Semi-Parametric Inference for the Partial Area Under the ROC Curve

Fangfang Sun

Follow this and additional works at: https://scholarworks.gsu.edu/math_theses

Part of the Mathematics Commons

Recommended Citation
https://scholarworks.gsu.edu/math_theses/61

This Thesis is brought to you for free and open access by the Department of Mathematics and Statistics at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Mathematics Theses by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.
ABSTRACT

Diagnostic tests are central in the field of modern medicine. One of the main factors for interpreting a diagnostic test is the discriminatory accuracy. For a continuous-scale diagnostic test, the area under the receiver operating characteristic (ROC) curve, AUC, is a useful one-number summary index for the diagnostic accuracy of the test. When only a particular region of the ROC curve would be of interest, the partial AUC (pAUC) is a more appropriate index for the diagnostic accuracy. In this thesis, we develop seven confidence intervals for the pAUC under the semi-parametric models for the diseased and non-diseased populations by using the normal approximation, bootstrap and empirical likelihood methods. In addition, we conduct simulation studies to compare the finite sample performance of the proposed confidence intervals for the pAUC. A real example is also used to illustrate the application of the recommended intervals.

INDEX WORDS: ROC, AUC, The partial AUC, Diagnostic test, Confidence interval
SEMI-PARAMETRIC INFERENCE FOR THE PARTIAL AREA
UNDER THE ROC CURVE

by

FANGFANG SUN

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science
In the College of Arts and Sciences
Georgia State University

2008
SEMI-PARAMETRIC INFERENCE FOR THE PARTIAL AREA
UNDER THE ROC CURVE

by

FANGFANG SUN

Major Professor: Dr. Gengsheng Qin
Committee: Dr. Yu-Sheng Hsu
Dr. Yixin Fang
Dr. Yuanhui Xiao

Electronic Version Approval:
Office of Graduate Studies
College of Arts and Sciences
Georgia State University
December 2008
ACKNOWLEDGEMENTS

I would like to acknowledge those who have helped me in the completion of this thesis. First of all, I like to thank my advisor, Dr. Gengsheng Qin, for his patience and guidance. I would also like to acknowledge the other members of my thesis committee, Dr. Yu-Sheng Hsu, Dr. Yixin Fang and Dr. Yuanhui Xiao for taking the time to read this thesis and provide useful comments. Lastly, I want to acknowledge the support of my family and friends.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS........................................................................................................iv

LIST OF TABLES......................................................................................................................vi

CHAPTERS

1. INTRODUCTION................................................................................................................1

2. NORMAL APPROXIMATION AND BOOTSTRAP BASED CONFIDENCE INTERVALS FOR THE PARTIAL AUC.................................................................6
   2.1 Normal approximation based confidence interval .........................................................7
   2.2 Bootstrap based confidence intervals...........................................................................9

3. EMPIRICAL LIKELIHOOD BASED SEMI-PARAMETRIC INFERENCE FOR THE PARTIAL AUC...........................................................................................11

4. SIMULATION STUDY.........................................................................................................16
   4.1 Normal distributions estimators................................................................................16
   4.2 Exponential distributions estimators.........................................................................17

5. DUCHENNE MUSCULAR DYSTROPHY EXAMPLES......................................................18

6. CONCLUSION AND DISCUSSION.................................................................................21

REFERENCES.......................................................................................................................22

APPENDICES.........................................................................................................................24

   APPENDIX A: SIMULATION TABLES..............................................................................24
   APPENDIX B: REAL DATA ANALYSIS TABLES...............................................................29
   APPENDIX C: THE S-PLUS CODE FOR SIMULATION STUDIES.................................32
   APPENDIX D: THE S-PLUS CODE FOR REAL DATA ANALYSIS.................................38
LIST OF TABLES

Table I: Normal distribution: level of 95 percent confidence interval for the AUC………………25
Table II: Normal distribution: level of 90 percent confidence interval for the pAUC……………26
Table III: Exponential distribution: level of 95 percent confidence interval for the pAUC …..27
Table IV: Exponential distribution: level of 90 percent confidence interval for the pAUC……28
Table V: Duchenne Muscular Dystrophy Example: Level of 95 percent confidence intervals for the normalized pAUC…………………………………………………………………….30
Table VI: Duchenne Muscular Dystrophy Example: Level of 90 percent confidence intervals for the normalized pAUC…………………………………………………………………….31
CHAPTER 1
INTRODUCTION

Diagnostic tests are central by helping physicians revise disease probability from their patients in the field of modern medicine. One of the main factors for interpreting a diagnostic test is the discriminatory accuracy which is the ability of the diagnostic test to distinguish between two groups of subjects, usually non-diseased and diseased subjects (Shapiro, 1999). As a result, statistical studies for assessing discriminatory accuracy have been recently received much attention. As Shapiro mentioned in his paper, most of diagnostic tests are continuous. Therefore, the discriminatory accuracy of continuous-scale tests will only be discussed in this thesis.

The accuracy of a binary diagnostic test can be measured by its specificity and sensitivity. The sensitivity or true positive rate (TPR) of the test is the proportion that the test is positive among diseased patients. The specificity or true negative rate (TNR) of the test is the proportion that the test will be negative among non-diseased patients. Let Y and X be the results of a continuous-scale test for a diseased subject and a non-diseased subject with cumulative distribution functions $G$ and $F$, respectively. For a given cut-off point $c$, the sensitivity and specificity of the test are defined as

$$Se = P(Y \geq c | D) = 1 - G(c), \quad Sp = P(X \leq c | \overline{D}) = F(c),$$

respectively, where $D$ denotes a diseased state and $\overline{D}$ denotes a non-diseased state.

Equivalently, the false positive rate (FPR) and false negative rate (FNR) of the test are defined as $(1$-specificity) $= 1 - F(c)$ and $(1$-sensitivity) $= G(c)$, respectively. In biological, medical, and health service research, the receiver operating characteristic (ROC) curve, which is a plot of sensitivity against FPR for all possible cut-off points $c$, is a useful
graphical summary of the diagnosis accuracy of a test. The ROC curve can mathematically be represented by \( R(p) = 1 - G(F^{-1}(1 - p)) \), where \( F^{-1} \) is the inverse function of \( F \). \( R(p) \) is indeed the sensitivity of the test when the specificity is fixed at level \((1 - p)\). The ROC curve was derived from statistical decision theory and originally developed in the literature of electronic signal detection (Hanley, 1989). As Shapiro (1999) reviewed in his paper, many approaches have been proposed to make inference about the ROC curve, such as, parametric approaches proposed by Strike (1995), Goddard et al. (1990), Egan (1975) and England WL (1988); a semi-parametric algorithm, LABROC4, proposed by Metz et al. (1998); and non-parametric approaches proposed by Zou et al. (1997), Le (1997) and Lloyd (1998). The ROC curve best summarizes the diagnostic accuracy by plotting sensitivity versus FPR for all possible cut-off point \( c \). However, since a one-number summary index for the diagnostic accuracy is often desired, the area under the ROC curve (AUC), defined as \( \delta = \int_0^1 R(p)dp \), becomes a popular summary of the accuracy of a diagnostic test across all the possible cut-off points. The larger is the AUC, the better performance the test will have. In other word, the value of \( \delta \) closer to 1 indicates the higher diagnostic accuracy of the test.

The AUC is a global summary of the diagnostic accuracy, but it has some limitations (Shapiro, 1999; Hilden 1991). For instance, when two ROC curves cross, the two tests can have similar AUC even though one test has higher sensitivity for certain specificities while the other test has better sensitivity for other specificities; the AUC includes regions of ROC space that would not be of practical interest (e.g. very high FPR, or very low TPR) (Shapiro, 1999). As a result, in case that it is critical to maintain a
particular range of specificity or sensitivities in some diagnostic testing, such as a low false positive rate or a high sensitivity, the AUC may not be an appropriate accuracy index for diagnostic tests because only the region of ROC curve corresponding to acceptable low false positive rates or high sensitivities would be of interest. For example, when the FPR must be very small to be acceptable for cancer screening (Lilienfeld, 1974), interest would only be in the lower tail of ROC curve. Alternatively, the partial AUC (pAUC) was proposed for some diagnostic tests in which only a particular region of the ROC curve would be of interest. The pAUC over the interval of false positive rates \((p_0, p_1)\), denoted by \(\delta_{p_0,p_1}\), is
\[
\delta_{p_0,p_1} = \int_{p_0}^{p_1} R(p) dp , \text{ for } 0 \leq p_0 < p_1 \leq 1.
\]
It can be described as the cumulative value of sensitivity for all possible values of the false positive rates in the interval \((p_0, p_1)\).

In recent years, many parametric and nonparametric approaches have been proposed in making inference about the partial AUC. Parametric methods based on the bi-normal model were proposed by McClish (1989), Thompson and Zucchini (1989), and Jiang, Metz, and Nishikawa (1996). However, a major concern of these parametric methods is that they are quite sensitive to model departures from the distributional assumptions (Walsh, 1997). A generalized nonparametric method was proposed by Wieand et al (1989) for the inference of both the full and the partial AUC. But the method is mathematically too complicated to be well applied in practice since it is involved in density and distribution function estimations (Qin et al. 2006). Another nonparametric method for the pAUC was proposed by Zhang et al. (2002). However, this
method is valid only for ordinal-scale data. Let $X_1, X_2, \ldots, X_m$ be the test results from a random sample of non-diseased population with distribution function $F$; let $Y_1, Y_2, \ldots, Y_n$ be the test results from a random sample of diseased population with distribution function $G$. Dodd and Pepe (2003) proposed the following consistent nonparametric estimator for the pAUC:

$$
\hat{\delta}_{p,p_i} = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} I(Y_j \geq X_i)I(X_i \in (\hat{q}_i, \hat{q}_0))
$$

where $\hat{q}_l = \hat{F}^{-1}(1 - p_l)$ ($l=0,1$) is the $(1 - p_l)$-th sample quantile, and $\hat{F}$ is the empirical distribution of $F$. But they did not provide the asymptotic variance of the estimator. Instead, they recommended using the bootstrap to obtain the variance estimate. Qin and Zhou (2006) proposed an Empirical Likelihood (EL) based approach to derive confidence intervals for the full AUC. Qin, Jin and Zhou (2007) derived the asymptotic normal distribution for the estimator $\hat{\delta}_{p,p_i}$ and developed EL-based nonparametric confidence intervals for the partial AUC.

Semi-parametric methods are frequently used in statistical literature because they share the advantages of both parametric and nonparametric methods. They are more flexible and robust than the traditional parametric methods. On the other hand they may be more efficient than the pure non-parametric methods and usually have better finite sample performances than their non-parametric counterparts. Semi-parametric methods have been used for making inference of ROC curves. For instance, Li et al. (1999) proposed a non-parametric approach to estimate the distribution of test results in non-diseased subjects, whereas assuming a parametric model for the distribution of test results in diseased subjects. Dodd and Pepe (2003) proposed a semi-parametric regression model
for evaluating covariate effects on ROC curves. Qin and Zhang (2003) developed a semi-parametric approach by assuming a density ratio model for disease and disease-free densities. In this thesis, we make a general semi-parametric model assumption for the test results $X$ and $Y$. We shall propose a semi-parametric estimator for the pAUC and derive the asymptotic normal distribution of the estimator. Seven confidence intervals for the partial AUC based on the semi-parametric models are developed by using the normal approximation, bootstrap and empirical likelihood methods.

The thesis is organized as follows: In Chapter 2, we propose a normal approximation based confidence interval and two bootstrap-based confidence intervals for the partial AUC. In Chapter 3 we propose four EL-based intervals for the partial AUC. Simulation studies are conducted in Chapter 4 to evaluate the finite sample performances of these intervals. Chapter 5 uses a real data set to illustrate the application of the proposed intervals. Finally, the conclusions are discussed in Chapter 6.
CHAPTER 2
NORMAL APPROXIMATION AND BOOTSTRAP BASED CONFIDENCE INTERVALS FOR THE PARTIAL AUC

Consider one diagnostic test which yields continuous measurements and is performed on \( m \) non-diseased subjects and \( n \) diseased cases. Let \( X_1, X_2, \ldots, X_m \) be the test results of a random sample of non-diseased subjects with an unknown cumulative distribution function \( F \), and \( Y_1, Y_2, \ldots, Y_n \) the test results of a random sample of diseased subjects with a parametric distribution function \( G(y; \theta) \), where \( \theta \in \Theta \) is an unknown parameter for some set \( \Theta \subset R^q, q \geq 1 \). In this setting, we can rewrite the partial AUC as

\[
\delta_{p_0p_1} = \int_{p_0}^{p_1} R(p)dp
= P\{Y > X, X \in (q_k, q_0)\;|\;X \in (q_1, q_0))\}
= E[P(Y > X; \theta \mid X \in (q_1, q_0))]
= E[V(X; \theta)]
\]

for \( 0 \leq p_0 < p_1 \leq 1 \), where \( R(p) = 1 - G(F^{-1}(1 - p); \theta) \) is the ROC curve of the test, and \( q_k = F^{-1}(1 - p_k), k = 0,1 \) are the \( (1 - p_k) \)-th quantiles of \( F \), and

\[
V(X; \theta) = P(Y > X; \theta \mid X \in (q_1, q_0))
= (1 - G(X; \theta))I(X \in (q_1, q_0))
\]

We propose the following estimator for the partial AUC:

\[
\delta_{p_0p_1}^\hat{=} = \frac{1}{m} \sum_{i=1}^{m} \hat{V}_i(X_i; \theta),
\]
where $\hat{V}_i(X_i; \theta) = (1-G(X_i; \theta))I(X_i \in (q_i^+, q_i))$, $i = 1, 2, \ldots, m$, $\hat{\theta}$ is the maximum likelihood estimator of $\theta$ based on $Y_j$'s, $j = 1, 2, \ldots, n$, and $\hat{q}_k = \hat{F}^{-1}(1-p_k)$ is the $(1-p_k)$-th sample quantiles of $X_i$'s, $i = 1, 2, \ldots, m$, and $\hat{F}$ is the empirical distribution of $X_i$'s.

In this section, our goal is to find the asymptotic distribution of the estimator and construct confidence intervals for $\delta_{p_0p_1}$ based on test results $X_i$'s and $Y_i$'s.

### 2.1 Normal approximation based confidence interval

We can show that the estimator $\hat{\delta}_{p_0p_1}$ asymptotically follows a normal distribution in the following theorem.

**Theorem 2.1:** Assume that $0 < \lim_{m,n \to \infty} \frac{m}{n} = \rho < \infty$ is a constant. Then

$$m^{1/2}(\hat{\delta}_{p_0p_1} - \delta_{p_0p_1}) \xrightarrow{L} N(0, \sigma^2_{p_0p_1}),$$

(2.1)

where

$$\sigma^2_{p_0p_1} = \text{Var}(B(X; \theta, q_0, q_1)) + \rho D'(\theta, q_0, q_1)\Sigma_{\theta}D(\theta, q_0, q_1),$$

$$D(\theta, q_0, q_1) = E[G_{\theta}(X; \theta)I(X \in (q_i, q_0))],$$

$$B(X; \theta, q_0, q_1) = [(1-G(X; \theta))I(X \in (q_i, q_0)) - \delta_{p_0p_1}]$$

$$- \sum_{k=0}^{1} [1-G(q_k; \theta)][I(X \leq q_k) - (1-p_k)],$$

and $\Sigma_{\theta}$ is the asymptotic covariance matrix of $\sqrt{n}(\hat{\theta} - \theta)$, $G_{\theta}(X; \theta)$ denotes the derivative of $G(X; \theta)$ with respect to $\theta \in \Theta \subseteq \mathbb{R}^d$, $D'(\theta, q_0, q_1)$ is the transpose of $D(\theta, q_0, q_1)$. 
The asymptotic normality in Theorem 2.1 cannot be directly used to produce a confidence interval for the partial AUC $\delta_{p\theta}$ because the asymptotic variance $\sigma^2_{p\theta}$ is a function of the unknown parameter $\theta$ and the distribution function $F$. However, we can empirically estimate the asymptotic variance $\sigma^2_{p\theta}$ by using the maximum likelihood estimator $\hat{\theta}$ and the empirical distribution $\hat{F}$. The following $\hat{\sigma}^2_{p\theta}$ is a consistent estimate for $\sigma^2_{p\theta}$:

$$\hat{\sigma}^2_{p\theta} = \hat{\text{Var}}(B(X; \theta, q_0, q_1)) + \frac{n}{m} \hat{D}(\theta, q_0, q_1) \hat{\Sigma}_D(\theta, q_0, q_1),$$

where

$$\hat{\text{Var}}(B(X; \theta, q_0, q_1)) = \frac{1}{m-1} \sum_{i=1}^{m} (B(X_i; \hat{\theta}, \hat{q}_0, \hat{q}_1) - \frac{1}{m} \sum_{i=1}^{m} B(X_i; \hat{\theta}, \hat{q}_0, \hat{q}_1))^2,$$

$$B(X_i; \hat{\theta}, \hat{q}_0, \hat{q}_1) = [(1 - G(X_i, \hat{\theta}))I(X_i \in (\hat{q}_1, \hat{q}_0)) - \hat{\delta}_{p\theta}],$$

$$\hat{D}(\theta, q_0, q_1) = \frac{1}{m} \sum_{i=1}^{m} G'(X_i, \hat{\theta})I(X_i \in (\hat{q}_1, \hat{q}_0)),$$

and $\hat{\Sigma}_D$ is a consistent estimator of $\Sigma_D$.

Therefore, a $(1 - \alpha)$-th normal approximation (NA) based confidence interval for the partial AUC $\delta_{p\theta}$ can be constructed as follows:

$$(\hat{\delta}_{p\theta} - z_{1-\alpha/2} \frac{\hat{\sigma}_{p\theta}}{\sqrt{m}}, \hat{\delta}_{p\theta} + z_{1-\alpha/2} \frac{\hat{\sigma}_{p\theta}}{\sqrt{m}}),$$

where $z_{1-\alpha/2}$ is the $(1 - \alpha / 2)$-th quantile of the standard normal distribution.
2.2 Bootstrap based confidence intervals

Bootstrap is a useful approach for the construction of confidence intervals of unknown parameters. It is frequently used to estimate the variance of a statistic when the variance is unknown and of a complex form. Clearly, the asymptotic variance $\sigma_p^2$ of $\delta_{p,p}$ is complex and can be empirically estimated by using the MLE $\hat{\theta}$ and sample quantiles. However, this empirical variance estimate may not be stable when sample size is small. To produce a better confidence interval for the partial AUC, we propose the following bootstrap procedure to estimate the asymptotic variance.

Step 1. Draw a bootstrap resample \{ $X_1^*$, $X_2^*$, ..., $X_m^*$ \} of size $m$ from the non-diseased sample \{ $X_1$, $X_2$, ..., $X_m$ \} and a bootstrap resample \{ $Y_1^*$, $Y_2^*$, ..., $Y_n^*$ \} of size $n$ from the diseased sample \{ $Y_1$, $Y_2$, ..., $Y_n$ \}.

Step 2. Calculate a bootstrap version of $\hat{\delta}_{p,p}$:

$$\hat{\delta}_{p,p}^* = \frac{1}{m} \sum_{i=1}^{m} V(X_i^*, \hat{\theta}^*)$$

where

$$V(X_i^*, \hat{\theta}^*) = [1 - G(X_i^*, \theta^*)] I(X_i^* \in (\hat{q}_k^*, \hat{q}_0^*)),$$

$\hat{\theta}^*$ is the bootstrap version of $\hat{\theta}$, and $\hat{q}_k^* = \hat{F}^{-1}(1 - p_k)$ is the $(1 - p_k)$-th sample quantile based on the bootstrap resample \{ $X_1^*$, $X_2^*$, ..., $X_m^*$ \}.

Step 3. Repeat the steps 1-2 $B$ ($B$ is recommended to be greater than or equal to 150) times to get $B$ bootstrap copies of $\hat{\delta}_{p,p}$.
Then, the bootstrap variance estimator is defined as:

$$\sigma^*_{p; p_b} = \frac{1}{B-1} \sum_{b=1}^{B} (\delta^*_{p; p_b} - \bar{\delta}^*_{p; p_b})^2,$$

where $$\bar{\delta}^*_{p; p_b} = \frac{1}{B} \sum_{b=1}^{B} \delta^*_{p; p_b; b}$$.

This leads to two $$(1-\alpha)$$-th bootstrap based confidence intervals for the partial AUC defined as follows:

**BI interval:**

$$\left( \delta_{p; p_b}^* - z \frac{\sigma_{p; p_b}^*}{\sqrt{m}}, \delta_{p; p_b}^* + z \frac{\sigma_{p; p_b}^*}{\sqrt{m}} \right) \quad (2.3)$$

**BII interval:**

$$\left( \bar{\delta}_{p; p_b}^* - z \frac{\sigma_{p; p_b}^*}{\sqrt{m}}, \bar{\delta}_{p; p_b}^* + z \frac{\sigma_{p; p_b}^*}{\sqrt{m}} \right) \quad (2.4)$$
CHAPTER 3  

Empirical Likelihood Based Confidence Intervals for the partial AUC

It is well known that the empirical likelihood (EL) (Owen, 1990, 2001) also is a popular non-parametric method used for providing confidence intervals of unknown parameter. Qin and Zhou (2006) proposed an empirical likelihood method for the statistical inference of the full AUC. In this chapter, we develop an EL-based semi-parametric inference for the partial AUC.

By assuming that the parameter $\theta$ and quantiles $q_1$ and $q_0$ are known, Dodd and Pepe (2003) defined the restricted placement value of $X$ as follows:

$$ V(X; \theta) = (1 - G(X; \theta))I(X \in (q_1, q_0)). $$

They interpreted $V(X; \theta)$ as the restricted placement value of a given non-diseased test value, $X$, in the survival function of test results of diseased subjects. As shown in Chapter 2,

$$ E[V(X; \theta)] = P\{Y > X, X \in (q_1, q_0); \theta\} = \delta_{p_0p_1}, $$

which implies that the expectation of $V(X; \theta)$ is equal to the partial AUC, $\delta_{p_0p_1}$.

By using this relationship between the restricted placement value and the partial AUC, an empirical likelihood for $\delta_{p_0p_1}$ can be defined by the following expression:

$$ \tilde{L}(\delta_{p_0p_1}) = \sup_{\mathbf{p}} \{ \prod_{i=1}^{m} p_i : \sum_{i=1}^{m} p_i (V(X_i; \theta) - \delta_{p_0p_1}) = 0 \}, \quad (3.1) $$

where $\mathbf{p}=(p_1, p_2, \ldots, p_m)$ is a probability vector, i.e. $\sum_{i=1}^{m} p_i = 1$ and $p_i \geq 0$ for all $i$.

Obviously, $\tilde{L}(\delta_{p_0p_1})$ cannot be calculated since it depends on the unknown parameter $\theta$. 


in the diseased distribution function $G$ and quantiles $q_i$’s of the unknown non-diseased
distribution function $F$. A natural solution is to replace $\theta$ by its MLE $\hat{\theta}$ and $q_i$’s by their
sample quantiles $\hat{q}_i$’s. Hence, an adjusted empirical likelihood for $\delta_{p_{0:1}}$ can be defined as
follows:

$$L(\delta_{p_{0:1}}) = \sup \left\{ \prod_{i=1}^{m} p_i : \sum_{i=1}^{m} p_i (\hat{V}(X_i, \hat{\theta}) - \delta_{p_{0:1}}) = 0 \right\},$$

where $\hat{V}(X_i, \hat{\theta}) = [1 - G(X_i, \theta)]I(X_i \in (\hat{q}_i, q_0)), i = 1, 2, \ldots, m$.

Then the corresponding empirical log-likelihood ratio (ELR) for $\delta_{p_{0:1}}$ is

$$l(\delta_{p_{0:1}}) = 2 \sum_{i=1}^{m} \log \left[ 1 + \lambda (\hat{V}(X_i, \hat{\theta}) - \delta_{p_{0:1}}) \right] ,$$

(3.2)

where $\lambda$ is the solution to

$$\frac{1}{m} \sum_{i=1}^{m} \frac{\hat{V}(X_i, \hat{\theta}) - \delta_{p_{0:1}}}{1 + \lambda (\hat{V}(X_i, \hat{\theta}) - \delta_{p_{0:1}})} = 0 .$$

(3.3)

We can prove that $l(\delta_{p_{0:1}})$ follow a scaled $\chi^2$ distribution as shown by the
following theorem:

**Theorem 3.1:** If $\delta_{p_{0:1}}$ is the true value of the partial AUC, and $\lim_{m,n \to \infty} \frac{m}{n} = \rho$ is a
constant, then the limiting distribution of $l(\delta_{p_{0:1}})$ is a scaled chi-square distribution with
one degree of freedom. i.e.,

$$C(\delta_{p_{0:1}}) l(\delta_{p_{0:1}}) \xrightarrow{L} \chi^2_1 ,$$

where the scale constant $C(\delta_{p_{0:1}}) = \frac{S_{p_{0:1}}^2}{\sigma_{p_{0:1}}^2}$ and $S_{p_{0:1}}^2 = Var(V(X; \theta))$. 

In order to construct empirical likelihood confidence intervals for the partial AUC, we need to estimate $S_{p_{p_1}}^2$ and $\sigma^2_{p_{p_1}}$ since the scale constant $C(\sigma_{p_{p_1}})$ is unknown.

Here, we will use the bootstrap variance estimate $\sigma^2_{p_{p_1}}$ defined in section (2.2). Then, using Theorem 3.1, four hybrid bootstrap and empirical likelihood (HBEL) based confidence intervals for the partial AUC can be constructed as follows:

**HBEL I interval:**

\[ \{ \delta : \hat{C}(\delta_{p_{p_1}})l(\delta) \leq \chi^2_1(1 - \alpha) \}, \quad (3.4) \]

where $\hat{C}(\delta_{p_{p_1}})$ is an estimator for $C(\delta_{p_{p_1}})$:

\[
\hat{C}(\delta_{p_{p_1}}) = \frac{\hat{S}_{p_{p_1}}^2}{m \sigma^2_{p_{p_1}}} \\
\hat{S}_{p_{p_1}}^2 = \frac{1}{m-1} \sum_{i=1}^{m} (\hat{U}(X_i; \theta) - \hat{\delta}_{p_{p_1}})^2
\]

**HBEL II interval:**

\[ \{ \delta : \hat{C}^*(\delta_{p_{p_1}})l(\delta) \leq \chi^2_1(1 - \alpha) \}, \quad (3.5) \]

where $\hat{C}^*(\delta_{p_{p_1}})$ is the following estimator for $C(\delta_{p_{p_1}})$:

\[
\hat{C}^*(\delta_{p_{p_1}}) = \frac{\bar{S}_{p_{p_1}}^2}{m \sigma^2_{p_{p_1}}},
\]

with $\bar{S}_{p_{p_1}}^2$ being the mean of $B$ bootstrap copies of $\hat{S}_{p_{p_1}}^2$.

**HBEL III interval:**

\[ \{ \delta : l(\delta) \leq C^* \}, \quad (3.6) \]
where $C^*$ is the $(1-\alpha)$-th quantile of $\{l_1^*(\hat{\delta}_{p,h}), l_2^*(\hat{\delta}_{p,h}), \ldots, l_B^*(\hat{\delta}_{p,h})\}$ which are the $B$ bootstrap copies of $l^*(\hat{\delta}_{p,h})$ with 

$$l^*(\hat{\delta}_{p,h}^{B}) = 2\sum_{i=1}^{m} \log[1 + \lambda^* (V(X_i^*; \theta^*) - \hat{\delta}_{p,h})],$$

and $\lambda^*$ being the solution to 

$$\frac{1}{m} \sum_{i=1}^{m} \frac{V(X_i^*, \hat{\theta}^{\lambda^*}) - \delta_{p,h}}{1 + \lambda^* (V(X_i^*, \hat{\theta}^{\lambda^*}) - \delta_{p,h})} = 0.$$

**HBELIV interval:**

$$\{\delta : \bar{C}^*(\hat{\delta}_{p,h}) l(\delta) \leq \chi^2_1(1-\alpha)\},$$

(3.7)

where $\bar{C}^*(\hat{\delta}_{p,h})$ is an estimator for $C(\hat{\delta}_{p,h})$:

$$\bar{C}^*(\hat{\delta}_{p,h}) = \left[\frac{1}{B} \sum_{b=1}^{B} l_b^*(\hat{\delta}_{p,h})\right]^{-1}.$$

To obtain the hybrid bootstrap and empirical likelihood based confidence intervals for the partial AUC, for example, the **HBELI** interval, we need to solve the following nonlinear equations for $\lambda$ and $\delta$:

$$\begin{align*}
\frac{1}{m} \sum_{i=1}^{m} \frac{\hat{V}(X_i; \hat{\theta}) - \delta}{1 + \lambda (\hat{V}(X_i; \hat{\theta}) - \hat{\delta})} &= 0 \\
\bar{C}^*(\hat{\delta}_{p,h}) l(\delta) &= \chi^2_1(1-\alpha)
\end{align*}$$
There will be two solutions for $\delta$, the smaller one and bigger one are the lower bound and the upper bound of the HBELI interval respectively. Use the similar procedure as above, we can find the HBELII, HBELIII and HBELIV intervals.
CHAPTER 4
SIMULATION STUDY

In this Chapter, we conduct two simulation studies to evaluate the small sample performances of the seven confidence intervals (NA, BI, BII, HBELI, HBELII, HBELIII, and HBELIV) presented in Chapter 2 and Chapter 3. Since we know the underlying parametric distributions in the simulation study, the maximum likelihood (ML) based confidence interval for the partial AUC is also included in the study as a comparison base.

In the simulation studies, binormal distributions and exponential distributions are chosen to be the underlying distributions respectively. For each study, 1000 random samples of size \( m \) from the non-diseased population \( F \) and of size \( n \) from the diseased population \( G \) have been generated respectively, where the sample sizes \((m, n) = (30, 30), (50, 50), (100, 100), (50, 30), \) and \((80, 50)\). The subinterval \((p_0, p_i)\) of \((0,1)\) for the partial AUC \(\delta_{p_0,p_i}\) is chosen to be \((0, 0.4), (0, 0.7), \) and \((0.05, 0.5)\), respectively. For the bootstrap method, we take \(B=150\) bootstrap re-samples from the original samples. Under these simulation settings, various confidence intervals with confidence levels at both 95% and 90% are computed for the partial AUC.

4.1 Normal distributions estimators

In the first study, the test results of non-diseased patients are generated from the standard normal distribution \(N(0,1)\), and the test results of diseased patients are generated
from the normal distribution $N(1,1)$. The simulation results are shown in Table I and Table II. From these tables, we observe that the NA intervals have the longest interval lengths and always over-cover the true partial AUC. Other six confidence intervals (BI, BII, HBELI, HBELII, HBELIII and HBELIV) based on bootstrap and empirical likelihood methods have similar coverage probabilities, but the HBELII and HBELIII intervals perform slightly better in most cases considered here.

4.2 Exponential distributions estimators

The following exponential models are considered in the second simulation study: a standard exponential distribution $\text{Exp}(\lambda)$ with rate $\lambda = 1$ for the non-diseased population and an exponential distribution $\text{Exp}(\gamma)$ with rate $\gamma = 0.43$ for the diseased population. Simulation results are presented in Table III and Table IV. From these two tables, we again observe that the NA intervals are the most conservative intervals for the partial AUC. Other six confidence intervals (BI, BII, HBELI, HBELII, HBELIII and HBELIV) based on bootstrap and empirical likelihood methods have similar coverage probabilities.

In summary, our simulation studies suggest that the bootstrap and EL-based methods perform better than the normal approximation (NA) based method due to the better coverage probability and shorter interval length. The parametric ML interval has good coverage accuracy, but it can only be used when the underlying parametric assumption is true.
CHAPTER 5

DUCHENNE MUSCULAR DYSTROPHY EXAMPLE

Duchenne muscular dystrophy (DMD) is one of nine types of muscular dystrophy and is characterized by rapidly-worsening muscle weakness that starts in the legs and pelvis, and later affects the whole body. It is inherited in what is known as an X-linked recessive pattern. In other word, DMD affects children who inherit the disease through their mothers. Unfortunately, women can be carriers of DMD but usually exhibit no symptoms. As a result, early screening of females who could be potential carriers is essential because there is no cure for DMD.

Serum enzyme tests can be important aids in the diagnosis of DMD since carriers lead to a significant increase in the muscle protein levels found in the blood. In this study, we shall consider the serum pyruvate kinase (PK) level measured in known carriers and in healthy females from serum enzyme test and want to know how accurate the test can be in detecting DMD.

The study data set, which is collected during a program run at the Hospital for Sick Children of Toronto, is given in Andrews and Herzberg (1985). Totally, 127 healthy females (non-carriers) and 67 carriers are included in the study. As Adimari and Chiogna (2005) discussed in their paper, transformations are needed for the original data set to assume normal distributions for the transformed values since there are variations in normality for both models (carriers and healthy females). The transformation with the power of -0.56 is suggested by the Box-Cox method for the PK levels of carriers. In other words, we assume normal distributions for $\tilde{Y} = Y^{-0.56}$. The same transformation is applied
to the healthy females PK levels, i.e. $\tilde{X} = X^{-0.56}$. The normal quantile-quantile plots for $\tilde{Y}_i$'s and $\tilde{X}_i$'s are shown in Figure 1.

![Q-Q Plot of the transformed $Y$](image1.png) ![Q-Q Plot of the transformed $X$](image2.png)

Figure 1: Normal quantile-quantile plots of the transformed PK levels.

This Q-Q plot suggests that the parametric method (ML) can not be used for the transformed data. However, we can use the newly proposed semi-parametric method to construct confidence intervals for

$$\delta_{p_{1}\geq p_{0}}/(p_{1} - p_{0}) = P\{Y > X, X \in (q_{1}, q_{0}); \theta\}/(p_{1} - p_{0}),$$

the normalized partial AUC of the test, where $X$ and $Y$ denote serum PK levels for a healthy and a carrier female, respectively.
Table V and Table VI show various confidence intervals for the normalized partial AUC’s and full AUC. All the intervals indicate that the serum enzyme test has low to moderate accuracy in detecting DMD.
CHAPTER 6
CONCLUSION AND DISCUSSION

Diagnostic tests are central in the field of modern medicine. Accurate diagnostic tests provide critical information about the disease states of patients. The partial AUC is one of the most important summary measures for the accuracy of a diagnostic test. Finding confidence regions of the partial AUC over an interested FPR interval is an appropriate way to evaluate the diagnostic accuracy of a test. In diagnostic studies, semi-parametric approaches often are useful alternative to the parametric and nonparametric counterparts since semi-parametric methods may inherit good properties of both parametric and nonparametric methods. In this thesis, we have developed new semi-parametric methods for the inference on the partial AUC. In particular, seven confidence intervals are proposed for the partial AUC by using the normal approximation, bootstrap and empirical likelihood methods. The simulation studies indicate that the bootstrap and EL-based methods perform better than the normal approximation (NA) based method. Therefore, the use of hybrid bootstrap and EL-based confidence intervals for the partial AUC is recommended in evaluating the accuracy of diagnostic tests when the test results follow a semi-parametric model.
REFERENCES


APPENDIX A: SIMULATION TABLES
Table I: Normal distribution: Level of 95 percent confidence interval for the pAUC.

| Sample size (m,n) | Method | \((p_0, p_1) = (0, 0.4)\) | | \((p_0, p_1) = (0, 0.7)\) | | \((p_0, p_1) = (0.05, 0.5)\) |
|------------------|--------|-----------------|-----------------|-----------------|-----------------|
|                  |        | Coverage probability | Length | Coverage probability | Length | Coverage probability | Length |
| (30, 30)         | NA     | 0.984            | 0.231   | 0.983            | 0.301   | 0.982            | 0.266   |
|                  | BI     | 0.965            | 0.181   | 0.960            | 0.243   | 0.967            | 0.195   |
|                  | BII    | 0.954            | 0.181   | 0.937            | 0.243   | 0.943            | 0.195   |
|                  | HBELI  | 0.959            | 0.178   | 0.966            | 0.239   | 0.968            | 0.192   |
|                  | HBELII | 0.947            | 0.167   | 0.936            | 0.230   | 0.964            | 0.179   |
|                  | HBELIII| 0.957            | 0.180   | 0.954            | 0.239   | 0.972            | 0.194   |
|                  | HBELIV | 0.963            | 0.187   | 0.961            | 0.245   | 0.979            | 0.200   |
|                  | ML     | 0.927            | 0.153   | 0.923            | 0.215   | 0.921            | 0.165   |
| (50, 50)         | NA     | 0.984            | 0.177   | 0.990            | 0.233   | 0.986            | 0.204   |
|                  | BI     | 0.954            | 0.134   | 0.952            | 0.182   | 0.959            | 0.144   |
|                  | BII    | 0.931            | 0.134   | 0.935            | 0.182   | 0.932            | 0.144   |
|                  | HBELI  | 0.957            | 0.133   | 0.954            | 0.180   | 0.961            | 0.142   |
|                  | HBELII | 0.960            | 0.127   | 0.953            | 0.176   | 0.963            | 0.137   |
|                  | HBELIII| 0.958            | 0.132   | 0.953            | 0.178   | 0.964            | 0.141   |
|                  | HBELIV | 0.960            | 0.136   | 0.961            | 0.182   | 0.969            | 0.146   |
|                  | ML     | 0.940            | 0.119   | 0.944            | 0.167   | 0.944            | 0.128   |
| (100, 100)       | NA     | 0.993            | 0.125   | 0.995            | 0.164   | 0.992            | 0.144   |
|                  | BI     | 0.961            | 0.091   | 0.954            | 0.124   | 0.953            | 0.099   |
|                  | BII    | 0.949            | 0.091   | 0.948            | 0.124   | 0.940            | 0.099   |
|                  | HBELI  | 0.958            | 0.091   | 0.954            | 0.124   | 0.957            | 0.099   |
|                  | HBELII | 0.961            | 0.088   | 0.959            | 0.122   | 0.959            | 0.097   |
|                  | HBELIII| 0.951            | 0.090   | 0.954            | 0.123   | 0.951            | 0.098   |
|                  | HBELIV | 0.957            | 0.092   | 0.958            | 0.125   | 0.960            | 0.101   |
|                  | ML     | 0.949            | 0.085   | 0.940            | 0.119   | 0.954            | 0.092   |
| (50, 50)         | NA     | 0.980            | 0.187   | 0.980            | 0.251   | 0.987            | 0.217   |
|                  | BI     | 0.950            | 0.150   | 0.945            | 0.204   | 0.957            | 0.162   |
|                  | BII    | 0.927            | 0.150   | 0.939            | 0.204   | 0.936            | 0.162   |
|                  | HBELI  | 0.946            | 0.148   | 0.949            | 0.201   | 0.959            | 0.160   |
|                  | HBELII | 0.933            | 0.140   | 0.934            | 0.197   | 0.951            | 0.154   |
|                  | HBELIII| 0.948            | 0.150   | 0.934            | 0.198   | 0.953            | 0.159   |
|                  | HBELIV | 0.956            | 0.151   | 0.943            | 0.204   | 0.956            | 0.163   |
|                  | ML     | 0.940            | 0.134   | 0.942            | 0.192   | 0.937            | 0.147   |
| (80, 50)         | NA     | 0.984            | 0.147   | 0.990            | 0.197   | 0.991            | 0.170   |
|                  | BI     | 0.951            | 0.114   | 0.949            | 0.157   | 0.951            | 0.125   |
|                  | BII    | 0.943            | 0.114   | 0.942            | 0.157   | 0.938            | 0.125   |
|                  | HBELI  | 0.951            | 0.113   | 0.952            | 0.156   | 0.950            | 0.124   |
|                  | HBELII | 0.944            | 0.110   | 0.942            | 0.153   | 0.947            | 0.121   |
|                  | HBELIII| 0.944            | 0.112   | 0.941            | 0.153   | 0.953            | 0.123   |
|                  | HBELIV | 0.956            | 0.115   | 0.948            | 0.158   | 0.959            | 0.126   |
|                  | ML     | 0.941            | 0.106   | 0.951            | 0.151   | 0.946            | 0.115   |
Table II: Normal distribution: Level of 90 percent confidence interval for the pAUC.

<table>
<thead>
<tr>
<th>Sample size (m,n)</th>
<th>Method</th>
<th>((p_0, p_1)=(0, 0.4))</th>
<th>((p_0, p_1)=(0, 0.7))</th>
<th>((p_0, p_1)=(0.05, 0.5))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Coverage probability</td>
<td>Length</td>
<td>Coverage probability</td>
</tr>
<tr>
<td>(30, 30)</td>
<td>NA</td>
<td>0.970</td>
<td>0.194</td>
<td>0.966</td>
</tr>
<tr>
<td></td>
<td>BI</td>
<td>0.915</td>
<td>0.154</td>
<td>0.910</td>
</tr>
<tr>
<td></td>
<td>BII</td>
<td>0.886</td>
<td>0.154</td>
<td>0.868</td>
</tr>
<tr>
<td></td>
<td>HBELI</td>
<td>0.920</td>
<td>0.152</td>
<td>0.913</td>
</tr>
<tr>
<td></td>
<td>HBELII</td>
<td>0.893</td>
<td>0.140</td>
<td>0.891</td>
</tr>
<tr>
<td></td>
<td>HBELIII</td>
<td>0.924</td>
<td>0.150</td>
<td>0.909</td>
</tr>
<tr>
<td></td>
<td>HBELIV</td>
<td>0.939</td>
<td>0.159</td>
<td>0.923</td>
</tr>
<tr>
<td></td>
<td>ML</td>
<td>0.882</td>
<td>0.128</td>
<td>0.876</td>
</tr>
<tr>
<td>(50, 50)</td>
<td>NA</td>
<td>0.963</td>
<td>0.148</td>
<td>0.974</td>
</tr>
<tr>
<td></td>
<td>BI</td>
<td>0.913</td>
<td>0.113</td>
<td>0.923</td>
</tr>
<tr>
<td></td>
<td>BII</td>
<td>0.895</td>
<td>0.113</td>
<td>0.896</td>
</tr>
<tr>
<td></td>
<td>HBELI</td>
<td>0.915</td>
<td>0.112</td>
<td>0.924</td>
</tr>
<tr>
<td></td>
<td>HBELII</td>
<td>0.900</td>
<td>0.106</td>
<td>0.907</td>
</tr>
<tr>
<td></td>
<td>HBELIII</td>
<td>0.920</td>
<td>0.112</td>
<td>0.919</td>
</tr>
<tr>
<td></td>
<td>HBELIV</td>
<td>0.930</td>
<td>0.114</td>
<td>0.925</td>
</tr>
<tr>
<td></td>
<td>ML</td>
<td>0.888</td>
<td>0.100</td>
<td>0.885</td>
</tr>
<tr>
<td>(100, 100)</td>
<td>NA</td>
<td>0.975</td>
<td>0.104</td>
<td>0.975</td>
</tr>
<tr>
<td></td>
<td>BI</td>
<td>0.912</td>
<td>0.076</td>
<td>0.891</td>
</tr>
<tr>
<td></td>
<td>BII</td>
<td>0.903</td>
<td>0.076</td>
<td>0.903</td>
</tr>
<tr>
<td></td>
<td>HBELI</td>
<td>0.912</td>
<td>0.076</td>
<td>0.893</td>
</tr>
<tr>
<td></td>
<td>HBELII</td>
<td>0.905</td>
<td>0.075</td>
<td>0.897</td>
</tr>
<tr>
<td></td>
<td>HBELIII</td>
<td>0.901</td>
<td>0.076</td>
<td>0.910</td>
</tr>
<tr>
<td></td>
<td>HBELIV</td>
<td>0.909</td>
<td>0.077</td>
<td>0.920</td>
</tr>
<tr>
<td></td>
<td>ML</td>
<td>0.897</td>
<td>0.071</td>
<td>0.887</td>
</tr>
<tr>
<td>(50, 30)</td>
<td>NA</td>
<td>0.961</td>
<td>0.157</td>
<td>0.953</td>
</tr>
<tr>
<td></td>
<td>BI</td>
<td>0.912</td>
<td>0.126</td>
<td>0.898</td>
</tr>
<tr>
<td></td>
<td>BII</td>
<td>0.893</td>
<td>0.126</td>
<td>0.881</td>
</tr>
<tr>
<td></td>
<td>HBELI</td>
<td>0.915</td>
<td>0.125</td>
<td>0.902</td>
</tr>
<tr>
<td></td>
<td>HBELII</td>
<td>0.894</td>
<td>0.118</td>
<td>0.902</td>
</tr>
<tr>
<td></td>
<td>HBELIII</td>
<td>0.909</td>
<td>0.122</td>
<td>0.893</td>
</tr>
<tr>
<td></td>
<td>HBELIV</td>
<td>0.919</td>
<td>0.127</td>
<td>0.897</td>
</tr>
<tr>
<td></td>
<td>ML</td>
<td>0.885</td>
<td>0.113</td>
<td>0.892</td>
</tr>
<tr>
<td>(80, 50)</td>
<td>NA</td>
<td>0.964</td>
<td>0.124</td>
<td>0.959</td>
</tr>
<tr>
<td></td>
<td>BI</td>
<td>0.926</td>
<td>0.096</td>
<td>0.919</td>
</tr>
<tr>
<td></td>
<td>BII</td>
<td>0.912</td>
<td>0.096</td>
<td>0.912</td>
</tr>
<tr>
<td></td>
<td>HBELI</td>
<td>0.928</td>
<td>0.096</td>
<td>0.919</td>
</tr>
<tr>
<td></td>
<td>HBELII</td>
<td>0.886</td>
<td>0.092</td>
<td>0.902</td>
</tr>
<tr>
<td></td>
<td>HBELIII</td>
<td>0.897</td>
<td>0.094</td>
<td>0.890</td>
</tr>
<tr>
<td></td>
<td>HBELIV</td>
<td>0.895</td>
<td>0.096</td>
<td>0.898</td>
</tr>
<tr>
<td></td>
<td>ML</td>
<td>0.888</td>
<td>0.089</td>
<td>0.893</td>
</tr>
</tbody>
</table>
Table III: Exponential distribution: Level of 95 percent confidence interval for the pAUC.

<table>
<thead>
<tr>
<th>Sample size (m,n)</th>
<th>Method</th>
<th>((P_0, P_1) = (0, 0.4))</th>
<th>((P_0, P_1) = (0, 0.7))</th>
<th>((P_0, P_1) = (0.05, 0.5))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Coverage probability</td>
<td>Length</td>
<td>Coverage probability</td>
</tr>
<tr>
<td>(30, 30)</td>
<td>NA</td>
<td>0.981</td>
<td>0.204</td>
<td>0.979</td>
</tr>
<tr>
<td></td>
<td>BI</td>
<td>0.966</td>
<td>0.152</td>
<td>0.958</td>
</tr>
<tr>
<td></td>
<td>BII</td>
<td>0.958</td>
<td>0.152</td>
<td>0.951</td>
</tr>
<tr>
<td></td>
<td>HBELI</td>
<td>0.968</td>
<td>0.150</td>
<td>0.960</td>
</tr>
<tr>
<td></td>
<td>HBELII</td>
<td>0.944</td>
<td>0.142</td>
<td>0.961</td>
</tr>
<tr>
<td></td>
<td>HBELIII</td>
<td>0.967</td>
<td>0.152</td>
<td>0.965</td>
</tr>
<tr>
<td></td>
<td>HBELIV</td>
<td>0.974</td>
<td>0.159</td>
<td>0.970</td>
</tr>
<tr>
<td></td>
<td>ML</td>
<td>0.922</td>
<td>0.130</td>
<td>0.942</td>
</tr>
<tr>
<td>(50, 50)</td>
<td>NA</td>
<td>0.991</td>
<td>0.158</td>
<td>0.985</td>
</tr>
<tr>
<td></td>
<td>BI</td>
<td>0.950</td>
<td>0.112</td>
<td>0.958</td>
</tr>
<tr>
<td></td>
<td>BII</td>
<td>0.939</td>
<td>0.112</td>
<td>0.942</td>
</tr>
<tr>
<td></td>
<td>HBELI</td>
<td>0.948</td>
<td>0.111</td>
<td>0.961</td>
</tr>
<tr>
<td></td>
<td>HBELII</td>
<td>0.951</td>
<td>0.108</td>
<td>0.950</td>
</tr>
<tr>
<td></td>
<td>HBELIII</td>
<td>0.956</td>
<td>0.111</td>
<td>0.949</td>
</tr>
<tr>
<td></td>
<td>HBELIV</td>
<td>0.964</td>
<td>0.113</td>
<td>0.958</td>
</tr>
<tr>
<td></td>
<td>ML</td>
<td>0.933</td>
<td>0.101</td>
<td>0.941</td>
</tr>
<tr>
<td>(100, 100)</td>
<td>NA</td>
<td>0.996</td>
<td>0.111</td>
<td>0.966</td>
</tr>
<tr>
<td></td>
<td>BI</td>
<td>0.957</td>
<td>0.077</td>
<td>0.940</td>
</tr>
<tr>
<td></td>
<td>BII</td>
<td>0.949</td>
<td>0.077</td>
<td>0.937</td>
</tr>
<tr>
<td></td>
<td>HBELI</td>
<td>0.957</td>
<td>0.077</td>
<td>0.939</td>
</tr>
<tr>
<td></td>
<td>HBELII</td>
<td>0.946</td>
<td>0.075</td>
<td>0.950</td>
</tr>
<tr>
<td></td>
<td>HBELIII</td>
<td>0.951</td>
<td>0.076</td>
<td>0.948</td>
</tr>
<tr>
<td></td>
<td>HBELIV</td>
<td>0.964</td>
<td>0.078</td>
<td>0.952</td>
</tr>
<tr>
<td></td>
<td>ML</td>
<td>0.951</td>
<td>0.072</td>
<td>0.949</td>
</tr>
<tr>
<td>(50, 30)</td>
<td>NA</td>
<td>0.980</td>
<td>0.168</td>
<td>0.987</td>
</tr>
<tr>
<td></td>
<td>BI</td>
<td>0.950</td>
<td>0.125</td>
<td>0.956</td>
</tr>
<tr>
<td></td>
<td>BII</td>
<td>0.943</td>
<td>0.125</td>
<td>0.949</td>
</tr>
<tr>
<td></td>
<td>HBELI</td>
<td>0.952</td>
<td>0.124</td>
<td>0.956</td>
</tr>
<tr>
<td></td>
<td>HBELII</td>
<td>0.941</td>
<td>0.120</td>
<td>0.947</td>
</tr>
<tr>
<td></td>
<td>HBELIII</td>
<td>0.956</td>
<td>0.126</td>
<td>0.949</td>
</tr>
<tr>
<td></td>
<td>HBELIV</td>
<td>0.966</td>
<td>0.130</td>
<td>0.952</td>
</tr>
<tr>
<td></td>
<td>ML</td>
<td>0.937</td>
<td>0.117</td>
<td>0.936</td>
</tr>
<tr>
<td>(80, 50)</td>
<td>NA</td>
<td>0.992</td>
<td>0.131</td>
<td>0.988</td>
</tr>
<tr>
<td></td>
<td>BI</td>
<td>0.954</td>
<td>0.097</td>
<td>0.954</td>
</tr>
<tr>
<td></td>
<td>BII</td>
<td>0.946</td>
<td>0.097</td>
<td>0.948</td>
</tr>
<tr>
<td></td>
<td>HBELI</td>
<td>0.952</td>
<td>0.096</td>
<td>0.954</td>
</tr>
<tr>
<td></td>
<td>HBELII</td>
<td>0.955</td>
<td>0.094</td>
<td>0.957</td>
</tr>
<tr>
<td></td>
<td>HBELIII</td>
<td>0.950</td>
<td>0.097</td>
<td>0.946</td>
</tr>
<tr>
<td></td>
<td>HBELIV</td>
<td>0.963</td>
<td>0.099</td>
<td>0.951</td>
</tr>
<tr>
<td></td>
<td>ML</td>
<td>0.933</td>
<td>0.092</td>
<td>0.939</td>
</tr>
</tbody>
</table>
Table IV: Exponential distribution: Level of 90 percent confidence interval for the pAUC.

<table>
<thead>
<tr>
<th>Sample size (m,n)</th>
<th>Method</th>
<th>((p_0, p_1)=(0, 0.4))</th>
<th>((p_0, p_1)=(0, 0.7))</th>
<th>((p_0, p_1)=(0.05, 0.5))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Coverage probability</td>
<td>Coverage probability</td>
<td>Coverage probability</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Length</td>
<td>Length</td>
<td>Length</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.974 0.172 0.980 0.225</td>
<td>0.975 0.195</td>
<td></td>
</tr>
<tr>
<td>(30, 30)</td>
<td>NA</td>
<td>0.918 0.127 0.928 0.182</td>
<td>0.936 0.137</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BI</td>
<td>0.892 0.127 0.921 0.182</td>
<td>0.902 0.137</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BII</td>
<td>0.918 0.125 0.930 0.180</td>
<td>0.942 0.136</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HBELII</td>
<td>0.901 0.119 0.926 0.179</td>
<td>0.926 0.131</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HBELIII</td>
<td>0.923 0.126 0.923 0.180</td>
<td>0.947 0.139</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HBELIV</td>
<td>0.936 0.134 0.934 0.189</td>
<td>0.952 0.147</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ML</td>
<td>0.877 0.109 0.887 0.160</td>
<td>0.881 0.117</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.977 0.132 0.965 0.173</td>
<td>0.979 0.150</td>
<td></td>
</tr>
<tr>
<td>(50, 50)</td>
<td>NA</td>
<td>0.910 0.094 0.911 0.135</td>
<td>0.929 0.102</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BI</td>
<td>0.898 0.094 0.898 0.135</td>
<td>0.911 0.102</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BII</td>
<td>0.912 0.094 0.913 0.135</td>
<td>0.932 0.102</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HBELII</td>
<td>0.889 0.091 0.909 0.135</td>
<td>0.923 0.099</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HBELIII</td>
<td>0.907 0.093 0.906 0.134</td>
<td>0.926 0.102</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HBELIV</td>
<td>0.914 0.095 0.916 0.138</td>
<td>0.937 0.105</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ML</td>
<td>0.877 0.085 0.897 0.124</td>
<td>0.897 0.091</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.977 0.093 0.983 0.122</td>
<td>0.984 0.106</td>
<td></td>
</tr>
<tr>
<td>(100, 100)</td>
<td>NA</td>
<td>0.909 0.064 0.907 0.093</td>
<td>0.899 0.070</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BI</td>
<td>0.901 0.064 0.904 0.093</td>
<td>0.887 0.070</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BII</td>
<td>0.912 0.064 0.907 0.093</td>
<td>0.900 0.070</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HBELII</td>
<td>0.901 0.064 0.900 0.092</td>
<td>0.906 0.070</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HBELIII</td>
<td>0.914 0.064 0.904 0.092</td>
<td>0.918 0.070</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HBELIV</td>
<td>0.922 0.065 0.909 0.094</td>
<td>0.922 0.072</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ML</td>
<td>0.887 0.061 0.893 0.088</td>
<td>0.903 0.065</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.972 0.141 0.958 0.187</td>
<td>0.981 0.189</td>
<td></td>
</tr>
<tr>
<td>(50, 30)</td>
<td>NA</td>
<td>0.917 0.105 0.893 0.154</td>
<td>0.905 0.114</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BI</td>
<td>0.910 0.105 0.886 0.154</td>
<td>0.891 0.114</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BII</td>
<td>0.920 0.104 0.898 0.152</td>
<td>0.911 0.113</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HBELII</td>
<td>0.895 0.101 0.897 0.151</td>
<td>0.916 0.111</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HBELIII</td>
<td>0.913 0.104 0.896 0.150</td>
<td>0.916 0.114</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HBELIV</td>
<td>0.924 0.110 0.914 0.157</td>
<td>0.933 0.120</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ML</td>
<td>0.886 0.098 0.892 0.144</td>
<td>0.889 0.106</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.972 0.110 0.957 0.147</td>
<td>0.982 0.148</td>
<td></td>
</tr>
<tr>
<td>(80, 50)</td>
<td>NA</td>
<td>0.902 0.081 0.906 0.118</td>
<td>0.910 0.089</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BI</td>
<td>0.896 0.081 0.900 0.118</td>
<td>0.900 0.089</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BII</td>
<td>0.906 0.081 0.907 0.117</td>
<td>0.911 0.089</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HBELII</td>
<td>0.891 0.079 0.912 0.116</td>
<td>0.920 0.087</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HBELIII</td>
<td>0.900 0.080 0.899 0.116</td>
<td>0.922 0.088</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HBELIV</td>
<td>0.909 0.083 0.908 0.120</td>
<td>0.930 0.091</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ML</td>
<td>0.879 0.077 0.878 0.113</td>
<td>0.877 0.083</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX B: REAL DATA ANALYSIS TABLES
Table V: Duchenne Muscular Dystrophy Example: Level of 95 percent confidence intervals for the normalized pAUC

<table>
<thead>
<tr>
<th>Method</th>
<th>$(p_0, p_1) = (0, 0.4)$</th>
<th>$(p_0, p_1) = (0, 0.7)$</th>
<th>$(p_0, p_1) = (0.05, 0.5)$</th>
<th>$(p_0, p_1) = (0, 1.0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>NA</td>
<td>(0.520, 0.770)</td>
<td>(0.663, 0.836)</td>
<td>(0.613, 0.824)</td>
<td>(0.749, 0.875)</td>
</tr>
<tr>
<td>BI</td>
<td>(0.543, 0.748)</td>
<td>(0.673, 0.824)</td>
<td>(0.620, 0.816)</td>
<td>(0.747, 0.876)</td>
</tr>
<tr>
<td>BII</td>
<td>(0.545, 0.750)</td>
<td>(0.673, 0.824)</td>
<td>(0.609, 0.804)</td>
<td>(0.746, 0.875)</td>
</tr>
<tr>
<td>HBELI</td>
<td>(0.535, 0.740)</td>
<td>(0.667, 0.819)</td>
<td>(0.616, 0.809)</td>
<td>(0.740, 0.870)</td>
</tr>
<tr>
<td>HBELII</td>
<td>(0.535, 0.740)</td>
<td>(0.669, 0.819)</td>
<td>(0.620, 0.804)</td>
<td>(0.741, 0.869)</td>
</tr>
<tr>
<td>HBELIII</td>
<td>(0.538, 0.738)</td>
<td>(0.661, 0.824)</td>
<td>(0.611, 0.811)</td>
<td>(0.740, 0.870)</td>
</tr>
<tr>
<td>HBELIV</td>
<td>(0.533, 0.743)</td>
<td>(0.660, 0.824)</td>
<td>(0.613, 0.809)</td>
<td>(0.737, 0.872)</td>
</tr>
<tr>
<td>Estimate of pAUC</td>
<td>0.6442331</td>
<td>0.7490747</td>
<td>0.7180316</td>
<td>0.8116641</td>
</tr>
</tbody>
</table>
Table VI: Duchenne Muscular Dystrophy Example: Level of 90 percent confidence intervals for the normalized pAUC

<table>
<thead>
<tr>
<th>Method</th>
<th>((p_0, p_1) = (0, 0.4))</th>
<th>((p_0, p_1) = (0, 0.7))</th>
<th>((p_0, p_1) = (0, 0.5))</th>
<th>((p_0, p_1) = (0, 1.0))</th>
</tr>
</thead>
<tbody>
<tr>
<td>NA</td>
<td>(0.54,0.750)</td>
<td>(0.676,0.821)</td>
<td>(0.629,0.807)</td>
<td>(0.759,0.864)</td>
</tr>
<tr>
<td>BI</td>
<td>(0.555,0.733)</td>
<td>(0.680,0.819)</td>
<td>(0.636,0.800)</td>
<td>(0.760,0.864)</td>
</tr>
<tr>
<td>BII</td>
<td>(0.558,0.738)</td>
<td>(0.680,0.819)</td>
<td>(0.640,0.804)</td>
<td>(0.759,0.863)</td>
</tr>
<tr>
<td>HBEI</td>
<td>(0.550,0.730)</td>
<td>(0.674,0.814)</td>
<td>(0.631,0.796)</td>
<td>(0.755,0.860)</td>
</tr>
<tr>
<td>HBEII</td>
<td>(0.550,0.728)</td>
<td>(0.674,0.814)</td>
<td>(0.633,0.796)</td>
<td>(0.756,0.859)</td>
</tr>
<tr>
<td>HBEIII</td>
<td>(0.560,0.720)</td>
<td>(0.687,0.804)</td>
<td>(0.638,0.791)</td>
<td>(0.755,0.860)</td>
</tr>
<tr>
<td>HBEIV</td>
<td>(0.553,0.728)</td>
<td>(0.686,0.804)</td>
<td>(0.633,0.796)</td>
<td>(0.756,0.859)</td>
</tr>
<tr>
<td>Estimate of pAUC</td>
<td>0.6442331</td>
<td>0.7490747</td>
<td>0.7180316</td>
<td>0.8116641</td>
</tr>
</tbody>
</table>
APPENDIX C: THE S-PLUS CODE FOR SIMULATION STUDIES

# solve non-linear equations (for solving lambda)
solveNonlinear<-function(f,y0,x,...)
{
g<-function(x,y0,f)sum((f(x)-y0)^2)
g$y0<-y0
g$f<-f
nlmin(g,x,max.fcal=1000,max.iter=1000,...)
}

iter<-1000        # number of iteration
m<-30            # sample size of non-diseased sample: x1,...xm
n<-30            # sample size of diseased sample: y1,...yn
rho<-m/n
mux<-0           # mean of the non-diseased population
stddx<-1         # standard deviation of the non-diseased population
muy<-1           # mean of the diseased population
stddy<-1         # standard deviation of the diseased population
p0<-0.05
p1<-0.5
alpha<-0.1
z<-qnorm(1-alpha/2)

# Function R(p)

rp<-function(p,muy,stddy)
{
y<-1-pnorm(qnorm(1-p),muy,stddy)
y
}

# true pauc
truepauc<-integrate(rp, muy=muy,stddy=stddy,lower=p0,upper=p1)$integral

# Function R(p)=1-pnorm((mu0-mu1+sigma0*qnorm(1-p))/sigma1) for ML estimate

rpml <- function(p, mu0, sigma0, mu1, sigma1)
{
  1-pnorm((mu0-mu1+sigma0*qnorm(1-p))/sigma1)
}

# function used to calculate the variance of ML estimate of pAUC
# when the underlying distribution is the normal distribution.

pvar1 <- function(p, mu0, sigma0, mu1, sigma1)
{
  dnorm((mu0-mu1+sigma0*qnorm(1-p))/sigma1)
}
pvar2 <- function(p, mu0, sigma0, mu1, sigma1)
{
  (qnorm(1-p))*dnorm((mu0-mu1+sigma0*qnorm(1-p))/sigma1)
}

normalcoverage <- 0
normallength <- 0
bootcoverage1 <- 0
bootlength1 <- 0
bootcoverage2 <- 0
bootlength2 <- 0
ecoverage1 <- 0
ecoverage2 <- 0
eellength1 <- 0
eellength2 <- 0
ecoverage3 <- 0
eellength3 <- 0
ecoverage4 <- 0
eellength4 <- 0
mlcoverage <- 0
mllength <- 0

# bootstrap function
mybootstrap <- function(B, x, y, m, n, p0, p1)
{
  Starrpauc <- 0
  Starssquare <- 0
  Starvm <- matrix(0, ncol = B, nrow = m)

  for (b in 1:B)
  {
    xb <- sample(x, m, replace = T)
    yb <- sample(y, n, replace = T)
    hatq0b <- 0
    hatq1b <- 0
    hatq0b <- quantile(xb, c(1-p0))      # hatq0b, hatq1b: sample quantiles
    hatq1b <- quantile(xb, c(1-p1))      # of F in bootstrap
    starmlemean <- mean(yb)              # MLE estimate of startheta
    starmlestdd <- (mean((yb - starmlemean)^2))^(1/2)
    starv <- 0                           # compute starv
    for (i in (1:m))
    {
      starv[i] <- (1-pnorm(xb[i], starmlemean, starmlestdd)) * (hatq1b <= xb[i]) * (xb[i] <= hatq0b)
      Starvm[i, b] <- (1-pnorm(xb[i], starmlemean, starmlestdd)) * (hatq1b <= xb[i]) * (xb[i] <= hatq0b)
    }
    Starrpauc[b] <- mean(starv)
    Starssquare[b] <- var(starv)
  }
  list(starrpauc, starssquare, Starvm)
}

# Loop
for (loop in c(1:iter))
{
  x <- rnorm(m, mux, stddx)     # obs from the non-diseased population
  y <- rnorm(n, muy, stddy)    # obs from the diseased population
hatq0<-quantile(x, c(1-p0))  # hatq0, hatq1: sample quantiles fo F
hatq1<-quantile(x, c(1-p1))

mlemean<-mean(y)              # MLE estimate of theta
mlestdd<-(mean((y-mlemean)^2))^(1/2)

hatv<-rep(0,m)                # hat v_i=(1-hat G(x_i, hat theta))*
for (i in 1:m)
{
  hatv[i]<-(1-pnorm(x[i],mlemean,mlestdd))*(hatq1 <=x[i])*(x[i] <= hatq0)
}

hatpauc<- mean(hatv)          # estimate of pauc

# compute the normal approxamation based C.I.

s<-matrix(0,ncol=2,nrow=2)
d<-matrix(0,ncol=1,nrow=2)

s[1,1]<-mlestdd^2
s[1,2]<0
s[2,1]<0
s[2,2]<2*mlestdd^4

# compute B(x,theta,q0,q1)
b<-rep(0,m)
for (i in 1:m)
{
  b[i]<-((1-pnorm(x[i],mlemean,mlestdd))*(hatq1 <=x[i])*(x[i] <= hatq0) -
  hatpauc) - (1-pnorm(hatq0,mlemean,mlestdd))*(x[i]<=hatq0)-(1-p0)) -
  (1-pnorm(hatq1,mlemean,mlestdd))*(x[i]<=hatq1)-(1-p1))
}

# compute the deravative of G(pnorm)
dg1<-rep(0,m)
dg2<-rep(0,m)
for (i in 1:m)
{
  dg1[i]<-dnorm(x[i],mlemean,mlestdd)*(hatq1 <=x[i])*(x[i] <= hatq0) -
  hatpauc) - (1-pnorm(hatq0,mlemean,mlestdd))*(x[i]-mlemean)*(hatq1 <=
  hatq0)/(2*mlestdd^2)
}

# compute matrix D

d[1,1]<mean(dg1)
d[2,1]<mean(dg2)

# compute hatsigmasquare
hatsigmasquare<-var(b)+rho*t(d)%*%s%*%d

normallow<-hatpauc-z*sqrt(hatsigmasquare)/sqrt(m)
normalup<-hatpauc+z*sqrt(hatsigmasquare)/sqrt(m)

length<-normalup-normallow
normallength<-normallength+length

if ((normallow<=truepauc) & (normalup>=truepauc)) normalcoverage<-normallength+1
# end of computing the normal approximation based C.I.

# compute the CI by using bootstrap method

B<-150
starvm<-matrix(0,m,B)
lstar<-rep(0,B)
bootlist<-mybootstrap(B,x,y,m,n,p0,p1)
starpauc<-bootlist[[1]]
starssquare<-bootlist[[2]]
starvm<-bootlist[[3]]

starsigmasquare<-var(starpauc)
meanstarpauc<-mean(starpauc)

bootlow1<-hatpauc-z*sqrt(starsigmasquare)
bootup1<-hatpauc+z*sqrt(starsigmasquare)

bootlow2<-meanstarpauc-z*sqrt(starsigmasquare)
bootup2<-meanstarpauc+z*sqrt(starsigmasquare)

length1<-bootup1-bootlow1
length2<-bootup2-bootlow2

bootlength1<-bootlength1+length1
bootlength2<-bootlength2+length2

if ((bootlow1<=truepauc) & (bootup1>=truepauc)) bootcoverage1<-bootcoverage1+1
if ((bootlow2<=truepauc) & (bootup2>=truepauc)) bootcoverage2<-bootcoverage2+1

# end of computing the CI by using bootstrap method

# compute the EL-based semi-parametric CI: HBELI, HBELII

starsigmasquare<-var(starpauc)
meanstarpauc<-mean(starpauc)

hatssquare<-mean((hatv-hatpauc)^2)
starc<-hatssquare/(m*starsigmasquare)

# compute lambda and el bound for first (hatc)
# x[1]: lambda. x[2]: pauc

# compute lambda and el bound for first (starc)
f<-function(x) c(mean((hatv-x[2])/(1+x[1]*(hatv-x[2]))),starc*2*sum(log(abs(1+x[1]*(hatv-x[2])))))-qchisq(1-alpha,1))
sollow<-solveNonlinear(f,c(0,0.01),c(0,0.01))
solup<-solveNonlinear(f,c(0,0),c(0,0.95))
eellow1<-sollow$x[2]
elup1<-solup$x[2]

starc<-meanstarssquare/(m*starsigmasquare)

# compute lambda and el bound for second (starc)
\begin{verbatim}
f<-function(x) c(mean((hatv-x[2])/(1+x[1]*(hatv-x[2]))),starc*2*sum(log(abs(1+x[1]*(hatv-x[2]))))-qchisq(1-alpha,1))
sollow<-solveNonlinear(f,c(0,0),c(0,0.01))
solup<-solveNonlinear(f,c(0,0),c(0,0.95))
ellow2<-sollow$x[2]
elpu2<-solup$x[2]

length1<-elup1-ellow1
length2<-elup2-ellow2

ellength1<-ellength1+length1
ellength2<-ellength2+length2

if ((ellow1<=truepauc) & (elup1>=truepauc)) elcoverage1<-elcoverage1+1
if ((ellow2<=truepauc) & (elup2>=truepauc)) elcoverage2<-elcoverage2+1

# end of computing the EL-based semi-parametric CI

# compute the CI by using bootstrap method
for (b in 1:B)
{
wjhatb<-rep(0,m)
for (i in 1:m)
{
wjhatb[i]<-starvm[i,b]-hatpauc
}
f<-function(lam)mean(wjhatb/(1+lam*wjhatb))
lambda<-solveNonlinear(f,c(0),c(0.01))$x
lstar[b]<-2*sum(log(abs(1+lambda*wjhatb)))
}
ctail<-1/mean(lstar)
cstar<-quantile(lstar, c(1-alpha))

# compute lambda and el bound for first (hatc)
# x[1]: lambda. x[2]: pauc

# compute lambda and el bound for third
f<-function(x) c(mean((hatv-x[2])/(1+x[1]*(hatv-x[2]))),2*sum(log(abs(1+x[1]*(hatv-x[2]))))-cstar)
sollow<-solveNonlinear(f,c(0,0),c(0,0.01))
solup<-solveNonlinear(f,c(0,0),c(0,0.95))
ellow3<-sollow$x[2]
elpu3<-solup$x[2]

# compute lambda and el bound for fourth
f<-function(x) c(mean((hatv-x[2])/(1+x[1]*(hatv-x[2]))),ctail*2*sum(log(abs(1+x[1]*(hatv-x[2]))))-qchisq(1-alpha,1))
sollow<-solveNonlinear(f,c(0,0),c(0,0.01))
solup<-solveNonlinear(f,c(0,0),c(0,0.95))
ellow4<-sollow$x[2]
elpu4<-solup$x[2]

length3<-elup3-ellow3
length4<-elup4-ellow4
\end{verbatim}
ellength3<-ellength3+length3
ellength4<-ellength4+length4

if ((ellow3<=truepauc) & (elup3>=truepauc)) elcoverage3<-elcoverage3+1
if ((ellow4<=truepauc) & (elup4>=truepauc)) elcoverage4<-elcoverage4+1

# end of computing the EL-based semi-parametric CI

#444444444444444444444444444444445555555555555555555555555555555
#666666666666666666666666666666666666666666666666666666666666666
#computing the ML CI
mu0<-mean(x)
mul<-mean(y)
sigma0<-sqrt(var(x))
sigma1<-sqrt(var(y))

hatpauc<-integrate(rpml, mu0=mu0, sigma0=sigma0, mul=mul, sigma1=sigma1, lower = p0, upper = p1)$integral
  # ML estimate of pauc
kl<-integrate(pvar1, mu0=mu0, sigma0=sigma0, mul=mul, sigma1=sigma1, lower = p0, upper = p1)$integral
k2<-integrate(pvar2, mu0=mu0, sigma0=sigma0, mul=mul, sigma1=sigma1, lower = p0, upper = p1)$integral

varpauc<-kl^2*(sigma0^2/m+sigma1^2/n)/sigma1^2+k2^2*sigma0^2/(2*m*sigma1^2)+((mu0-mul)*kl+sigma0*k2)^2/(2*n*sigma1^2)
  # variance of ML estimate of pauc

## compute the ML-based interval.
mllow<-hatpauc-z*sqrt(varpauc)
mllup<-hatpauc+z*sqrt(varpauc)

if ((mllow <= truepauc) & (mlup >= truepauc))mlcoverage<-mlcoverage+1

# end of ML CI
#666666666666666666666666666666666666666666666666666666666666666
}
# end of loop

coverage<-c(normalcoverage/iter, bootcoverage1/iter,
  bootcoverage2/iter, elcoverage1/iter, elcoverage2/iter, elcoverage3/iter, elcoverage4/iter, mlcoverage/iter)
length<-c(normallength/iter, bootlength1/iter, bootlength2/iter, ellength1/iter, ellength2/iter, ellength3/iter, ellength4/iter, mllength/iter)
cat("m=", m, "n=", n, "iter=", iter, ",\\n")
cat("meanx=", mux, "meany=", muy, "STDX=", stddx, "STDY=", stddy, "p0=", p0, 
  "p1=", "AUC(p0,p1)=", truepauc, ",\\n")
cat("Coverage of (NA, BOOT1, BOOT2, EL1, EL2):", coverage, ",\\n")
cat("Average length of (NA, BOOT1, BOOT2, EL1, EL2):", length, ",\\n")
APPENDIX D: THE S-PLUS CODE FOR REAL DATA ANALYSIS

# solve non-linear equations (for solving lambda)
solveNonlinear <- function(f, y0, x, ...)
{
  g <- function(x, y0, f) sum((f(x) - y0)^2)
  g$y0 <- y0
  g$f <- f
  nlmin(g, x, max.fcal = 1000, max.iter = 1000, ...)
}

m <- 127            # sample size of non-diseased sample: x1, ... xm
n <- 67            # sample size of diseased sample: y1, ... yn
p0 <- 0
p1 <- 0.7
alpha <- 0.05
z <- qnorm(1 - alpha / 2)

mybootstrap <- function(B, x, y, m, n, p0, p1){
  starpauc <- 0
  starssquare <- 0
  starvm <- matrix(0, ncol = B, nrow = m)
  for (b in 1:B)
  {
    xxb <- sample(x, m, replace = T)
    yyb <- sample(y, n, replace = T)
    hatq0b <- 0
    hatq1b <- 0
    xb <- xxb**(-0.56)
    yb <- yyb**(-0.56)
    hatqlb <- quantile(xxb, c(1 - p0))**(-0.56)  # hatq0b, hatqlb: sample quantiles
    hatq0b <- quantile(xxb, c(1 - p1))**(-0.56)  # of F in bootstrap
    starmlemean <- mean(yb)                # MLE estimate of startheta
    starmlestdd <-(mean((yb - starmlemean)^2))^(1/2)
    starv <- 0                           # compute starv
    for (i in (1:m))
    {
      starv[i] <- (1 - pnorm(xb[i], starmlemean, starmlestdd)) * (hatqlb <- xb[i])**(-0.56)
      starvm[i, b] <- (1 - pnorm(xb[i], starmlemean, starmlestdd)) * (hatqlb <- xb[i])**(-0.56)
      starpauc[b] <- mean(starv)
      starssquare[b] <- var(starv)
    }
    list(starpauc, starssquare, starvm)
  }
}

# main code
xx <- c(10.9, 11, 13.2, 22.6, 15.2, 9.6, 13.5, 17.5, 13.3, 17.1, 22.7, 6.9, 14.6, 18.2, 5.6, 7.9, 12.6, 16.1, 9.9, 10.1, 16.4, 15.3, 8.1, 3.5, 20.7, 19.9, 11.8, 13.7, 4.6, 8.8, 9.9, 10.2, 16.8, 6.4, 10.9, 8.6, 16.4, 10.3, 2.8,
\begin{verbatim}
17.1, 10.9, 8, 15.6, 11.8, 5.1, 12, 16.6, 15.3, 4.4, 9.3, 15.1, 16.5, 21.8, 15.8, 10.3, 12, 1
0.5, 6.7, 11.3,
15.3, 13.7, 12.2, 17.9, 15.4, 22.3, 8.7, 5.3, 16.1, 9.8, 12.9, 3.9, 14.2, 16.2, 9.7, 10.3, 5
.8, 10.6, 11.9, 14.5,
14.8, 9.7, 17.1, 10.3, 10, 12.3, 10, 5.9, 9.9, 9.9, 13.7, 12.7, 11.3, 6.9, 15.1, 6.1, 12.2, 7.3, 10.
7, 7, 11.9,
16.6, 7.6, 18.1, 21.6, 12.5, 5.6, 8.9, 3.8, 11.1, 10.8, 10.7, 17.4, 14.5, 15.3, 15.3, 11.9, 12
, 16.4, 12.1, 12.9,
11.7, 14, 10.9, 6.15.3, 11.4, 20.3)

yy<-c(25.6, 26.8, 8.8, 17.4, 23.8, 20.2, 11, 18.3, 16.7, 21.6, 16.1, 36.4, 49.1, 32.2, 14, 16.9
, 12.7, 9.7, 110, 63.7,
73, 23.3, 31.9, 41.6, 18.8, 17, 22, 10.9, 19.9, 18.8, 12.9, 15.5, 20.6, 16, 19.8, 16.4, 10.4
, 17.1, 25.3, 62.9,
51.6, 33.9, 22.2, 20.9, 36, 12.8, 11.7, 18.6, 19.4, 11.2, 21, 20.1, 8.3, 25.2, 16.6, 22.7, 2
1.3, 10.2, 12.1, 22,
14.4, 8.9, 27.1, 49.1, 11.8, 15.1, 14.2)
x<-xx^(-0.56)
y<-yy^(-0.56)
hatq1<-quantile(xx, c(1-p0))^(-0.56)   # hatq0, hatq1: sample quantiles fo F
hatq0<-quantile(xx, c(1-p1))^(-0.56)
mlemean<-mean(y)              # MLE estimate of theta
mlestdd<-(mean((y-mlemean)^2))^(1/2)
hatv<-rep(0,m)                # hat v_i=(1-hat G(x_i, hat theta))*
t<-rep(0,m)                              # I(x_i \in (hatq1,hatq0))
for (i in 1:m)
{
  hatv[i]<-(1-pnorm(x[i],mlemean,mlestdd))*(hatq1 <=x[i])*(x[i] <= hatq0)
  t[i]<-(hatq1 <=x[i])*(x[i] <= hatq0)
}
hatpauc<mean(hatv)          # estimate of pauc
p<mean(t)

# compute the normal approximation based C.I.
s<-matrix(0,ncol=2,nrow=2)
d<-matrix(0,ncol=1,nrow=2)
s[1,1]<-mlestdd^2
s[1,2]<-0
s[2,1]<-0
s[2,2]<-2*mlestdd^4
b<-rep(0,m)
for (i in 1:m)
{
b[i]<-((1-pnorm(x[i],mlemean,mlestdd))*(hatq1 <=x[i])*(x[i] <= hatq0)-
hatpauc)
- (1-pnorm(hatq0,mlemean,mlestdd))*(x[i] <= hatq0) -(1-p0))
\end{verbatim}
(1 - pnorm(hatq1, mlemean, mlestdd)) * ((x[i] <= hatq1) - (1 - p1))
}

# compute the derivative of G(pnorm)
dg1 <- rep(0, m)
dg2 <- rep(0, m)
for (i in 1:m)
{
dg1[i] <- -dnorm(x[i], mlemean, mlestdd) * (hatq1 <= x[i]) * (x[i] <= hatq0)
dg2[i] <- -dnorm(x[i], mlemean, mlestdd) * (x[i] - mlemean) * (hatq1 <= x[i]) * (x[i] <= hatq0) / (2 * mlestdd^2)
}

# compute matrix D
d[1,1] <- mean(dg1)
d[2,1] <- mean(dg2)

# compute hatsigmasquare
hatsigmasquare <- var(b) + rho * t(d) %*% s %*% d

normallow <- p - hatpauc - z * sqrt(hatsigmasquare) / sqrt(m)
normalup <- p - hatpauc + z * sqrt(hatsigmasquare) / sqrt(m)

# end of computing the normal approximation based C.I.

# compute the CI by using bootstrap method
B <- 150
starvm <- matrix(0, m, B)
lstar <- rep(0, B)
bootlist <- mybootstrap(B, xx, yy, m, n, p0, p1)
starpauc <- bootlist[[1]]
starssquare <- bootlist[[2]]
starvm <- bootlist[[3]]

starsigmasquare <- var(starpauc)
meanstarpauc <- mean(starpauc)

bootlow1 <- p - hatpauc - z * sqrt(starsigmasquare)
bootup1 <- p - hatpauc + z * sqrt(starsigmasquare)

bootlow2 <- p - meanstarpauc - z * sqrt(starsigmasquare)
bootup2 <- p - meanstarpauc + z * sqrt(starsigmasquare)

# end of computing the CI by using bootstrap method

# compute the EL-based semi-parametric CI: HBELI, HBELII

starsigmasquare <- var(starpauc)
meanstarssquare <- mean(starssquare)

hatssquare <- mean((hatv - hatpauc)^2)
starc <- hatssquare / (m * starsigmasquare)

# compute lambda and el bound for first (hatc)
# x[1]: lambda. x[2]: pauc

# compute lambda and el bound for first (starc)
\begin{verbatim}
f<-function(x) c(mean((hatv-x[2])/(1+x[1]*(hatv-x[2]))),starc*2*sum(log(abs(1+x[1]*(hatv-x[2]))))-qchisq(1-alpha,1))
sollow<-solveNonlinear(f,c(0,0),c(0,0.01))
solup<-solveNonlinear(f,c(0,0),c(0,0.95))
ellow1<-p-solup$x[2]
elup1<-p-sollow$x[2]

starc<-(meanstarssquare/(m*starsigmasquare))

# compute lambda and el bound for second (starc)
f<-function(x) c(mean((hatv-x[2])/(1+x[1]*(hatv-x[2]))),starc*2*sum(log(abs(1+x[1]*(hatv-x[2]))))-qchisq(1-alpha,1))
sollow<-solveNonlinear(f,c(0,0),c(0,0.01))
solup<-solveNonlinear(f,c(0,0),c(0,0.95))
ellow2<-p-solup$x[2]
elup2<-p-sollow$x[2]

# end of computing the EL-based semi-parametric CI

# compute the CI by using bootstrap method
for (b in 1:B)
{
wjhatb<-rep(0,m)
  for (i in 1:m)
  {
    wjhatb[i]<-starvm[i,b]-hatpauc
  }
  f<-function(lam)mean(wjhatb/(1+lam*wjhatb))
  lambda<-solveNonlinear(f,c(0),c(0.01))$x
  lstar[b]<-2*sum(log(abs(1+lambda*wjhatb)))
}
ctail<-1/mean(lstar)
cstar<-quantile(lstar, c(1-alpha))

# compute lambda and el bound for first (hatc)
# x[1]: lambda. x[2]: pauc

# compute lambda and el bound for third
f<-function(x) c(mean((hatv-x[2])/(1+x[1]*(hatv-x[2]))),2*sum(log(abs(1+x[1]*(hatv-x[2])))))-cstar)
sollow<-solveNonlinear(f,c(0,0),c(0,0.01))
solup<-solveNonlinear(f,c(0,0),c(0,0.95))
ellow3<-p-solup$x[2]
elup3<-p-sollow$x[2]

# compute lambda and el bound for fourth
f<-function(x) c(mean((hatv-x[2])/(1+x[1]*(hatv-x[2]))),ctail*2*sum(log(abs(1+x[1]*(hatv-x[2]))))-qchisq(1-alpha,1))
sollow<-solveNonlinear(f,c(0,0),c(0,0.01))
solup<-solveNonlinear(f,c(0,0),c(0,0.95))
ellow4<-p-solup$x[2]
elup4<-p-sollow$x[2]

# end of computing the EL-based semi-parametric CI
\end{verbatim}
# output
cat \{normal/\{p1-p0\}, normalup/\{p1-p0\}, bootlow1/\{p1-p0\}, bootup1/\{p1-p0\}, bootlow2/\{p1-p0\}, bootup2/\{p1-p0\}, elow1/\{p1-p0\}, elup1/\{p1-p0\}, elow2/\{p1-p0\}, elup2/\{p1-p0\}, elow3/\{p1-p0\}, elup3/\{p1-p0\}, elow4/\{p1-p0\}, elup4/\{p1-p0\}\}