the data evidence, accelerated failure time model and the proportional odds model are the possible alternatives. There are some situations that the data provides evidence of crossing hazard functions: a treatment has some adverse effects at the early time but later on it shows strong effect. In this case, hazard functions may cross. A choice is to use the Cox regression model with time-dependent

covariates with the effect that the hazard ratio of the two groups is a fully parametric function. Yang and Prentice (2005) developed a new two-sample semi-parametric model (crossing model) that may accommodate crossing survival functions. The two parameters in the crossing model have special properties: they are short-term and long-term hazard ratios respectively. The crossing model includes the Cox model and proportional odds model as submodels. Suppose we have two study groups: treatment and control. Their corresponding survival functions (hazard functions) are S_1 (h_1) and S_0 (h_0) respectively.

Definition 4.1.1. *Suppose two study groups' corresponding hazard rates have the relation:*

$$h_1(t) = \frac{\theta_1\theta_2}{\theta_1 + (\theta_2 - \theta_1)S_0(t)} h_0(t) \quad t < \tau_0, \tag{4.1.1}$$

where $\tau_0 = \sup\{t : S_0(t) > 0\}$*, and both* θ_1 *and* θ_2 *are positive. Then we call 4.1.1 a crossing model.*

For the crossing model, the hazard ratio between the two groups is nonconstant. If $t < \tau_0$, then the hazard ratio is

$$\frac{h_1(t)}{h_0(t)} = \frac{\theta_1\theta_2}{\theta_1 + (\theta_2 - \theta_1)S_0(t)},$$

which depends on the two parameters θ_1, θ_2 and the control survival function $S_0(t)$. The monotonicity of the ratio depends on the comparison of the two parameters θ_1 and θ_2: when $\theta_2 > \theta_1$, it is monotonic increasing; when $\theta_2 < \theta_1$, it is monotonic decreasing. Figure 4.1.1 illustrates the hazard ratios for parameters $\theta = (\theta_1, \theta_2)$ and $S_0(t) = e^{-t}$ under the two settings: $\theta = (2, 0.5)$ and $\theta = (1, 2)$.

We also find that

$$\theta_1 = \lim_{t \to 0^+} \frac{h_1(t)}{h_0(t)}, \quad \theta_2 = \lim_{t \to \tau_0} \frac{h_1(t)}{h_0(t)}, \quad \text{provided the limits exist.}$$

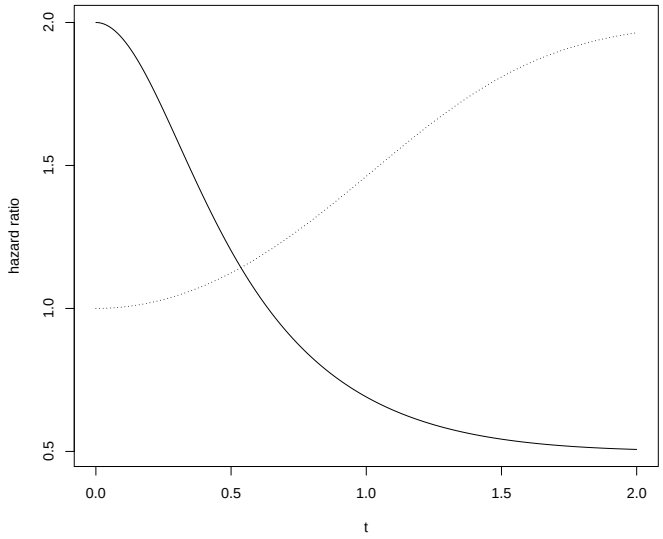

Figure 4.1.1: solid line for $\theta = (2, 0.5)$ and dotted line for $\theta = (1, 2)$

From above two equalities, the parameters θ_1 and θ_2 can be interpreted as the short-term and long-term hazard ratios respectively. The crossing model includes the Cox model and the proportional odds model as two submodels.

- When $\theta_2 = 1$, the crossing model reduces to be $\dfrac{h_1(t)}{h_0(t)} = \dfrac{\theta_1}{S_0(t) + \theta_1 F_0(t)}$. It is the proportional odds model (Bennett, 1983).
- When $\theta_1 = \theta_2$, the crossing model reduces to be $\dfrac{h_1(t)}{h_0(t)} = \theta_2$. It is the Cox proportional model (the Lehmann family in Gonen and Heller, 2007).

Let

$$R(t) = \frac{1 - S_0(t)}{S_0(t)} \quad t < \tau_0$$

be the odds of the control group. Then the survival functions of the two groups can be rewritten as

$$S_0(t) = (1 + R(t))^{-1} \quad \text{and} \quad S_1(t) = (1 + \frac{\theta_1}{\theta_2} R(t))^{-\theta_2}.$$

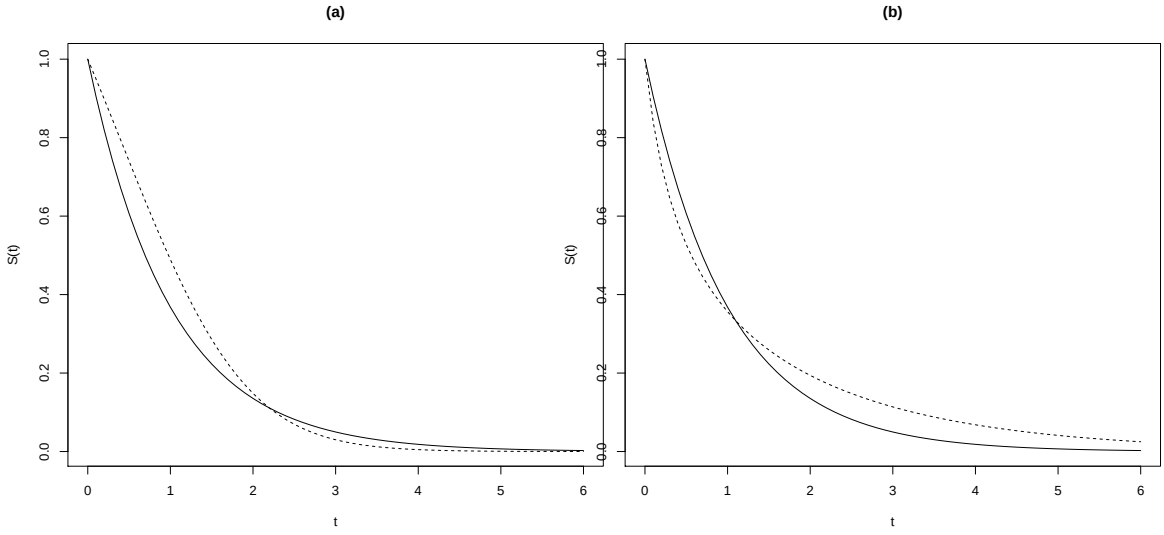

Figure 4.1.2: $S_0(t) = e^{-t}$, (a): $\theta = (0.5, 2), S_1(t) = (0.25e^t + 0.75)^{-2}$ (b): $\theta = (2, 0.5), S_1(t) = (4e^t - 3)^{-0.5}$, S_0: solid line, S_1: dotted line.

We have the following proposition:

Proposition 4.1.1. *(Yang and Prentice, 2005) For the above defined crossing model, when* $\theta_1 < 1, \theta_2 > 1$ *or* $\theta_1 > 1, \theta_2 < 1$*, the two survival functions* $S_0(t)$ *and* $S_1(t)$ *cross.*

The results of the proposition are illustrated in the Figure 4.1.2.

4.2 AUC ESTIMATION UNDER THE CROSSING MODEL

Recall that

$$ROC(t) = S_1(S_0^{-1}(t)), \quad t \in (0, 1).$$

The ROC curve is generally represented as a function of survival functions of two groups. Pepe (2003) classified the analysis of ROC curves by three approaches. Each of them requires the estimation of the survival functions. The bi-normal model as a semi-parametric model has attracted most attention. In this section, we present an alternative to the popular bi-normal model based on the crossing model. We will see

that different from the bi-normal model, the ROC curve under the crossing model only depends on the two parameters, therefore so does AUC.

Proposition 4.2.1. *Under the crossing model, the ROC function:* $ROC(t) = \left(\frac{\theta_2 t}{\theta_1 + (\theta_2 - \theta_1)t}\right)^{\theta_2}$, *where* $0 \leq t \leq 1$.

Proof: From the crossing model, we have:

$$\begin{aligned} S_1(x) &= exp\left(-\int_0^x h_1(s)ds\right) \\ &= exp\left(-\int_0^x \frac{\theta_1\theta_2 h_0(s)}{\theta_1 + (\theta_2 - \theta_1)S_0(s)}ds\right) \\ &= exp\left(\int_0^x \frac{\theta_1\theta_2 S_0'(s)}{S_0(s)(\theta_1 + (\theta_2 - \theta_1)S_0(s))}ds\right) \\ &= exp\left(\int_0^x \frac{\theta_1\theta_2 dS_0(s)}{S_0(s)(\theta_1 + (\theta_2 - \theta_1)S_0(s))}\right) \\ &= \left(\frac{\theta_2 S_0(x)}{\theta_1 + (\theta_2 - \theta_1)S_0(x)}\right)^{\theta_2}. \end{aligned}$$

Let $t = S_0(x)$, we have

$$\begin{aligned} ROC(t) &= S_1(S_0^{-1}(t)) = S_1(x) \\ &= \left(\frac{\theta_2 t}{\theta_1 + (\theta_2 - \theta_1)t}\right)^{\theta_2}. \end{aligned}$$

It is easy to see when $\theta_1 = \theta_2, ROC(t)$ reduces to be t^{θ_1}. It is the result in Gonen and Heller (2007). In following, we obtain a closed form expression for AUC under the crossing model.

Theorem 4.2.1. *Under the crossing model, if* $\frac{\theta_1}{\theta_2} < 2$ *and* $\frac{\theta_1}{\theta_2} \neq 1$, *then we have*

$$AUC = 1 - \theta_1 \sum_{i=1}^{\infty} \frac{1}{i + \theta_2}\left(\frac{\theta_2 - \theta_1}{\theta_2}\right)^{i-1}. \tag{4.2.2}$$

Proof:

$$AUC = \int_0^1 ROC(t)dt = \theta_2^{\theta_2}\int_0^1 \left(\frac{t}{\theta_1 + (\theta_2 - \theta_1)t}\right)^{\theta_2} dt.$$

Let $s = \dfrac{t}{\theta_1 + (\theta_2 - \theta_1)t}$, $dt = \dfrac{\theta_1 ds}{(1 - s(\theta_2 - \theta_1))^2}$. Then, by the condition of $\dfrac{\theta_1}{\theta_2} < 2$, we have

$$\begin{aligned}
AUC &= \frac{\theta_1 \theta_2^{\theta_2}}{\theta_2 - \theta_1} \int_0^{\frac{1}{\theta_2}} s^{\theta_2} d\frac{1}{1 - s(\theta_2 - \theta_1)} \qquad (\theta_1 \neq \theta_2) \\
&= \frac{\theta_1 \theta_2^{\theta_2}}{\theta_2 - \theta_1} \left(\frac{\theta_2}{\theta_1 \theta_2^{\theta_2}} - \theta_2 \int_0^{\frac{1}{\theta_2}} s^{\theta_2 - 1} \sum_{i=0}^{\infty} s^i (\theta_2 - \theta_1)^i ds \right) \\
&= \frac{\theta_2}{\theta_2 - \theta_1} - \frac{\theta_1 \theta_2}{\theta_2 - \theta_1} \sum_{i=0}^{\infty} \frac{1}{i + \theta_2} (\frac{\theta_2 - \theta_1}{\theta_2})^i \\
&= 1 - \frac{\theta_1 \theta_2}{\theta_2 - \theta_1} \sum_{i=1}^{\infty} \frac{1}{i + \theta_2} \left(\frac{\theta_2 - \theta_1}{\theta_2} \right)^i \\
&= 1 - \theta_1 \sum_{i=1}^{\infty} \frac{1}{i + \theta_2} \left(\frac{\theta_2 - \theta_1}{\theta_2} \right)^{i-1} .
\end{aligned}$$

Similarly, we have a partial AUC closed-form expression under the crossing model.

Theorem 4.2.2. *Under the crossing model, if* $\dfrac{\theta_1}{\theta_2} < 1$ *or* $1 < \dfrac{\theta_1}{\theta_2} < 2$, *the partial AUC at point* t_0 *is:*

$$pAUC = \frac{\theta_1 \theta_2^{\theta_2}}{\theta_2 - \theta_1} (\frac{t_0}{\theta_1 + (\theta_2 - \theta_1)t_0})^{\theta_2} \left(\frac{\theta_1 + (\theta_2 - \theta_1)t_0}{\theta_1} - \theta_2 \sum_{i=0}^{\infty} \frac{1}{i + \theta_2} \left(\frac{(\theta_2 - \theta_1)t_0}{\theta_1 + (\theta_2 - \theta_1)t_0} \right)^i \right).$$

Although the ROC curve is generally represented as a function of survival functions of diseased and non-diseased populations, the ROC curve of crossing models only depends on the two parameters, and does not require the estimation of survival functions explicitly.

4.3 THE ESTIMATED AUC AND ITS ASYMPTOTIC PROPERTY

Under the crossing model, Yang and Prentice (2005) reparametrized (θ_1, θ_2) by $\beta_1 = \ln \theta_1, \beta_2 = \ln \theta_2$ and use a pseudo maximum likelihood approach to give estimators of the two scale parameters and their asymptotic distribution.

1) Let $T_1, ..., T_n ....., T_{n+m}$ be the pooled lifetimes of the two groups, and $T_1, ..., T_n$ from the control population. Let $C_1, ..., C_{n+m}$ be the censoring variables and $Z_i = I(i > n), i = 1, 2, ...n + m$. Then the data consists of the triplets (X_i, δ_i, Z_i), for $i = 1, 2.....n+m$, where $X_i = min(T_i, C_i)$ and $\delta_i = I(T_i \leq C_i)$. Assume that $\{C_i\}$ are independent and also independent of $\{T_i\}$. Then the crossing model can be expressed as

$$h_i(t) = \frac{1}{\gamma_{1i} + \gamma_{2i}(\beta)R(t)} \frac{dR(t)}{dt}, \tag{4.3.3}$$

where $\beta = (\beta_1, \beta_2)$ and $\gamma_{ji} = exp(-\beta_j Z_i)$, for $j = 1, 2$ and $i = 1.....n + m$.

2) $R(t)$ car be rewritten as

$$R(t) = \int_0^t R(s) \sum_{i=1}^{m+n} a_i \gamma_{2i}(\beta) h_i(s) ds + \int_0^t \sum_{i=1}^{m+n} a_i \gamma_{1i}(\beta) h_i(s) ds, \tag{4.3.4}$$

where $\sum_i a_i = 1$ for any $a_i \geq 0$. Details are shown in Yang and Prentice (1999). Since $h_i(t) = \frac{f_i(t)}{S_i(t)}, h_i(t)$ can be estimated by the empirical method in (4.3.4). Assuming the existence of the solution in (4.3.4), denoted by $\widehat{R}(t, \beta)$. Let $L(\beta, R)$ be the likelihood function of β for the model (4.3.3) with the corresponding score vector $S(\beta, R) = \partial \ln L(\beta, R)/\partial\beta$. Then the pseudo maximum likelihood estimator $\widehat{\beta}$ is defined as the root of $Q(\beta) = S(\beta, R)|_{R(t)=\widehat{R}(t,\beta)} = 0$. Plugging the corresponding estimators of θ into (4.2.2), we obtain an estimator of AUC.

Suppose we have samples from the two groups, and the sample sizes from treatment population and control population are m and n, respectively. We require the following conditions:

(C1). $\lim_{m,n\to\infty} \frac{n}{m+n} = p \in (0, 1)$.

(C2). Survival function S_0 is nonincreasing and has a bounded derivative for $t \in (0, \tau)$, $\tau < \tau_0$.

(C3). The cumulative distribution function G_i of C_i has a bounded derivative for $t \in (0, \tau)$,

and

$$\frac{1}{n}\sum_{i\le n} G_i(t) \to \Gamma_1, \quad \frac{1}{m}\sum_{i>n} G_i(t) \to \Gamma_2$$

uniformly for $t \le \tau$ for two functions Γ_1 and Γ_2 that satisfy $\Gamma_j(\tau) < 1, (j = 1, 2)$.

Then we have the following asymptotic normality of the estimators $\widehat{\beta}_i, i = 1, 2$, where $\widehat{\beta}_i = \ln \widehat{\theta}_i, i = 1, 2$.

Theorem 4.3.1. *Under the conditions of (C1)-(C3), we have that*

$$(m+n)^{1/2}(\widehat{\beta} - \beta) \to N\{0, U(\beta)\}, \tag{4.3.5}$$

where $U(\beta) = A(\beta)^{-1}V(\beta)\{A(\beta)\}^{-1}$, *with* $A(\beta)$ *is the limit in probability of* $-(m+n)^{-1}\partial Q(\beta)/\partial\beta$ *and* $V(\beta)$ *is from* $(A8)$ *in Yang and Prentice (2005).*

The detailed proof is shown in Yang and Prentice (2005, p13-16).

Now we can find the asymptotic normality of $\widehat{AUC}$ under the crossing model.

Theorem 4.3.2. *Suppose that the samples from two groups and their survival functions satisfy the conditions of (C1)-(C3). Then we have:*

$$(m+n)^{1/2}(\widehat{AUC} - AUC) \to N(0, V),$$

where $V = (\frac{\partial AUC}{\partial \theta_1})^2 Var(\widehat{\theta_1}) + 2\frac{\partial^2 AUC}{\partial\theta_1\partial\theta_2} Cov(\widehat{\theta_1}, \widehat{\theta_2}) + (\frac{\partial AUC}{\partial\theta_2})^2 Var(\widehat{\theta_2})$, AUC *as a function of* θ_1, θ_2 *in (4.2.2) and*

$$\begin{aligned}
\frac{\partial AUC}{\partial\theta_1} &= \sum_{i=1}^{\infty}(\frac{\theta_2-\theta_1}{\theta_2})^{i-1}\left(\frac{i\theta_1-\theta_2}{(i+\theta_2)(\theta_2-\theta_1)}\right), \\
\frac{\partial^2 AUC}{\partial\theta_1\partial\theta_2} &= \sum_{i=1}^{\infty}(\frac{\theta_2-\theta_1}{\theta_2})^{i-1}\left(\frac{i^2(i-1)\theta_1^2 + i^2\theta_1^2\theta_2 + (1-3i)\theta_1\theta_2^2 + 2i(1-i)\theta_1\theta_2 + \theta_2^3}{\theta_2(\theta_2-\theta_1)^2(i+\theta_2)^2}\right), \\
\frac{\partial AUC}{\partial\theta_2} &= \frac{\theta_1}{\theta_2(\theta_2-\theta_1)}\sum_{i=1}^{\infty}(\frac{\theta_2-\theta_1}{\theta_2})^{i-1}\frac{i\theta_1(1-i-\theta_2)+\theta_2^2}{(i+\theta_2)^2}.
\end{aligned}$$

$Var(\widehat{\theta_1})$, $Var(\widehat{\theta_2})$ and $Cov(\widehat{\theta_1}, \widehat{\theta_2})$ can be found from (4.3.5).

The proof follows from Theorem 4.3.1.

As a popular semi-parametric model, the bi-normal model assumes that the marker values are normally distributed under some monotonic increasing transformation. The crossing model provides another approach to model the ROC curve. Since these two methods do not overlap, one can consider the crossing model and bi-normal model to be complementary.

Compared with the Lehmann family, the crossing model involves two scale parameters which have good interpretation. It does not have available statistical packages to do parameter estimation, inference and model diagnostics and its parameter estimation involves a complicated process.

CHAPTER 5
DISCUSSION

In this dissertation, we discuss the estimation of AUC under various bi-models. Many literatures discussed relatively simple models: bi-Exponential, bi-Burr X, bi-Gamma, bi-Generalized Exponential, bi-Weibull and so on. They only have two parameters in a single model (a scale and a shape parameter) and derived score equations can be reduced to solve an equation on a parameter. Other two parameters can be solved analytically in most cases. Therefore, the computation process can be simplified largely. My dissertation studies the bi-Exponentiated Weibull distribution which has two shape parameters and one scale parameter. Score equations are much more complicated than those of other bi-models. We obtain a general expression of AUC under the bi-EW. It is an extension of the ones under both bi-WE and bi-GE. Our simulations show that the MLE of AUC can behave very well under moderate sample sizes although the performances of the MLE of some shape parameters are somewhat unsatisfactory but still acceptable. We may further consider the Bayesian method to estimate the parameters and AUC, and then compare with the MLE method. Since the Fisher information matrix of the parameters under bi-EW is computationally complicated, we may consider parametric bootstrap methods (bootstrap p and bootstrap t) to construct the confidence intervals.

However, all these studies are only based on some specific parametric survival models. For further study, we will propose a semi-parametric family to include all popular survival models (GE, WE, EW, Gamma and so on) as submodels. Under the

new semi-parametric model, we hope to derive estimators of the parameters, AUC and the corresponding asymptotic distributions. A possible approach is to use a mixture of two models.

We have noticed that under bi-WE, bi-GE and bi-EW models, AUC is always independent of some common parameter. We are interested in finding a large class of bi-model which also has this property.

In ROC analysis, a promising development is to consider the use of multiple diagnostic tests since most single diagnostic tests are not perfect. Combination of several diagnostic tests may provide a better diagnostic tool than any single test. For instance, in clinical practice, multiple sources of information are often available to a clinician. Sign and symptoms of disease, family and personal medical histories, in addition to results of formal medical tests, can be used to decide on a diagnosis.

The second part of my dissertation studies the AUC under the crossing model. It is a very effective approach. The crossing model is proposed to accommodate the case when two survival curves cross while AUC is a summary index to check an accuracy of a biomarker in medical diagnosis. My research builds up a bridge between two different fields of survival analysis and medical diagnosis. Further work is required to check the accuracy of the estimated AUC under the moderate sample sizes and different censoring rates through simulation studies.

BIBLIOGRAPHY

[1] Ahmad, K. E., Fakhry, M. E. and Jaheen, Z. F. Empirical Bayes estimation of $P(Y < X)$ and characterizations of the Burr-type X model, *Journal of Statistical Planning and Inference*, 64 (1997) 297-308.

[2] Ashour, S. K. and Afify, W. M. Estimation of the parameters of exponentiated Weibull family with type II progressive interval censoring with random removals. *Applied science research*, 4 (2008) 11, 1428-1442.

[3] Awad, A. M. and Hamadan, M. A. Some inference results in $P(X > Y)$ in the bivariate exponential model. *Communications in Statistics-Theory and Methods*, 10 (1981) 2515-2524.

[4] Barlow, R. E. and Proschan, F. *Statistical Theory of Reliability and Life Testing: Probability Models*, New York: Holt, Reinhardt and Winston (1975).

[5] Bennett, S. Analysis of survival data by the proportional odds model. *Statistics in Medicine*, 2 (1983) 273-277.

[6] Cai, T. X., Chaya, S. and Skowitz, M. Semiparametric estimation of the binormal ROC curve for a continuous diagnostic test. *Biostatistica*, 5 (2004) 573-586.

[7] Campbell, G. Advances in statistical methodology for evaluation of diagostic and laboratory tests. *Statistics in Medicine*, 13 (1994) 499-508.

[8] Carrasco, J. M. F., Ortega, E. M. M. and Corderio, G. M. A generalized modified Weibull distribution for lifetime modeling. *Computational Statistics and Data Analysis*, 53 (2008) 2, 450-462.

[9] Choudhury, A. A Simple Derivation of Moments of the Exponentiated Weibull Distribution. *Metrika* 62 (2005) 1, 1722.

[10] Constantine, K. and Karson, M. The Estimation of $P(X > Y)$ in Gamma case. *Communication in Statistics-Computations and Simulations*, 15 (1986) 65-388.

[11] Dielman, T. A comparison of quantile estimators. *Communications in Statistics - Simulation and Computation*, 23 (1994) 2, 355-371.

[12] Delong, E. R. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics*. 44 (1988) 837-845.

[13] Faraggi, D., Reiser, B. and Schisterman, E. F. ROC curve analysis for biomarkers based on pooled assessments. *Statistics in Medicine*, 22 (2003) 2515-2527.

[14] Fushing, H. and Turnbull, B. W. Nonparametric and semi-parametric Estimation of the Receiver Operating characteristic curve. *The Annuals of Statistics*, 24 (1996) l, 25-40.

[15] Gail, M. H. and Green, S. B. A generalization of the one-sided two sample Kolmogorov-Smirnov statistics for evaluating diagostic tests. *Biometrics*, 32 (1976) 561-570.

[16] Ghosh, D. Semi-parametric models and estimation procedures for binormal ROC curves with multiple biomarkers. The University of Michigan Dept. of Biostatistics Working Paper Series. paper 39 (2004).

[17] Gonen, M. and Heller, G. Lehmann Family of ROC curves. Memorial SloanKettering Cancer Department of Epodemiology and Biostatistics Working Paper Series (2007).

[18] Graver, D. P. and Acar, M. Analytical hazard representations for use in reliability, mortality and simulation studies. *Communications in Statistics*, Part B-Simulation and Computation, 8 (1979) 91-111.

[19] Green, D. M and Swets, J. A. *Singal detection theory and psychophysics*. Wiley & Sons, Inc. New York (1966).

[20] Gu, J., Ghosal, S. and Roy, A. Non-parametric estimation of ROC curve. *Institute of Statistics Mimeo Series*, (2005) # 2592.

[21] Hanley, J. A. and McNeil, B. J. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, (1982) 143, 29-36.

[22] Hanley, J. A. Receiver operating characteristic (ROC) methdology: the state of the art. *Critical reviews diagnostics imaging*, 29 (1989) 307-335.

[23] Harrell, F. E. and Davis, C. E. A new distribution-free quantile estimator. *Biometrika*, 69 (1982) 143, 635-640.

[24] Hjorth, U. A reliability distribution with increasing, decreasing, constant and bathtub shaped failure rates, *Technometrics*, 22 (1980) 99-107.

[25] Jiang, R. and Murthy, D. N. P. The exponentiated Weibull family: a graphical approach. *IEEE Transaction Reliability*, 48 (1999) 68-72.

[26] Khedhairi, A. L., Sarhan, A. and Tadj, L. Estimation of the generalized Rayleigh distribution parameters, King Saud University (2007).

[27] Kim, C., Jung, J. and Chung, Y. Bayesian estimation for the exponentiated Weibull model type II progressive censoring. *Statistical Papers*, 52 (2009) 1, 53-70.

[28] Kotz, S., Lumelskii, Y. and Pensky, M. *The stress-Strength Model and its Generalizations* (2002).

[29] Kundu, D. and Gupta, R. D. Estimation of $P(X > Y)$ For Generalized Exponential Distribution. *Metrika*, 34 (2005) 201-226.

[30] Kundu, D. and Gupta, R. D. Estimation of $P(X > Y)$ For Weibull Distribution. *IEEE Transactions on Reliability*, 34 (2006) 201-226.

[31] Leger, C. and Altman, N. Bandwidth selection for kernel distribution function estimation. *Journal of Statistical Planning and Inference*, 46 (1995) 2, 195-214.

[32] Lee, W. C. Probabilistic analysis of global performances of diagnostic tests: Interpreting the Lorenz curve-based summary measures. *Statistics in Medicine*, 18 (1999) 455-471.

[33] Liu, A. and Zhu, Y. On linear combinations of biomarkers to improve diagnostic accuracy. *Statistics in medicine*, (2005) 24, 37-47.

[34] Lloyd, C. J. Using smoothed receiver operating characteristic curves to summarize and compare diagnostic systems. *Journal of the American Statistical Association* 16 (1998) 93: 1356-1364.

[35] Lloyd, C. J. and Yong, Z. Kernel estimators of the ROC curve are better than empirical. *statistics and probability Leter*, (1999) 44 , 221-228.

[36] Metz, C. E., Herman, B. A. and Shen, J. H. Maximum likelihood estimation of receiver operating characteristic curve from continuous distributed data. *Statistics in Medicine*, (1998) 17, 1033-1-53.

[37] Mudholkar, G. S. and Hutson, A. D. The exponentiated Weibull family: some properties and a flood data application. *Communications Statistics- Theory and Methods*, 25 (1996) 12, 3059-3083.

[38] Mudholkar, G. S., Sivastava, D. K. and Freimer, M. The Exponentiated Weibull Family: A reanalysis of the Bus-Motor-Failure Data, *Technometrics*, 137 (1995) 4, 426-445.

[39] Nadarajah, S. and Gupta, A. K. On the moments of the exponentiated weibull distribution. *Commun Stat Theory Meth*, 34 (2005) 253-256.

[40] Owen, D. B., Craswell, K. J. and Hanson, D. L. Non-parametric upper confidence bounds for $P(X > Y)$ and confidence limits for $P(X > Y)$ when X and Y are normal. *Journal of American Statistical Association*, 59 (1997) 906-924.

[41] Pepe, M. S. *The Statistical Evaluation of Medical Tests for Classification and Prediction*, Oxford University Press (2003).

[42] Pepe, M. S. and Cai, T. X. The analysis of placement values for evaluating discriminatory measures. *Biometrics*, (2004) 60, 528-535.

[43] Prentice, R. L. Discrimination among some parametric models. *Biometrica*, (1975) 62, 607-614.

[44] Qian, L. F. The Fisher information matrix for a three-parameter exponentiated Weibull distribution under type II censoring. *Statistical Methodology*, 9 (2012) 3, 320-329.

[45] Qiu, P. H. and Le, C. ROC curve estimation based on local smoothing. *Journal of Statistical Computation and Simulation*, 70 (2001) 1, 55-69.

[46] Raqab, M. Z. and Kundu, D. Comparison of Different Estimation of $P(X > Y)$ for a Scaled Burr Type X Distribution. *Communications in Statistica-Simulation and Computation*, 22 (2005) 122-150.

[47] Simpson, A. J. and Fitter, M. J. What is the best index of detectability. *Psychological Bulletin*, 80 (1973) 481-488.

[48] Singh, U., Gupta, P. K. and Upadlhaya, S. K. Some point estimates for the shape parameters of exponentiated Weibull family. *Korean Statistical Society*, 35 (2006) 63-77.

[49] Slymen, D. J. and Lachenbruch, P. A. Survival distributions arising from two families and generated by transformations. *Communications in Statistics*, Part A-Theory and Methods, (1984) 13, 1179-1201.

[50] Stacy, E. W. A generalization of the Gamma distribution, *The annals of Statistics*. (1962) 33, 1187-1192.

[51] Su, J. Q. and Liu, J. S. Linear combination of multiple diagnostic markers. *Journal of the American Statistical Association*, 88 (2003) 424, 1350-1355.

[52] Surles, J. G. and Padgett, W. J. Inference for $P(X > Y)$ in the Burr type X model. *Journal of Applied Statistical Sciences*, 7 (1998) 225-238.

[53] Surles, J. G. and Padgett, W. J. Inference for reliability and stress-strength for a scaled Burr-type X distribution. *Lifetime Data Analysis*, 7 (2001) 187-200.

[54] Yang, S. and Prentice, R. Semiparametric analysis of short-term and long-term hazard ratios with two-sample survival data. *Biometrika*, 92 (2005) 1, 1-17.

[55] Wahed, A. S. A general method of constructing extended families of distribution from an existing continuous class. *Journal of Probability and Statistics Science*. (2006) 4, 165-177.

[56] Wahed, A. S., Luong, T. M. and Jeong, J. H. A new generalization of Weibull distribution with application to a breast cancer data set. *Statistics in Medicine*, (2009) 28, 2077-2094.

[57] Wand, M. P. and Jones, M. C. *Kernel Smoothing*, (1995).

[58] Woodward, W. A. and Kelley, G. D. Minimum variance unbiased estimation of $P(Y < X)$ in the normal case, *Technometrics*, 19 (1997) 95-98.

[59] Zhou, X. H. and Lin, H. Z. Semi-parametric likelihood estimates for ROC curves of continuous-scale tests. *Statistics in Medicine*, (2008) 27, 5271-5290.

[60] Zou, K. H., Tempany, C. M., Fielding, J. R. and Silverman, S. G. Original smooth receiver operating characteristic curves estimation from continuous data: statistical methods for analyzing the predictive value of spiral CT of ureteral stones. *Academic Radiology*, (1998) 5, 680-687.

[61] Zou, K. H. and Hall, W. J. two transformation models for estimating an ROC curve derived from continuous data. *Journal of Applied Statistics*, (2000) 5, 621-631.