5-13-2021

Kernel-based Empirical Likelihood Inference for the Area under the ROC Curve using Ranked Set Samples

Iyanuoluwa Ayodele

KERNEL-BASED EMPIRICAL LIKELIHOOD INFERENCE FOR THE AREA UNDER THE ROC CURVE USING RANKED SET SAMPLES

by

IYANUOLUWA AYODELE

Under the Direction of Yichuan Zhao, PhD

ABSTRACT

In diagnostic medicine, it is important to be able to accurately distinguish between a diseased and non-diseased population. The area under the curve (AUC) is a commonly used measure index to evaluate the accuracy of the diagnostic test. Sometimes in research, it is costly and time consuming to sample the variables of interest, ranked set samples (RSS) is a more effective sampling method than the simple random sampling which can be obtained by ranking, thereby providing samples which are representative of the population of interest, in balanced ranked set samples (BRSS), there is an equal number of cycles for each set. In this thesis, we propose the empirical likelihood and jackknife empirical likelihood methods using BRSS and multistage RSS for the AUC. The simulation results show that our proposed method improves on the estimation of AUC. We performed two real data analysis to illustrate the proposed methods.

INDEX WORDS: Area under the ROC curve, empirical likelihood, jackknife empirical likelihood, adjusted jackknife empirical likelihood

KERNEL-BASED EMPIRICAL LIKELIHOOD INFERENCE FOR THE AREA UNDER THE ROC CURVE USING RANKED SET SAMPLES

by

IYANUOLUWA AYODELE

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

2021

KERNEL-BASED EMPIRICAL LIKELIHOOD INFERENCE FOR THE AREA UNDER THE ROC CURVE USING RANKED SET SAMPLES

by

IYANUOLUWA AYODELE

Committee Chair: Yichuan Zhao

Committee: Jun Kong

Jing Zhang

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

May 2021

DEDICATION

To all dedicated Educators.

ACKNOWLEDGEMENTS

My gratitude goes to God Almighty for the divine strength, divine speed and knowledge he bestowed on me, because without him, this thesis would have remained incomplete. My sincere gratitude goes to my thesis advisor Professor Yichuan Zhao for his guidance, patience, and assistance while completing this thesis work. He was very helpful in providing materials that would aid better understanding of the key concepts in this thesis and always ready to answer any question I might have and proofreading this thesis very carefully to ensure I have the correct notations.

I would like to acknowledge the professors and staff of the Department of Mathematics and Statistics for the knowledge they instilled in me throughout my Masters degree and to the department and Georgia State University for granting me full tuition waiver and stipends for my Master's degree, this scholarship made it possible for me to be here today.

I would like to extend my thank you to Ali Jinnah, for his assitance with my R codes while completing this thesis. My sincere appreciation goes to Dr. Chul Moon for sharing the R codes for his paper with me, the codes were of immerse help.

I am grateful to my friend Lauren Drinkard, who has made my journey in graduate school an interesting one, she has been by my side since the first day I stepped foot in the department, for her words of encouragement, support, and assistance with this thesis. I could not have imagined going through graduate school without her.

I would like to acknowledge my father and sisters who have always been my support system, always cheering me on, believing that I can do whatever I set my mind to. Thank you to my friends and family who encourage and motivate me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS

- CI - Confidence interval
- EL - Empirical likelihood
- LLR - Log likelihood ratio
- JEL - Jackknife empirical likelihood
- AJEL - Adjusted jackknife empirical likelihood
- RSS - Ranked set samples
- BRSS - Balanced ranked set samples
- MSRSS - Multistage ranked set samples
- ROC - Receiver operating characteristics curve
- AUC - Area under the receiver operating characteristics curve

CHAPTER 1

INTRODUCTION

1.1 Introduction

In diagnostic medicine, it is important to be able to accurately distinguish a population as being diseased or non-diseased, which can be measured by the specificity and sensitivity. Let X and Y be continuous random variables denoting the non-diseased and diseased individuals respectively, and F and G be their corresponding distribution functions. The receiver operating characteristic (ROC) curve is a plot of the true positive rate (sensitivity), $p = P(Y \geq c) = 1 - G(c)$ against the false negative rate (1 - specificity), $q = 1 - P(X \leq c)$, allowing the diagnostic test to classify an individual as diseased if the measurement is greater than a predetermined threshold, denoted by c. The ROC curve is then given as

$$R(p) = 1 - G(F^{-1}(p)), 0 \leq p \leq 1.$$

The area under ROC curve (AUC) denoted as δ is a summary index of the ROC curve commonly used to evaluate the ability of a diagnostic test to distinguish a diseased one from a non-diseased one. Bamber (1975) established that the AUC is given as

$$\delta = P(Y \geq X) = \int_0^1 R(p)dp,$$

which is interpreted as the probability that the measurement of a randomly selected diseased one would be greater than or equal to that of the non-diseased one.

In the past, several studies have been conducted on the estimation of AUC using non-parametric approaches. An unbiased estimator of the AUC is the Mann-Whitney (MW) U statistic which can be used to construct the confidence interval for AUC using asymptotic normality (Bamber, 1975; Hanley and McNeil, 1982). Although the MW is asymptotically correct, the MW intervals result in a low coverage accuracy when the AUC is high (Qin and Zhou, 2006). The MW statistics is given as:

$$\hat{\delta} = \frac{1}{nx}\frac{1}{ny}\sum_{i=1}^{nx}\sum_{j=1}^{ny} I(Y_j \geq X_i).$$

Zou et al. (1997) and Lloyd (1998) proposed a kernel method which uses smoothed estimators rather than the empirical ones. Yin et al. (2016) evaluated the efficiency of the kernel estimator of AUC using ranked based sampling (RSS) instead of the simple random sample, however, the kernel estimates results in an under-estimation of the AUC, which usually does not include the true AUC value (Moon et al., 2020).

Qin and Zhou (2006) proposed an empirical likelihood (EL) approach for the inference of the AUC based on the MW estimator using simple random samples (SRS) and they determined that this approach is better than the previous methods. Recently, Moon et al. (2020) proposed an EL based approach to estimate the AUC using balanced and unbalanced ranked set samples and it was shown that the approach performs better than the SRS-EL even when the concomitant variable used for judgment ranking is poor.

Liu et al. (2009) proposed an EL approach for hypothesis testing and confidence interval estimation with balanced RSS and the results show that RSS is a more efficient sampling technique than the SRS. Also, Zhang et al. (2016) proposed a jackknife empirical likelihood approach to make inference for the population mean and the difference between two population mean using RSS, and the efficiency of JEL approach using RSS was determined over the EL approach using SRS.

The empirical likelihood is a nonparametric approach used to make inferences about the confidence region and was first introduced by Owen (1988,1990). The EL method is advantageous because it does not make assumptions about the distribution yet retains the asymptotic properties of the conventional likelihood method such as Wilk's theorem (Owen, 1990) and Bartlett correction (DiCiccio et al.,1991). The EL combines the effectiveness of the likelihood method with the reliability of the nonparametric method which makes it a powerful method. The coverage probabilities of the EL are frequently lower than the nominal level when the sample size is small (Owen (2001)). The EL method becomes increasingly difficult to compute as the sample size increases and this is why the jackknife empirical likelihood (JEL) was proposed.

The jackknife empirical likelihood (JEL) method was proposed by Jing et al. (2009), by employing a combination of the EL and jackknife approaches. Few of the advantages of the JEL method are the simplicity and the effectiveness when dealing with U-statistics (Jing et al. (2009)). In this thesis, we review BRSS-EL, then we propose BRSS-JEL, kernel-based empirical likelihood BRSS (KERNEL-EL), kernel-based jackknife empirical likelihood BRSS, (KERNEL-JEL), and kernel-based adjusted jackknife empirical likelihood BRSS (KERNEL-AJEL) based approaches to make inference on the AUC using BRSS approach. In addition, we propose kernel-based empirical likelihood MSRSS (KERNEL-MSRSS-EL), kernel-based jackknife empirical likelihood MSRSS (KERNEL-MRSS-JEL), empirical likelihood MSRSS (MSRSS-EL), and jackknife empirical likelihood MSRSS (MRSS-JEL) based approaches to make inference on the AUC using MSRSS approach.

The adjusted jackknife empirical likelihood (AJEL) method combines the advantages of the adjusted empirical likelihood (AEL) and EL and was proposed by Chen et al. (2008). AJEL improves on the JEL method especially when the sample size is small.

This thesis is organized as follows. In Chapter 2, we review the EL-BRSS and then we will propose the BRSS-JEL, KERNEL-EL, KERNEL-JEL, KERNEL-AJEL, KERNEL-MSRSS-EL, and KERNEL-MSRSS-JEL approaches to construct confidence intervals for δ. In Chapter 3, we conduct an extensive simulation study in terms of coverage probablity and average length of the confidence interval. In Chapter 4, we illustrate the proposed methods using two real data sets. Finally, in Chapter 5, we make a conclusion for the proposed methodology.

CHAPTER 2

METHODOLOGY

2.1 Ranked set sampling

Ranked set sampling (RSS) is a more effective sampling method that reduces cost when compared with SRS. RSS is useful when it is costly or time consuming to sample the variables of interest but observations can be easily obtained by ranking which in-turn provides a sample that is more representative of the population. Let X and Y be non-diseased and diseased variables generated by RSS respectively.

The RSS procedure is described as follows, a set size, say m is randomly generated and ranked using a concomitant variable used for judgement. The smallest of the sample is measured and the remaining $m-1$ samples are discarded, another SRS of size m is generated, ranked, and the second smallest is taken note of. This process is repeated m times while taking note of the i-th smallest sample where $i = 1, ..., m$. This entire process is known as a cycle. The cycle is repeated k times, $j = 1, ..., k$ which generates an RSS of size $n_x = mk$ and the corresponding sample is given as $X_{[i]j}$ which is the i-th non-diseased observation in the j-th cycle judged to be the the i-th smallest. The same procedure is applied to the diseased observations with set size n, $r = 1, ..., n$, repeated l times, $s = 1, ..., l$ with total RSS size of $n_y = nl$ and corresponding sample is given as $Y_{[r]s}$ which is the r-th diseased observation in the s-th cycle judged to be the the r-th smallest. Furthermore, this type of RSS is referred to a balanced ranked set sample because in each ranked statistics,there is an equal number of measurements.

2.2 Multiple stage ranked set sampling

Multistage ranked set sampling (MSRSS) was proposed by Al-Saleh and Al-omari (2002) and is an extension of the RSS sampling technique. MSRSS improves the efficiency of the mean estimator of the variable of interest with respect to the SRS and RSS. However, the technique can be more complicated than the RSS. Let X and Y be non-diseased and diseased variables generated by MSRSS respectively. The procedure is described as follows.

Let a be the number of stages, m^{a+1} sample units are randomly selected from the population of interest and allocated randomly into m^{a-1} sets with set size m^2 each, the RSS procedure described in Section 2.1 is applied to the m^{a-1} sets to obtain m^{a-1} sets of m set size each. The RSS procedure is continued until we end up with a sample of set size m. The process can be repeated k times and the corresponding sample is given as $X_{[i]j}^{(a)}$, $i = 1, ..., m$, $j = 1, ..., k$. The same procedure is applied to the diseased individuals with b number stages , set size n and number of cycles l to obtain a corresponding sample $Y_{[r]s}^{(b)}$, $r = 1, ..., n$, $s = 1, ..., l$. In the MSRSS procedure, only the last set size generated is quantified.

2.3 Empirical likelihood confidence interval for δ

In this section, we review EL inference method for the AUC using BRSS as proposed by Moon et al., (2020). For completeness, the full details of the proposed method is as follows. We adopt the same notations as Moon et al. (2020) did. Let

$$\phi(X_{[i]j}, Y_{[r]s}) = \begin{cases} 1, & X_{[i]j} < Y_{[r]s} \\ 0, & otherwise. \end{cases}$$

The MW statistic of the AUC based on BRSS is given as

$$\hat{\delta}_{BRSS} = U_{BRSS}(Y_{[1]1}, ..., Y_{[1]l}, ..., Y_{[n]1}, ..., Y_{[n]l}; X_{[1]1}, ..., X_{[1]k}, ..., X_{[m]1}, ..., X_{[m]k})$$

$$= \frac{1}{mknl} \sum_{i=1}^{m} \sum_{j=1}^{k} \sum_{r=1}^{n} \sum_{s=1}^{l} \phi(X_{[i]j}, Y_{[r]s}).$$

The placement value of $Y_{[r]s}$ can then be defined similar to Pepe and Cai (2004) as

$$U_{rs} = 1 - F(Y_{[r]s}).$$

We can easily show that

$$E\left[\sum_{r=1}^{n}\sum_{s=1}^{l}\frac{1}{nl}(1-U_{rs})\right]=E\left[F(Y)\right]=\delta.$$

EL procedure can be derived using the relationship between δ and U_{rs}. Let $p_{rs} = dG(Y_{[r]s})$ for $r = 1,...,n$ and $s = 1,...,l$. The EL for the AUC using balanced ranked set samples evaluated at AUC value δ can be given as

$$L(\delta)=\sup\left\{\prod_{r=1}^{n}\prod_{s=1}^{l}p_{rs}:\sum_{r=1}^{n}\sum_{s=1}^{l}p_{rs}=1,p_{rs}\geq 0,\sum_{r=1}^{n}\sum_{s=1}^{l}p_{rs}(1-U_{rs}-\delta)=0\right\}.$$

By replacing U_{rs} with $\hat{U}_{rs}$, we have

$$\hat{U}_{rs}=(1-\frac{1}{mk}\sum_{i=1}^{m}\sum_{j=1}^{k}\phi(X_{[i]j},Y_{[r]s}))$$

and solving the Lagrange multipliers, we obtain the empirical log-likelihood ratio which is given as

$$l(\delta)=2\sum_{r=1}^{n}\sum_{s=1}^{l}\log\left(1+\lambda(1-\hat{U}_{rs}-\delta)\right),$$

where λ is the solution to

$$\frac{1}{nl}\sum_{r=1}^{n}\sum_{s=1}^{l}\frac{(1-\hat{U}_{rs}-\delta)}{1+\lambda(1-\hat{U}_{rs}-\delta)}=0.$$

The maximum empirical likelihood estimator (MELE) $\tilde{\delta}_{BRSS}$ is given as

$$\tilde{\delta}_{BRSS}=\arg\underset{\delta}{min} l(\delta).$$

We set $\tilde{\delta}_{BRSS} = \hat{\delta}_{BRSS}$ because the dimension of $\hat{\delta}$ and the constraint are equivalent and thus $\hat{\delta}_{BRSS}$ is the solution of $\sum_{r=1}^{n}\sum_{s=1}^{l}(1-\hat{U}_{rs}-\delta)=0$ (Qin and Lawless, 1994). Since U_{rs}

are not independent, we cannot apply the standard EL theory to the empirical log-likelihood ratio. Instead, we study a limiting distribution of the scaled EL (Wang and Rao, 2002a,b; Wang et al., 2004). The theorem below shows an asymptotic distribution of $l(\delta_0)$ follows a scaled chi-square distribution.

Theorem 2.1 (cf. Moon et. al., 2020). Assume that ranking of BRSS are consistent, the true value of the AUC is δ_0 and $E(|F(Y)|^3) < \infty$. For fixed m and n as $k \to \infty$ and $l \to \infty$,

$$r(\delta_0)l(\delta_0) \to \chi_1^2,$$

where

$$r(\delta_0) = \frac{mk}{mk+nl} \frac{\sum_{r=1}^{n}\sum_{s=1}^{l} \frac{1}{nl}(1-\hat{U}_{rs}-\delta_0)^2}{S^2},$$

$$S^2 = \frac{nl(S^{10})^2 + mk(S^{01})^2}{mk+nl},$$

$$(S^{10})^2 = \sum_{i=1}^{m}\sum_{j=1}^{k} \frac{1}{m(k-1)}(V^{10}(X_{[i]j} - \bar{V}_{[i]}^{10})^2,$$

$$(S^{01})^2 = \sum_{r=1}^{n}\sum_{s=1}^{l} \frac{1}{n(l-1)}(V^{01}(Y_{[r]s} - \bar{V}_{[r]}^{10})^2,$$

$$V^{10}(X_{[i]j}) = \frac{1}{nl}\sum_{r=1}^{n}\sum_{s=1}^{l}\phi(X_{[i]j}, Y_{[r]s}), V^{01}(Y_{[r]s}) = \frac{1}{mk}\sum_{i=1}^{m}\sum_{j=1}^{k}\phi(X_{[i]j}, Y_{[r]s}), \bar{V}_{[i]}^{10} = \frac{1}{k}\sum_{j=1}^{k}V^{10}(X_{[i]j})$$

and $\bar{V}_{[r]}^{01} = \frac{1}{l}\sum_{s=1}^{l} V^{01}(Y_{[r]s})$.

Theorem 2.1 can then be used to obtain the confidence interval for δ_0. A BRSS-EL $100(1-\alpha)\%$ confidence interval for AUC can be found as

$$CI_\alpha(\delta) = \left\{\delta : r(\hat{\delta}_{BRSS})l(\delta) \le \chi_{1,1-\alpha}^2\right\},$$

where $\chi_{1,1-\alpha}^2$ is the $(1-\alpha)$-th quantile of the chi-square distribution with 1 degree of freedom.

2.4 Jackknife empirical likelihood confidence interval for δ

In this section, we propose the JEL inference method for the AUC using BRSS. The basic idea of the JEL is to use the jackknife pseudo-values to turn the variable of interest into a sample mean (Jing et al. (2009)).

Recall that the MW statistic for AUC using the BRSS is given as

$$\hat{\delta}_{BRSS} = \frac{1}{mknl}\sum_{i=1}^{m}\sum_{j=1}^{k}\sum_{r=1}^{n}\sum_{s=1}^{l}\phi(X_{[i]j}, Y_{[r]s})$$

$$= U_{BRSS}(Y_{[1]1}, ..., Y_{[1]l}, ..., Y_{[n]1}, ..., Y_{[n]l}; X_{[1]1}, ..., X_{[1]k}, ..., X_{[m]1}, ..., X_{[m]k}) = U_{BRSS}(Z_1, ..., Z_w).$$

where $i = 1, ..., w$, $w = nl + mk$,.

Let $\hat{V}_i$ be the jackknife pseudo-value that are asymptotically independent random variables which are given as

$$\hat{V}_i = w * (U_{BRSS}) - (w-1) * U_{BRSS}^{(-i)},$$

$$= \frac{w(w-1)}{w-2}\left[\frac{V_{i,0}}{mk}I\{1 \le i \le mk\} + \frac{V_{0,i-mk}}{nl}I\{mk+1 \le i \le w\}\right] - \frac{w}{w-2}U_{BRSS}, i = 1, ..., w,$$

and

$$E\hat{V}_i \approx \frac{w\delta}{(w-2)}\left[\frac{nl-1}{mk}I\{1 \le i \le mk\} + \frac{mk-1}{nl}I\{mk+1 \le i \le w\}\right],$$

where

$$V_{i,0} = mkU_{BRSS}^{0} - (mk-1)U_{BRSS}^{-[i]j,[0]0}, \qquad i = 1, ..., w,$$

$$V_{0,i-mk} = nlU_{BRSS}^{0} - (nl-1)U_{BRSS}^{[0]0,-[r]s}, \qquad i = 1, ..., w,$$

U_{BRSS}^{0} is the original AUC based on all observations,
$U_{BRSS}^{-[i]j,[0]0} = \frac{1}{(mk-1)(nl)}\sum_{i=1}^{m}\sum_{j=1}^{k}\sum_{r=1}^{n}\sum_{s=1}^{l}\phi(X_{[i]j}, Y_{[r]s})$ is the AUC, when $X_{[i]j}$ is left out, for $i = 1, ..., m, j = 1, ..., k$,

$U_{BRSS}^{[0]0,-[r]s} = \frac{1}{mk(nl-1)} \sum_{i=1}^{m} \sum_{j=1}^{k} \sum_{r=1}^{n} \sum_{s=1}^{l} \phi(X_{[i]j}, Y_{[r]s})$ is the AUC, when $Y_{[r]s}$ is left out, for $r = 1, ..., n, s = 1, ..., l,$

Then the jackknife estimator of the δ is given as

$$U_{BRSS,jel} = \frac{1}{w} \sum_{i=1}^{w} \hat{V}_{i,0} = \frac{1}{w} \sum_{i=1}^{l} \hat{V}_{0,i-mk}.$$

When the pseudo-values are applied, the JEL ratio at δ using BRSS is given as

$$R(\delta) = \sup\Big\{ \prod_{i=1}^{w} (wp_i) : \sum_{i=1}^{w} p_i = 1, p_i \geq 0, \sum_{i=1}^{w} p_i(\hat{V}_i(\delta) - E\hat{V}_i(\delta)) = 0 \Big\}.$$

Using the Lagrange multiplier, the jackknife empirical log-likelihood function at δ can then be given as

$$-2\text{log}R(\delta) = 2 \sum_{i=1}^{w} \log(1 + \lambda(\hat{V}_i(\delta) - E\hat{V}_i(\delta))),$$

where λ is the solution to the equation

$$0 = \frac{1}{w} \sum_{i=1}^{w} \frac{(\hat{V}_i(\delta) - E\hat{V}_i(\delta))}{1 + \lambda(\hat{V}_i(\delta) - E\hat{V}_i(\delta))}.$$

We establish the Wilk's theorem as follows.

Theorem 2.2 Assume that the true value of the AUC is δ_0. There is a consistent ranking of BRSS, and $E(|F(Y)|^3) < \infty$. For fixed m and n as $k \to \infty$ and $l \to \infty$,

$$-2\text{log}R(\delta_0) \to \chi_1^2.$$

Theorem 2.2 can then be used to obtain the confidence interval for δ_0. A BRSS-JEL $100(1 - \alpha)\%$ confidence interval for AUC can be found as

$$CI_\alpha(\delta) = \big\{\delta : -2\text{log}R(\delta) \leq \chi_{1,1-\alpha}^2\big\}.$$

2.5 Kernel-based empirical likelihood confidence interval for δ

In this section, we propose the kernel-based EL inference method for the AUC using BRSS. The kernel-based AUC estimator using the BRSS is given as

$$\hat{\delta}_{BRSS} = \frac{1}{mknl}\sum_{i=1}^{m}\sum_{j=1}^{k}\sum_{r=1}^{n}\sum_{s=1}^{l} K(X_{[i]j}, Y_{[r]s}),$$

where $K(X_{[i]j}, Y_{[r]s}) = \Phi\left((Y_{[r]s} - X_{[i]j})/\sqrt{h_x^2 + h_y^2}\right)$, Φ is the CDF of the standard normal, and $h_x = 0.9\min(s_x, iqr_x/1.34)n_x^{-0.2}$ and $h_y = 0.9\min(s_y, iqr_y/1.34)n_y^{-0.2}$ are the bandwidths where sd$(\cdot)$ is the sample standard deviation and iqr$(\cdot)$ is the inter quartile range, such bandwidths were recommended by Silverman (1986). The kernel-based EL ratio at δ using BRSS is given as

$$R(\delta) = \sup\left\{\prod_{r=1}^{n}\prod_{s=1}^{l}(nlp_{rs}) : \sum_{r=1}^{n}\sum_{s=1}^{l} p_{rs} = 1, p_{rs} \geq 0, \sum_{r=1}^{n}\sum_{s=1}^{l} p_{rs}(1 - \hat{U}_{rs} - \delta) = 0\right\}.$$

Using the Lagrange multiplier, the kernel-based empirical log-likelihood ratio is given as

$$l(\delta) = 2\sum_{r=1}^{n}\sum_{s=1}^{l}\log\left(1 + \lambda(1 - \hat{U}_{rs} - \delta)\right),$$

where $\hat{U}_{rs} = (1 - \frac{1}{mk}\sum_{i=1}^{m}\sum_{j=1}^{k} K(X_{[i]j}, Y_{[r]s}))$ and λ is the solution to

$$\frac{1}{nl}\sum_{r=1}^{n}\sum_{s=1}^{l}\frac{(1 - \hat{U}_{rs} - \delta)}{1 + \lambda(1 - \hat{U}_{rs} - \delta)} = 0.$$

The theorem below shows an asymptotic distribution of $l(\delta_0)$ follows a scaled chi-square distribution. We establish the Wilk's theorem as follows.

Theorem 2.3 Assume that the true value of the AUC is δ_0. There is a consistent ranking of BRSS, and $E(|F(Y)|^3) < \infty$. For fixed m and n as $k \to \infty$ and $l \to \infty$,

$$r(\delta_0)l(\delta_0) \to \chi_1^2,$$

where

$$r(\delta_0) = \frac{mk}{mk+nl} \frac{\sum_{r=1}^{n}\sum_{s=1}^{l} \frac{1}{nl}(1-\hat{U}_{rs}-\delta_0)^2}{S^2},$$

$$S^2 = \frac{nl(S^{10})^2 + mk(S^{01})^2}{mk+nl},$$

$$(S^{10})^2 = \sum_{i=1}^{m}\sum_{j=1}^{k} \frac{1}{m(k-1)}(V^{10}(X_{[i]j} - \bar{V}_{[i]}^{10})^2,$$

$$(S^{01})^2 = \sum_{r=1}^{n}\sum_{s=1}^{l} \frac{1}{n(l-1)}(V^{01}(Y_{[r]s} - \bar{V}_{[r]}^{10})^2,$$

$$V^{10}(X_{[i]j}) = \frac{1}{nl}\sum_{r=1}^{n}\sum_{s=1}^{l} K(X_{[i]j}, Y_{[r]s}), V^{01}(Y_{[r]s}) = \frac{1}{mk}\sum_{i=1}^{m}\sum_{j=1}^{k} K(X_{[i]j}, Y_{[r]s}), \bar{V}_{[i]}^{10} = \frac{1}{k}\sum_{j=1}^{k} V^{10}(X_{[i]j})$$

and $\bar{V}_{[r]}^{01} = \frac{1}{l}\sum_{s=1}^{l} V^{01}(Y_{[r]s})$.

Theorem 2.3 can then be used to obtain the confidence interval for δ_0. A KERNEL-EL $100(1-\alpha)\%$ confidence interval for AUC can be found as

$$CI_\alpha(\delta) = \left\{\delta : r(\hat{\delta})l(\delta) \le \chi^2_{1,1-\alpha}\right\}.$$

.

2.6 Kernel-based jackknife empirical likelihood confidence interval for δ

In this section, we propose the kernel-based JEL inference method for the AUC using BRSS. Recall that the kernel-based AUC estimator using the BRSS is given as

$$\hat{\delta}_{BRSS} = \frac{1}{mknl}\sum_{i=1}^{m}\sum_{j=1}^{k}\sum_{r=1}^{n}\sum_{s=1}^{l} K(X_{[i]j}, Y_{[r]s})$$

$$= U_{BRSS}(Y_{[1]1}, ..., Y_{[1]l}, ..., Y_{[n]1}, ..., Y_{[n]l}; X_{[1]1}, ..., X_{[1]k}, ..., X_{[m]1}, ..., X_{[m]k}) = U_{BRSS}(Z_1, ..., Z_w).$$

where $i = 1, ..., w,\ w = nl + mk,$.

Let $\hat{V}_i$ be the jackknife pseudo-value that are asymptotically independent random variables which is given as

$$\hat{V}_i = w * (U_{BRSS}) - (w-1) * U_{BRSS}^{(-i)},$$

$$= \frac{w(w-1)}{w-2}\left[\frac{V_{i,0}}{mk} I\{1 \leq i \leq mk\} + \frac{V_{0,i-mk}}{nl} I\{mk+1 \leq i \leq w\}\right] - \frac{w}{w-2} U_{BRSS}, i = 1, ..., w,$$

and

$$E\hat{V}_i \approx \frac{w\delta}{(w-2)}\left[\frac{nl-1}{mk} I\{1 \leq i \leq mk\} + \frac{mk-1}{nl} I\{mk+1 \leq i \leq w\}\right],$$

where

$$V_{i,0} = mkU^0_{BRSS} - (mk-1)U^{-[i]j,[0]0}_{BRSS}, \qquad i = 1, ..., w,$$

$$V_{0,i-mk} = nlU^0_{BRSS} - (nl-1)U^{[0]0,-[r]s}_{BRSS}, \qquad i = 1, ..., w,$$

U^0_{BRSS} is the original AUC based on all observations,
$U^{-[i]j,[0]0}_{BRSS} = \frac{1}{(mk-1)(nl)}\sum_{i=1}^{m}\sum_{j=1}^{k}\sum_{r=1}^{n}\sum_{s=1}^{l} K(X_{[i]j}, Y_{[r]s})$ is the AUC, when $X_{[i]j}$ is left out, for $i = 1, ..., m, j = 1, ..., k,$
$U^{[0]0,-[r]s}_{BRSS} = \frac{1}{mk(nl-1)}\sum_{i=1}^{m}\sum_{j=1}^{k}\sum_{r=1}^{n}\sum_{s=1}^{l} K(X_{[i]j}, Y_{[r]s})$ is the AUC, when $Y_{[r]s}$ is left out, for $r = 1, ..., n, s = 1, ..., l,$
Then the kernel-based jackknife estimator of the δ is given as

$$U_{KernelBRSS,jel} = \frac{1}{w}\sum_{i=1}^{w}\hat{V}_{i,0} = \frac{1}{w}\sum_{i=1}^{w}\hat{V}_{0,i-mk}.$$

When the pseudo-values are applied, the kernel-based JEL ratio at δ using BRSS is given as

$$R(\delta) = \sup\Big\{ \prod_{i=1}^{w}(wp_i) : \sum_{i=1}^{w} p_i = 1, p_i \geq 0, \sum_{i=1}^{w} p_i(\hat{V}_i(\delta) - E\hat{V}_i(\delta)) = 0 \Big\}.$$

Using the Lagrange multiplier, the kernel-based jackknife empirical log-likelihood function at δ can then be given as

$$-2\log R(\delta) = 2\sum_{i=1}^{w} \log(1 + \lambda(\hat{V}_i(\delta) - E\hat{V}_i(\delta)),$$

where λ is the solution to the equation

$$0 = \frac{1}{w}\sum_{i=1}^{w} \frac{(\hat{V}_i(\delta) - E\hat{V}_i(\delta))}{1 + \lambda(\hat{V}_i(\delta) - E\hat{V}_i(\delta))}.$$

We establish the Wilk's theorem as follows.

Theorem 2.4 Assume that the true value of the AUC is δ_0. There is a consistent ranking of BRSS, and $E(|F(Y)|^3) < \infty$. For fixed m and n as $k \to \infty$ and $l \to \infty$,

$$-2\log R(\delta_0) \to \chi_1^2.$$

Theorem 2.4 can then be used to obtain the confidence interval for δ_0. A KERNEL-JEL $100(1-\alpha)\%$ confidence interval for AUC can be found as

$$CI_\alpha(\delta) = \big\{\delta : -2\log R(\delta) \leq \chi^2_{1,1-\alpha}\big\}.$$

2.7 Kernel-based adjusted jackknife empirical likelihood confidence interval for δ

In this section, we propose the kernel-based AJEL inference method for the AUC using BRSS. The basic idea of the AJEL is to add an observation to the jackknife pseudo-values (Chen et al. (2008)). To simplify the notation, let

$$D_i(\delta) = \hat{V}_i(\delta) - E\hat{V}_i(\delta), i = 1, ..., w.$$

Also, let $D_{w+1}(\delta)$ be the observation added which is given as

$$D_{w+1}(\delta) = -\frac{a_w}{w}\sum_{i=1}^{w} D_i(\delta),$$

where $a_w = \max(1, \log(w)/2)$ and $a_w > 0$. When the new pseudo-values is added, the kernel-based AJEL ratio at δ using BRSS is given as

$$R^{adj}(\delta) = \sup\left\{\prod_{i=1}^{w+1}(w+1)p_i : \sum_{i=1}^{w+1} p_i = 1, p_i \geq 0, \sum_{i=1}^{w+1} p_i D_i(\delta) = 0\right\}.$$

Using the Lagrange multipliers, the kernel-based adjusted jackknife empirical log-likelihood function at δ can then be given as

$$-2\log R^{adj}(\delta) = 2\sum_{i=1}^{w+1}\log(1+\lambda D_i(\delta)),$$

where λ is the solution to the equation

$$0 = \frac{1}{w+1}\sum_{i=1}^{w+1}\frac{D_i(\delta)}{1+\lambda D_i(\delta)}.$$

The same conditions in Jing et al. (2009) are applied and we establish the Wilk's theorem as follows.

Theorem 2.5 Assume that the true value of the AUC is δ_0. There is a consistent ranking of BRSS, and $E(|F(Y)|^3) < \infty$. For fixed m and n as $k \to \infty$ and $l \to \infty$,

$$-2\log R^{adj}(\delta_0) \to \chi_1^2.$$

Theorem 2.5 can then be used to obtain the confidence interval for δ_0. A KERNEL-AJEL $100(1-\alpha)\%$ confidence interval for AUC can be found as

$$CI_\alpha(\delta)^{adj} = \left\{\delta : -2\log R^{adj}(\delta) \leq \chi_{1,1-\alpha}^2\right\}.$$

2.8 Kernel-based empirical likelihood confidence interval for δ using multi-stage RSS

In this section, we propose the kernel-based EL inference method for the AUC using MSRSS. The kernel-based AUC estimator using the MSRSS is given as

$$\hat{\delta} = \frac{1}{mknl} \sum_{i=1}^{m} \sum_{j=1}^{k} \sum_{r=1}^{n} \sum_{s=1}^{l} K(X_{[i]j}^{(a)}, Y_{[r]s}^{(b)}),$$

where $K(X_{[i]j}^{(a)}, Y_{[r]s}^{(b)}) = \Phi\left((Y_{[r]s}^{(b)} - X_{[i]j}^{(a)}) / \sqrt{h_x^2 + h_y^2} \right)$, Φ is the CDF of the standard normal, and $h_x = 0.9\text{min}(s_x, iqr_x/1.34)n_x^{-0.2}$ and $h_y = 0.9\text{min}(s_y, iqr_y/1.34)n_y^{-0.2}$ are the bandwidths where $sd(\cdot)$ is the sample standard deviation and $iqr(\cdot)$ is the inter quartile range, such bandwidths were recommended by Silverman (1986). The kernel-based EL ratio at δ using MSRSS is given as

$$R(\delta) = \sup\left\{ \prod_{r=1}^{n} \prod_{s=1}^{l} (nlp_{rs}) : \sum_{r=1}^{n} \sum_{s=1}^{l} p_{rs} = 1, p_{rs} \geq 0, \sum_{r=1}^{n} \sum_{s=1}^{l} p_{rs}(1 - \hat{U}_{rs} - \delta) = 0 \right\}.$$

Using the Lagrange multipliers, the kernel-based empirical log-likelihood ratio using MSRSS is given as

$$l(\delta) = 2 \sum_{r=1}^{n} \sum_{s=1}^{l} \log\left(1 + \lambda(1 - \hat{U}_{rs}^{(b)} - \delta)\right),$$

where $\hat{U}_{rs}^{(b)}(\delta) = (1 - \frac{1}{mk} \sum_{i=1}^{m} \sum_{j=1}^{k} K(X_{[i]j}^{(a)}, Y_{[r]s}^{(b)}))$ and λ is the solution to

$$\frac{1}{nl} \sum_{r=1}^{n} \sum_{s=1}^{l} \frac{(1 - \hat{U}_{rs}^{(b)} - \delta)}{1 + \lambda(1 - \hat{U}_{rs}^{(b)} - \delta)} = 0.$$

The theorem below shows an asymptotic distribution of $l(\delta_0)$ follows a scaled chi-square distribution. We establish the Wilk's theorem as follows.

Theorem 2.6 Assume that the true value of the AUC is δ_0. There is a consistent ranking of MSRSS, and $E(|F^{(b)}(Y)|^3) < \infty$. For fixed m and n as $k \to \infty$ and $l \to \infty$,

$$r(\delta_0)l(\delta_0) \to \chi_1^2,$$

where

$$r(\delta_0) = \frac{mk}{mk+nl} \frac{\sum_{r=1}^{n}\sum_{s=1}^{l} \frac{1}{nl}(1-\hat{U}_{rs}^{(b)}-\delta_0)^2}{S^2},$$

$$S^2 = \frac{nl(S^{10})^2 + mk(S^{01})^2}{mk+nl},$$

$$(S^{10})^2 = \sum_{i=1}^{m}\sum_{j=1}^{k} \frac{1}{m(k-1)}(V^{10}(X_{[i]j}^{(a)} - \bar{V}_{[i]}^{10})^2,$$

$$(S^{01})^2 = \sum_{r=1}^{n}\sum_{s=1}^{l} \frac{1}{n(l-1)}(V^{01}(Y_{[r]s}^{(b)} - \bar{V}_{[r]}^{10})^2,$$

$$V^{10}(X_{[i]j}) = \frac{1}{nl}\sum_{r=1}^{n}\sum_{s=1}^{l} K(X_{[i]j}^{(a)}, Y_{[r]s}^{(b)}), V^{01}(Y_{[r]s}) = \frac{1}{mk}\sum_{i=1}^{m}\sum_{j=1}^{k} K(X_{[i]j}^{(a)}, Y_{[r]s}^{(b)}), \bar{V}_{[i]}^{10} = \frac{1}{k}\sum_{j=1}^{k} V^{10}(X_{[i]j}^{(a)})$$

and $\bar{V}_{[r]}^{01} = \frac{1}{l}\sum_{s=1}^{l} V^{01}(Y_{[r]s}^{(b)})$.
Theorem 2.6 can then be used to obtain the confidence interval for δ_0. A KERNEL-MRSS-EL $100(1-\alpha)\%$ confidence interval for AUC using MSRSS can be found as

$$CI_\alpha(\delta) = \left\{\delta : r(\hat{\delta})l(\delta) \le \chi^2_{1,1-\alpha}\right\}.$$

2.9 Kernel-based jackknife empirical likelihood confidence interval for δ using multi-stage RSS

In this section, we propose the kernel-based JEL inference method for the AUC using MSRSS. Recall that the kernel-based AUC estimator using the MSRSS is given as

$$\hat{\delta}_{MSRSS} = \frac{1}{mknl}\sum_{i=1}^{m}\sum_{j=1}^{k}\sum_{r=1}^{n}\sum_{s=1}^{l} K(X_{[i]j}^{(a)}, Y_{[r]s}^{(b)})$$

$$= U_{MSRSS}(Y_{[1]1}^{(b)}, ..., Y_{[1]l}^{(b)}, ..., Y_{[n]1}^{(b)}, ..., Y_{[n]l}^{(b)}; X_{[1]1}^{(a)}, ..., X_{[1]k}^{(a)}, ..., X_{[m]1}^{(a)}, ..., X_{[m]k}^{(a)}) = U_{MSRSS}(Z_1, ..., Z_w).$$

where $i = 1, ..., w,\ w = nl + mk,.$

Let $\hat{V}_i$ be the jackknife pseudo-value that are asymptotically independent random variables which is given as

$$\hat{V}_i = w * (U_{MSRSS}) - (w-1) * U_{MSRSS}^{(-i)},$$

$$= \frac{w(w-1)}{w-2}\left[\frac{V_{i,0}}{mk}I\{1 \le i \le mk\} + \frac{V_{0,i-mk}}{nl}I\{mk+1 \le i \le w\}\right] - \frac{w}{w-2}U_{MSRSS}, i = 1, ..., w,$$

and

$$E\hat{V}_i \approx \frac{w\delta}{(w-2)}\left[\frac{nl-1}{mk}I\{1 \le i \le mk\} + \frac{mk-1}{nl}I\{mk+1 \le i \le w\}\right],$$

where

$$V_{i,0} = mkU_{BRSS}^{0} - (mk-1)U_{BRSS}^{-[i]j,[0]0}, \qquad i = 1, ..., w,$$

$$V_{0,i-mk} = nlU_{BRSS}^{0} - (nl-1)U_{BRSS}^{[0]0,-[r]s}, \qquad i = 1, ..., w,$$

U_{BRSS}^{0} is the original AUC based on all observations,
$U_{BRSS}^{-[i]j,[0]0} = \frac{1}{(mk-1)(nl)}\sum_{i=1}^{m}\sum_{j=1}^{k}\sum_{r=1}^{n}\sum_{s=1}^{l} K(X_{[i]j}^{(a)}, Y_{[r]s}^{(b)})$ is the AUC when $X_{[i]j}^{(a)}$ is left out, for $i = 1, ..., m, j = 1, ..., k,$
$U_{BRSS}^{[0]0,-[r]s} = \frac{1}{mk(nl-1)}\sum_{i=1}^{m}\sum_{j=1}^{k}\sum_{r=1}^{n}\sum_{s=1}^{l} K(X_{[i]j}^{(a)}, Y_{[r]s}^{(b)})$ is the AUC when $Y_{[r]s}^{(b)}$ is left out, for $r = 1, ..., n, s = 1, ..., l,$

Then the kernel-based jackknife estimator of the δ is given as

$$U_{KernelMSRSS,jel} = \frac{1}{w}\sum_{i=1}^{w}\hat{V}_{i,0} = \frac{1}{w}\sum_{i=1}^{w}\hat{V}_{0,i-mk}.$$

When the pseudo-values are applied, the kernel-based JEL ratio at δ using MSRSS is given as

$$R(\delta) = \sup\left\{ \prod_{i=1}^{w}(wp_i) : \sum_{i=1}^{w} p_i = 1, p_i \geq 0, \sum_{i=1}^{w} p_i(\hat{V}_i(\delta) - E\hat{V}_i(\delta)) = 0 \right\}.$$

Using the Lagrange multiplier, the kernel-based jackknife empirical log-likelihood function at δ using MSRSS can then be given as

$$-2\log R(\delta) = 2\sum_{i=1}^{w} \log(1 + \lambda(\hat{V}_i(\delta) - E\hat{V}_i(\delta)),$$

where λ is the solution to the equation

$$0 = \frac{1}{w}\sum_{i=1}^{w} \frac{(\hat{V}_i(\delta) - E\hat{V}_i(\delta))}{1 + \lambda(\hat{V}_i(\delta) - E\hat{V}_i(\delta))}.$$

We establish the Wilk's theorem as follows.

Theorem 2.7 Assume that the true value of the AUC is δ_0. There is a consistent ranking of MSRSS, and $E(|F^{(b)}(Y)|^3) < \infty$. For fixed m and n as $k \to \infty$ and $l \to \infty$,

$$-2\log R(\delta_0) \to \chi_1^2.$$

Theorem 2.7 can then be used to obtain the confidence interval for δ_0. A KERNEL-MSRSS-JEL $100(1-\alpha)\%$ confidence interval for AUC using MSRSS can be found as

$$CI_\alpha(\delta) = \left\{\delta : -2\log R(\delta) \leq \chi^2_{1,1-\alpha}\right\}.$$

2.10 Empirical likelihood confidence interval for δ using multi-stage RSS

In this section, we propose the EL inference method for the AUC using MSRSS. The AUC estimator using the MSRSS is given as

$$\hat{\delta} = \frac{1}{mknl}\sum_{i=1}^{m}\sum_{j=1}^{k}\sum_{r=1}^{n}\sum_{s=1}^{l} \phi(X_{[i]j}^{(a)}, Y_{[r]s}^{(b)}),$$

The EL ratio for δ using MSRSS is given as

$$R(\delta) = \sup\Big\{ \prod_{r=1}^{n}\prod_{s=1}^{l}(nlp_{rs}) : \sum_{r=1}^{n}\sum_{s=1}^{l} p_{rs} = 1, p_{rs} \geq 0, \sum_{r=1}^{n}\sum_{s=1}^{l} p_{rs}(1 - \hat{U}_{rs} - \delta) = 0 \Big\}.$$

Using the Lagrange multipliers, the empirical log-likelihood ratio using MSRSS is given as

$$l(\delta) = 2\sum_{r=1}^{n}\sum_{s=1}^{l} \log\Big(1 + \lambda(1 - \hat{U}_{rs}^{(b)} - \delta)\Big),$$

where $\hat{U}_{rs}^{(b)}(\delta) = (1 - \frac{1}{mk}\sum_{i=1}^{m}\sum_{j=1}^{k}\phi(X_{[i]j}^{(a)}, Y_{[r]s}^{(b)}))$ and λ is the solution to

$$\frac{1}{nl}\sum_{r=1}^{n}\sum_{s=1}^{l}\frac{(1 - \hat{U}_{rs}^{(b)} - \delta)}{1 + \lambda(1 - \hat{U}_{rs}^{(b)} - \delta)} = 0.$$

The theorem below shows an asymptotic distribution of $l(\delta_0)$ follows a scaled chi-square distribution. We establish the Wilk's theorem as follows.

Theorem 2.8 Assume that the true value of the AUC is δ_0. There is a consistent ranking of MSRSS, and $E(|F^{(b)}(Y)|^3) < \infty$. For fixed m and n as $k \to \infty$ and $l \to \infty$,

$$r(\delta_0)l(\delta_0) \to \chi_1^2,$$

where

$$r(\delta_0) = \frac{mk}{mk + nl}\frac{\sum_{r=1}^{n}\sum_{s=1}^{l}\frac{1}{nl}(1 - \hat{U}_{rs}^{(b)} - \delta_0)^2}{S^2},$$

$$S^2 = \frac{nl(S^{10})^2 + mk(S^{01})^2}{mk + nl},$$

$$(S^{10})^2 = \sum_{i=1}^{m}\sum_{j=1}^{k}\frac{1}{m(k-1)}(V^{10}(X_{[i]j}^{(a)} - \bar{V}_{[i]}^{10})^2,$$

$$(S^{01})^2 = \sum_{r=1}^{n}\sum_{s=1}^{l} \frac{1}{n(l-1)}(V^{01}(Y_{[r]s}^{(b)} - \bar{V}_{[r]}^{10})^2,$$

$$V^{10}(X_{[i]j}) = \frac{1}{nl}\sum_{r=1}^{n}\sum_{s=1}^{l}\phi(X_{[i]j}^{(a)}, Y_{[r]s}^{(b)}), V^{01}(Y_{[r]s}) = \frac{1}{mk}\sum_{i=1}^{m}\sum_{j=1}^{k}\phi(X_{[i]j}^{(a)}, Y_{[r]s}^{(b)}), \bar{V}_{[i]}^{10} = \frac{1}{k}\sum_{j=1}^{k} V^{10}(X_{[i]j}^{(a)})$$

and $\bar{V}_{[r]}^{01} = \frac{1}{l}\sum_{s=1}^{l} V^{01}(Y_{[r]s}^{(b)})$.
Theorem 2.8 can then be used to obtain the confidence interval for δ_0. A MRSS-EL $100(1-\alpha)\%$ confidence interval for AUC using MSRSS can be found as

$$CI_\alpha(\delta) = \left\{\delta : r(\hat{\delta})l(\delta) \le \chi^2_{1,1-\alpha}\right\}.$$

2.11 Jackknife empirical likelihood confidence interval for δ using multi-stage RSS

In this section, we propose the JEL inference method for the AUC using MSRSS. Recall that the AUC estimator using the MSRSS is given as

$$\hat{\delta}_{MSRSS} = \frac{1}{mknl}\sum_{i=1}^{m}\sum_{j=1}^{k}\sum_{r=1}^{n}\sum_{s=1}^{l}\phi(X_{[i]j}^{(a)}, Y_{[r]s}^{(b)})$$

$$= U_{MSRSS}(Y_{[1]1}^{(b)}, ..., Y_{[1]l}^{(b)}, ..., Y_{[n]1}^{(b)}, ..., Y_{[n]l}^{(b)}; X_{[1]1}^{(a)}, ..., X_{[1]k}^{(a)}, ..., X_{[m]1}^{(a)}, ..., X_{[m]k}^{(a)}) = U_{MSRSS}(Z_1, ..., Z_w).$$

where $i = 1, ..., w$, $w = nl + mk$,.

Let $\hat{V}_i$ be the jackknife pseudo-value that are asymptotically independent random variables which is given as

$$\hat{V}_i = w * (U_{MSRSS}) - (w-1) * U_{MSRSS}^{(-i)},$$

$$= \frac{w(w-1)}{w-2}\left[\frac{V_{i,0}}{mk} I\{1 \le i \le mk\} + \frac{V_{0,i-mk}}{nl} I\{mk+1 \le i \le w\}\right] - \frac{w}{w-2} U_{MSRSS}, i = 1, ..., w,$$

and

$$E\hat{V}_i \approx \frac{w\delta}{(w-2)}\left[\frac{nl-1}{mk}I\{1 \leq i \leq mk\} + \frac{mk-1}{nl}I\{mk+1 \leq i \leq w\}\right],$$

where

$$V_{i,0} = mkU^0_{BRSS} - (mk-1)U^{-[i]j,[0]0}_{BRSS}, \qquad i = 1, ..., w,$$

$$V_{0,i-mk} = nlU^0_{BRSS} - (nl-1)U^{[0]0,-[r]s}_{BRSS}, \qquad i = 1, ..., w,$$

U^0_{BRSS} is the original AUC based on all observations,
$U^{-[i]j,[0]0}_{BRSS} = \frac{1}{(mk-1)(nl)}\sum_{i=1}^{m}\sum_{j=1}^{k}\sum_{r=1}^{n}\sum_{s=1}^{l}\phi(X^{(a)}_{[i]j}, Y^{(b)}_{[r]s})$ is the AUC when $X^{(a)}_{[i]j}$ is left out, for $i = 1, ..., m, j = 1, ..., k$,
$U^{[0]0,-[r]s}_{BRSS} = \frac{1}{mk(nl-1)}\sum_{i=1}^{m}\sum_{j=1}^{k}\sum_{r=1}^{n}\sum_{s=1}^{l}\phi(X^{(a)}_{[i]j}, Y^{(b)}_{[r]s})$ is the AUC when $Y^{(b)}_{[r]s}$ is left out, for $r = 1, ..., n, s = 1, ..., l$,
Then the kernel-based jackknife estimator of the δ is given as

$$U_{KernelMSRSS,jel} = \frac{1}{w}\sum_{i=1}^{w}\hat{V}_{i,0} = \frac{1}{w}\sum_{i=1}^{w}\hat{V}_{0,i-mk}.$$

When the pseudo-values are applied, the JEL ratio at δ using MSRSS is given as

$$R(\delta) = \sup\left\{\prod_{i=1}^{w}(wp_i) : \sum_{i=1}^{w}p_i = 1, p_i \geq 0, \sum_{i=1}^{w}p_i(\hat{V}_i(\delta) - E\hat{V}_i(\delta)) = 0\right\}.$$

Using the Lagrange multiplier, the jackknife empirical log-likelihood function at δ using MSRSS can then be given as

$$-2\log R(\delta) = 2\sum_{i=1}^{w}\log(1 + \lambda(\hat{V}_i(\delta) - E\hat{V}_i(\delta)),$$

where λ is the solution to the equation

$$0 = \frac{1}{w}\sum_{i=1}^{w}\frac{(\hat{V}_i(\delta) - E\hat{V}_i(\delta))}{1 + \lambda(\hat{V}_i(\delta) - E\hat{V}_i(\delta))}.$$

We establish the Wilk's theorem as follows.

Theorem 2.9 Assume that the true value of the AUC is δ_0. There is a consistent ranking of MSRSS, and $E(|F^{(b)}(Y)|^3) < \infty$. For fixed m and n as $k \to \infty$ and $l \to \infty$,

$$-2\log R(\delta_0) \to \chi_1^2.$$

Theorem 2.9 can then be used to obtain the confidence interval for δ_0. A MSRSS-JEL $100(1-\alpha)\%$ confidence interval for AUC using MSRSS can be found as

$$CI_\alpha(\delta) = \left\{\delta : -2\log R(\delta) \leq \chi^2_{1,1-\alpha}\right\}.$$

CHAPTER 3

SIMULATION STUDY

In this chapter, we report the performance of the proposed methods; BRSS-JEL,KERNEL-EL, KERNEL-JEL, KERNEL-AJEL, MRSS-EL, MRSS-JEL, KERNEL-MSRSS-EL, KERNEL-MSRSS-JEL and compare with BRSS-EL of (Moon et al., 2020), and KERNEL-BRSS of Yin et al. (2016) using the average length and the coverage probabilities of the confidence interval. We run simulation for sample sizes $n_x = n_y$= $20, 40, 80$, and with δ $0.8, 0.9, 0.95$. We run simulations 5000 times for each combination of the sample size pairs and δ referred to in the sentence above.

For the Kernel-based approaches, the Gaussian kernel would be applied with bandwidths; $h_x = 0.9\min(s_x, iqr_x/1.34)n_x^{-0.2}$ and $h_y = 0.9\min(s_y, iqr_y/1.34)n_y^{-0.2}$ where $sd(\cdot)$ is the sample standard deviation and $iqr(\cdot)$ is the inter quartile range.

The chosen set sizes are $n = m = 2$, the number of cycles is given as $k = n_x/m, l = n_y/n$ and for the MSRSS, the number of stages are $a = b$: $2, 3$. The concomitant variables used for the judgement ranking are related to the samples and given as:

$$C_x = \rho_x \left(\frac{X - \mu_x}{\sigma_x} \right) + (1 - \rho_x^2) Z_x$$

$$C_y = \rho_y \left(\frac{Y - \mu_y}{\sigma_y} \right) + (1 - \rho_y^2) Z_y,$$

where $\mu(\cdot)$ denotes the mean, $\sigma(\cdot)$ denotes the standard deviation, $Z(\cdot)$ follows a standard normal distribution, and $\rho(\cdot)$ is the Pearson correlation coefficient representing accuracy of the rankings and is set as $\rho_x = \rho_y$= 0.9 and 1, which signifies good and perfect judgment respectively.

For this simulation study, three distributions; Normal, Uniform distribution, and Log-normal will be used. In the first simulation, X is generated from a standard normal distribution $N(0,1)$ and Y follows a normal distribution $N(\sqrt{5}\Phi^{-1}(\delta),4)$, where Φ is the CDF of a standard normal distribution. For the second simulation, X is generated from a standard uniform distribution $U(0,1)$ and Y follows a uniform distribution $U\left(0,\frac{1}{2(1-\delta)}\right)$. For the third simulation, X is generated from a standard log-normal distribution $LN(0,1)$ and Y follows a log-normal distribution $LN(\sqrt{5}\Phi^{-1}(\delta),4)$. The nominal confidence level $1-\alpha$ is 0.95. From Tables 3.1, 3.2, and 3.2, which show the simulation studies, we can conclude that.

1) As the AUC value increases, the coverage probability and average length of the confidence interval decreases and as the sample size increases, the average length of the confidence intervals decreases.

2) The kernel-based methods tend to have shorter average lengths than their corresponding methods based on the Mann-Whitney statistic.

3) Although KERNEL-BRSS has the shortest average lengths, the coverage probability of the confidence intervals is lower than all of the other methods.

4) The JEL methods have longer average length and better coverage probability of the confidence intervals than the EL methods and the AJEL method also have longer average length and better coverage probability of the confidence intervals than the JEL methods.

5) The two-stage KERNEL-MSRSS-EL has shorter average lengths than the KERNEL-EL and the three-stage KERNEL-MSRSS-EL tends to have shorter average lengths than the KERNEL-EL in some cases. The kernel-based JEL approaches for the MSRSS tend to have comparable average length of the confidence intervals with KERNEL-JEL.

6) The two-stage MRSS-EL approaches has shorter average lengths of the confidence intervals than the three-stage MRSS-EL.

7) The MSRSS-EL have longer average length and better coverage probability of the confidence interval than the KERNEL-MSRSS-EL. However, the KERNEL-MRSS-JEL

have longer average length and better coverage probability of the confidence interval than the MSRSS-EL in most cases.

8) As the sample size increases, the coverage probability of the confidence interval tends to be more stable and be closer to the nominal level.

9) For the Log-normal distribution, the coverage probability of the confidence interval reduces drastically as the AUC value increases.

Table (3.1) Normal Distribution: 95% Coverage Probabilities and Average Interval Lengths

ρ	(nx, ny)	AUC	KERNEL-BRSS	BRSS-EL	KERNEL-EL	BRSS-JEL	KERNEL-JEL	KERNEL-AJEL
1	(20, 20)	0.80	.876 (.193)	.942 (.250)	.912 (.223)	.972 (.295)	.967 (.279)	.983 (.300)
		0.90	.854 (.143)	.906 (.186)	.864 (.170)	.931 (.209)	.938 (.211)	.968 (.227)
		0.95	.853 (.102)	.864 (.129)	.825 (.122)	.883 (.141)	.922 (.153)	.956 (.165)
	(40, 40)	0.80	.876 (.138)	.949 (.177)	.914 (.161)	.976 (.205)	.966 (.197)	.978 (.205)
		0.90	.848 (.103)	.939 (.133)	.875 (.124)	.959 (.147)	.949 (.148)	.966 (.155)
		0.95	.838 (.073)	.901 (.093)	.848 (.090)	.919 (.099)	.927 (.106)	.949 (.110)
	(80, 80)	0.80	.871 (.099)	.950 (.124)	.910 (.116)	.977 (.144)	.961 (.139)	.971 (.143)
		0.90	.851 (.074)	.954 (.094)	.875 (.089)	.966 (.103)	.941 (.104)	.952 (.106)
		0.95	.839 (.052)	.934 (.067)	.864 (.065)	.945 (.071)	.928 (.074)	.940 (.076)
0.9	(20, 20)	0.80	.860 (.191)	.938 (.257)	.916 (.231)	.967 (.295)	.960 (.279)	.982 (.300)
		0.90	.841 (.141)	.906 (.189)	.875 (.175)	.923 (.209)	.931 (.211)	.960 (.227)
		0.95	.837 (.101)	.865 (.132)	.822 (.124)	.881 (.141)	.918 (.153)	.951 (.165)
	(40, 40)	0.80	.862 (.137)	.952 (.182)	.917 (.166)	.974 (.205)	.959 (.197)	.972 (.205)
		0.90	.847 (.101)	.943 (.135)	.886 (.127)	.953 (.147)	.943 (.149)	.960 (.155)
		0.95	.839 (.072)	.910 (.095)	.855 (.092)	.912 (.099)	.923 (.106)	.942 (.111)
	(80, 80)	0.80	.854 (.099)	.951 (.128)	.916 (.120)	.970 (.144)	.954 (.140)	.967 (.142)
		0.90	.836 (.073)	.950 (.095)	.884 (.090)	.962 (.103)	.941 (.104)	.950 (.106)
		0.95	.812 (.051)	.934 (.067)	.860 (.066)	.945 (.071)	.926 (.074)	.939 (.076)

Table (3.2) Uniform Distribution: 95% Coverage Probabilities and Average Interval Lengths

ρ	(nx, ny)	AUC	KERNEL-BRSS	BRSS-EL	KERNEL-EL	BRSS-JEL	KERNEL-JEL	KERNEL-AJEL
1	(20, 20)	0.80	.888 (.187)	.943 (.249)	.927 (.211)	.972 (.296)	.975 (.264)	.990 (.284)
		0.90	.729 (.131)	.913 (.196)	.866 (.158)	.934 (.934)	.953 (.201)	.980 (.215)
		0.95	.664 (.104)	.901 (.154)	.685 (.123)	.919 (.168)	.914 (.170)	.961 (.182)
	(40, 40)	0.80	.895 (.135)	.949 (.176)	.935 (.155)	.978 (.206)	.976 (.188)	.987 (.196)
		0.90	.828 (.095)	.945 (.140)	.871 (.117)	.959 (.155)	.955 (.140)	.972 (.145)
		0.95	.610 (.071)	.916 (.106)	.656 (.089)	.920 (.112)	.899 (.110)	.934 (.115)
	(80, 80)	0.80	.901 (.098)	.953 (.125)	.933 (.113)	.979 (.145)	.976 (.134)	.980 (.137)
		0.90	.812 (.069)	.948 (.100)	.873 (.085)	.966 (.109)	.956 (.099)	.967 (.101)
		0.95	.520 (.050)	.943 (.075)	.601 (.064)	.946 (.079)	.871 (.075)	.895 (.077)
0.9	(20, 20)	0.80	.888 (.185)	.953 (.256)	.936 (.217)	.970 (.295)	.970 (.264)	.985 (.283)
		0.90	.833 (.131)	.941 (.143)	.864 (.162)	.931 (.221)	.950 (.200)	.977 (.215)
		0.95	.675 (.104)	.899 (.156)	.686 (.126)	.920 (.168)	.909 (.170)	.963 (.182)
	(40, 40)	0.80	.882 (.134)	.954 (.181)	.938 (.160)	.973 (.206)	.970 (.188)	.981 (.196)
		0.90	.813 (.094)	.941 (.143)	.953 (.147)	.958 (.155)	.960 (.135)	.966 (.146)
		0.95	.612 (.070)	.908 (.105)	.666 (.091)	.921 (.112)	.896 (.110)	.932 (.115)
	(80, 80)	0.80	.888 (.097)	.951 (.128)	.936 (.116)	.974 (.144)	.973 (.134)	.978 (.137)
		0.90	.801 (.068)	.947 (.101)	.881 (.087)	.964 (.109)	.953 (.099)	.963 (.101)
		0.95	.519 (.050)	.937 (.076)	.660 (.065)	.947 (.079)	.864 (.075)	.887 (.077)

Table (3.3) Log-normal Distribution: 95% Coverage Probabilities and Average Interval Lengths

ρ	(nx, ny)	AUC	KERNEL-BRSS	BRSS-EL	KERNEL-EL	BRSS-JEL	KERNEL-JEL	KERNEL-AJEL
1	(20, 20)	0.80	.636 (.149)	.945 (.251)	.847 (.201)	.968 (.278)	.852 (.246)	.898 (.264)
		0.90	.051 (.131)	.907 (.187)	.484 (.158)	.945 (.193)	.672 (.209)	.730 (.228)
		0.95	.004 (.125)	.866 (.128)	.204 (.129)	.912 (.126)	.554 (.192)	.618 (.214)
	(40, 40)	0.80	.433 (.105)	.950 (.177)	.771 (.146)	.981 (.194)	.818 (.175)	.837 (.182)
		0.90	.005 (.092)	.943 (.134)	.303 (.114)	.968 (.135)	.578 (.147)	.610 (.157)
		0.95	.000 (.087)	.901 (.093)	.073 (.093)	.946 (.089)	.432 (.132)	.492 (.147)
	(80, 80)	0.80	.236 (.075)	.949 (.125)	.674 (.105)	.981 (.136)	.775 (.125)	.767 (.128)
		0.90	.000 (.065)	.951 (.094)	.156 (.082)	.970 (.094)	.479 (.102)	.503 (.108)
		0.95	.000(.061)	.932 (.067)	.012 (.066)	.957 (.063)	.277(.091)	.384 (.099)
0.9	(20, 20)	0.80	.636 (.149)	.945 (.251)	.888 (.222)	.961 (.278)	.847 (.247)	.899 (.265)
		0.90	.051 (.131)	.907 (.187)	.556 (.174)	.933 (.192)	.683 (.210)	.735 (.229)
		0.95	.004 (.125)	.866 (.128)	.236 (.143)	.901 (.126)	.554 (.192)	.634 (.215)
	(40, 40)	0.80	.433 (.105)	.950 (.177)	.818 (.160)	.971 (.194)	.818 (.175)	.817 (.182)
		0.90	.005 (.092)	.943 (.134)	.369 (.125)	.960 (.135)	.575 (.146)	.596 (.157)
		0.95	.000 (.087)	.901 (.093)	.090 (.102)	.936 (.089)	.443 (.133)	.488 (.146)
	(80, 80)	0.80	.765 (.102)	.946 (.150)	.744 (.115)	.965 (.136)	.758 (.124)	.749 (.127)
		0.90	.000 (.065)	.951 (.094)	.215 (.089)	.960 (.094)	.473 (.102)	.496 (.107)
		0.95	.000 (.061)	.932 (.067)	.025 (.072)	.951 (.063)	.456 (.091)	.477 (.099)

Table (3.4) Normal Distribution: 95% Coverage Probabilities and Average Interval Lengths for Multistage Ranked Set Sampling

			Two-stage RSS				Three-stage RSS			
ρ	(nx, ny)	AUC	MSRSS-EL	KERNEL-MSRSS-EL	MSRSS-JEL	KERNEL-MSRSS-JEL	MSRSS-EL	KERNEL-MSRSS-EL	MSRSS-JEL	KERNEL-MSRSS-JEL
1	(20, 20)	0.80	.945 (.229)	.894 (.197)	.982 (.296)	.982 (.280)	.943 (.251)	.910 (.223)	.971 (.295)	.963 (.279)
		0.90	.911 (.180)	.852 (.163)	.939 (.210)	.954 (.211)	.914 (.187)	.863 (.170)	.933 (.209)	.939 (.211)
		0.95	.865 (.126)	.815 (.117)	.890 (.141)	.919 (.150)	.874 (.128)	.823 (.122)	.887 (.140)	.911 (.150)
	(40, 40)	0.80	.949 (.161)	.894 (.143)	.985 (.206)	.971 (.203)	.951 (.177)	.909 (.161)	.977 (.206)	.969 (.197)
		0.90	.941 (.128)	.864 (.116)	.967 (.147)	.956 (.157)	.940 (.133)	.885 (.125)	.955 (.147)	.948 (.149)
		0.95	.912 (.091)	.844 (.087)	.923 (.100)	.935 (.115)	.906 (.093)	.852 (.090)	.913 (.100)	.926 (.106)
	(80, 80)	0.80	.947 (.113)	.890 (.103)	.986 (.144)	.976 (.139)	.955 (.952)	.903 (.116)	.981 (.144)	.968 (.140)
		0.90	.947 (.090)	.875 (.084)	.969 (.103)	.959 (.106)	.952 (.094)	.877 (.089)	.968 (.103)	.949 (.104)
		0.95	.933 (.065)	.859 (.063)	.947 (.070)	.935 (.074)	.936 (.067)	.856 (.065)	.944 (.071)	.929 (.075)
0.9	(20, 20)	0.80	.945 (.231)	.897 (.198)	.982 (.296)	.982 (.280)	.945 (.251)	.910 (.223)	.972 (.295)	.969 (.279)
		0.90	.913 (.180)	.861 (.160)	.942 (.210)	.950 (.212)	.914 (.186)	.865 (.170)	.932 (.208)	.939 (.210)
		0.95	.864 (.127)	.817 (.118)	.887 (.141)	.919 (.150)	.863 (.128)	.825 (.121)	.880 (.139)	.906 (.149)
	(40, 40)	0.80	.948 (.162)	.894 (.144)	.986 (.206)	.977 (.197)	.952 (.177)	.917 (.161)	.978 (.206)	.966 (.197)
		0.90	.941 (.128)	.868 (.117)	.965 (.147)	.953 (.149)	.941 (.133)	.883 (.124)	.957 (.147)	.946 (.149)
		0.95	.911 (.092)	.844 (.087)	.922 (.010)	.930 (.115)	.907 (.093)	.848 (.091)	.914 (.099)	.927 (.106)
	(80, 80)	0.80	.949 (.114)	.892 (.104)	.985 (.114)	.976 (.139)	.958 (.125)	.911 (.115)	.982 (.144)	.968 (.140)
		0.90	.948 (.090)	.875 (.084)	.971 (.103)	.945 (.106)	.957 (.094)	.874 (.089)	.970 (.103)	.948 (.104)
		0.95	.937 (.065)	.856 (.063)	.948 (.071)	.933 (.078)	.941 (.067)	.855 (.060)	.950 (.071)	.932 (.074)

Table (3.5) Uniform Distribution: 95% Coverage Probabilities and Average Interval Lengths for Multistage Ranked Set Sampling

			Two-stage RSS				Three-stage RSS			
ρ	(nx, ny)	AUC	MSRSS-EL	KERNEL-MSRSS-EL	MSRSS-JEL	KERNEL-MSRSS-JEL	MSRSS-EL	KERNEL-MSRSS-EL	MSRSS-JEL	KERNEL-MSRSS-JEL
1	(20, 20)	0.80	.936 (.225)	.912 (.183)	.982 (.297)	.986 (.265)	.940 (.248)	.929 (.210)	.969 (.295)	.971 (.264)
		0.90	.909 (.191)	.831 (.147)	.935 (.221)	.958 (.200)	.910 (.195)	.857 (.157)	.931 (.219)	.942 (.198)
		0.95	.906 (.154)	.639 (.117)	.926 (.169)	.916 (.161)	.889 (.153)	.689 (.122)	.910 (.166)	.894 (.158)
	(40, 40)	0.80	.951 (.159)	.924 (.135)	.987 (.206)	.988 (.188)	.949 (.176)	.931 (.155)	.974 (.206)	.974 (.188)
		0.90	.941 (.136)	.851 (.109)	.965 (.155)	.966 (.140)	.934 (.140)	.872 (.124)	.953 (.154)	.954 (.139)
		0.95	.911 (.104)	.626 (.086)	.926 (.112)	.901 (.110)	.911 (.105)	.653 (.089)	.926 (.112)	.896 (.109)
	(80, 80)	0.80	.946 (.112)	.926 (.099)	.986 (.145)	.988 (.136)	.951 (.124)	.933 (.113)	.975 (.144)	.972 (.134)
		0.90	.946 (.097)	.850 (.080)	.968 (.109)	.964 (.100)	.943 (.099)	.867 (.085)	.962 (.108)	.948 (.099)
		0.95	.942 (.075)	.560 (.062)	.954 (.079)	.879 (.077)	.933 (.075)	.602 (.064)	.944 (.079)	.876 (.075)
0.9	(20, 20)	0.80	.938 (.225)	.910 (.185)	.982 (.297)	.987 (.265)	.942 (.248)	.931 (.210)	.970 (.294)	.974 (.265)
		0.90	.912 (.192)	.831 (.147)	.938 (.223)	.954 (.200)	.908 (.196)	.860 (.157)	.928 (.219)	.941 (.198)
		0.95	.906 (.154)	.637 (.118)	.924 (.170)	.912 (.161)	.896 (.154)	.674 (.123)	.916 (.167)	.900 (.158)
	(40, 40)	0.80	.948 (.160)	.922 (.137)	.988 (.206)	.988 (.188)	.942 (.176)	.923 (.155)	.971 (.206)	.969 (.187)
		0.90	.937 (.136)	.851 (.109)	.963 (.155)	.962 (.140)	.938 (.140)	.869 (.116)	.956 (.154)	.950 (.139)
		0.95	.911 (.104)	.626 (.086)	.924 (.112)	.901 (.109)	.914 (.105)	.651 (.090)	.927 (.112)	.893 (.109)
	(80, 80)	0.80	.952 (.113)	.925 (.100)	.985 (.145)	.988 (.136)	.946 (.124)	.940 (.113)	.975 (.144)	.974 (.134)
		0.90	.949 (.097)	.850 (.080)	.972 (.109)	.965 (.100)	.949 (.099)	.872 (.085)	.967 (.108)	.957 (.098)
		0.95	.940 (.074)	.565 (.062)	.954 (.079)	.880 (.077)	.935 (.075)	.592 (.064)	.947 (.079)	.875 (.075)

Table (3.6) Log-normal Distribution: 95% Coverage Probabilities and Average Interval Lengths for Multistage Ranked Set Sampling

			Two-stage RSS				Three-stage RSS			
ρ	(nx,ny)	AUC	MSRSS-EL	KERNEL-MSRSS-EL	MSRSS-JEL	KERNEL-MSRSS-JEL	MSRSS-EL	KERNEL-MSRSS-EL	MSRSS-JEL	KERNEL-MSRSS-JEL
1	(20, 20)	0.80	.951 (.213)	.759 (.172)	.986 (.280)	.873 (.244)	.948 (.239)	.843 (.202)	.973 (.279)	.859 (.246)
		0.90	.931 (.162)	.580 (.145)	.963 (.194)	.655 (.208)	.928 (.172)	.481 (.158)	.949 (.193)	.671 (.209)
		0.95	.895 (.112)	.011 (.115)	.932 (.127)	.535 (.192)	.890 (.115)	.200 (.129)	.921 (.127)	.556 (.191)
	(40, 40)	0.80	.952 (.148)	.674 (.126)	.991 (.194)	.819 (.174)	.957 (.168)	.764 (.146)	.978 (.194)	.818 (.175)
		0.90	.946 (.155)	.234 (.100)	.976 (.135)	.568 (.146)	.950 (.124)	.313 (.114)	.968 (.135)	.567 (.146)
		0.95	.929 (.082)	.058 (.082)	.950 (.089)	.431 (.132)	.930 (.085)	.070 (.093)	.943 (.089)	.424 (.133)
	(80, 80)	0.80	.945 (.103)	.580 (.091)	.988 (.136)	.787 (.124)	.959 (.117)	.674 (.106)	.981 (.136)	.764 (.125)
		0.90	.946 (.080)	.127 (.073)	.976 (.094)	.490 (.103)	.954 (.086)	.168 (.082)	.969 (.095)	.469 (.102)
		0.95	.942 (.058)	.011(.059)	.959 (.062)	.263 (.090)	.945 (.060)	.014 (.066)	.960 (.063)	.287 (.091)
0.9	(20, 20)	0.80	.946 (.217)	.780 (.178)	.989 (.280)	.875 (.245)	.948 (.239)	.841 (.202)	.975 (.279)	.855 (.246)
		0.90	.927 (.164)	.401 (.142)	.962 (.194)	.663 (.209)	.921 (.172)	.481 (.158)	.943 (.192)	.675 (.209)
		0.95	.898 (.114)	.164 (.116)	.933 (.127)	.585 (.193)	.888 (.115)	.204 (.129)	.918 (.127)	.567 (.193)
	(40, 40)	0.80	.951 (.152)	.697 (.130)	.990 (.194)	.821 (.175)	.953 (.168)	.765 (.146)	.976 (.194)	.802 (.175)
		0.90	.944 (.117)	.245 (.104)	.976 (.135)	.574 (.146)	.946 (.124)	.312 (.114)	.965 (.135)	.578 (.146)
		0.95	.927 (.082)	.058 (.085)	.950 (.089)	.419 (.133)	.929 (.085)	.080 (.093)	.944 (.089)	.430 (.133)
	(80, 80)	0.80	.949 (.106)	.606 (.094)	.985 (.136)	.663 (.125)	.958 (.117)	.673 (.106)	.979 (.136)	.769 (.129)
		0.90	.947 (.081)	.139 (.075)	.974 (.094)	.487 (.102)	.952 (.086)	.166 (.082)	.971 (.095)	.468 (.103)
		0.95	.941 (.058)	.012 (.061)	.961 (.062)	.440 (.090)	.941 (.060)	.014 (.066)	.960 (.063)	.289 (.091)

CHAPTER 4

REAL DATA ANALYSIS

In this chapter, we would assess the performance of the proposed methodologies using two real data sets and make conclusions based on the analysis. The confidence intervals and interval lengths for the AUC would be used to demonstrate these application using KERNEL-BRSS, EL-BRSS, KERNEL-EL, KERNEL-JEL, KERNEL-AJEL, MRSS-EL, MRSS-JEL, KERNEL-MSRSS-EL, and KERNEL-MRSS-JEL.

4.1 Application using Diabetes data

For our first real data analysis application, we would be using the National Health and Nutrition Examination Survey (NHANES) for the 2009 to 2012 sample years. The NHANES is a survey data which was collected by the US National Center for Health Statistics (NCHS) and contains 10,000 observations and 75 variables which includes demographic variables, physical measurements, health variables, and lifestyle variables.

For the purpose of this thesis, we would focus on the two groups; those who have diabetes (y) and those who do not have Diabetes (x). The body mass index (BMI) would be used as the predictor for diabetes with an estimated AUC value of 0.73 and the concomitant variable chosen is weight which has a strong Pearson correlation coefficient with BMI ($\rho_x = 0.901, \rho_y = 0.881$). Our chosen set size is $m = n = 2$ and sample sizes are $n_x = n_y = 20, 40, 80$, the number of cycle is determined by $k = nx/m, l = ny/n$. The 95% confidence intervals and interval lengths for the AUC using BRSS and MRSS summarized in Table 4.1 indicates that the Two-stage KERNEL-MSRSS-EL has the shortest lengths, the kernel-based methods have shorter confidence lengths than the MW statistic methods, the JEL methods have longer interval lengths than the EL methods, and all other results are consistent with the simulation study.

Table (4.1) 95% Confidence Intervals and Interval Lengths using the Diabetes data

	(20,20)			(40,40)			(80,80)		
Method	Length	LB	UB	Length	LB	UB	Length	LB	UB
KERNEL-BRSS	0.2100	0.5842	0.7942	0.1553	0.5993	0.7546	0.1069	0.6475	0.7544
BRSS-EL	0.3073	0.5425	0.8498	0.2051	0.5852	0.7904	0.1402	0.6379	0.7781
KERNEL-EL	0.2635	0.5556	0.8190	0.1907	0.5758	0.7664	0.1295	0.6333	0.7628
BRSS-JEL	0.3279	0.5328	0.8607	0.2335	0.5721	0.8056	0.1579	0.6303	0.7882
KERNEL-JEL	0.2844	0.5361	0.8205	0.2190	0.5657	0.7847	0.1459	0.6139	0.7598
KERNEL-AJEL	0.3305	0.4807	0.8112	0.2392	0.5369	0.7762	0.1558	0.5974	0.7532
Two-stage MSRSS-EL	0.2248	0.6491	0.8739	0.1471	0.6768	0.8239	0.1121	0.6389	0.7510
Two-stage KERNEL-MSRSS-EL	0.2043	0.6308	0.8352	0.1388	0.6629	0.8017	0.1024	0.6332	0.7356
Two-stage MSRSS-JEL	0.2934	0.6161	0.9095	0.2108	0.6442	0.8551	0.1606	0.6146	0.7752
Two-stage KERNEL-MSRSS-JEL	0.2806	0.6287	0.9093	0.1938	0.6629	0.8017	0.1470	0.6137	0.7607
Three-stage MSRSS-EL	0.2876	0.4989	0.7866	0.1653	0.7044	0.8697	0.1264	0.6927	0.8190
Three-stage KERNEL-MSRSS-EL	0.2436	0.5289	0.7725	0.1530	0.6951	0.8480	0.1201	0.6758	0.7959
Three-stage MSRSS-JEL	0.3458	0.4693	0.8151	0.1969	0.6933	0.8901	0.1463	0.6839	0.8302
Three-stage KERNEL-MSRSS-JEL	0.2924	0.5126	0.8050	0.1921	0.6742	0.8663	0.1386	0.6738	0.8124

4.2 Application using Heart disease data

For our second real data analysis application, we would be using the coronary heart disease (CHD) data from the Framingham Heart Study. The study began in 1948 and is under the direction of the National Heart, Lung, and Blood Institute (NHLBI). When the study began, 5,209 women and men from the town of Framingham between the ages of 30 and 62 who have not shown any symptom of cardiovascular disease or suffered from stroke or heart attack were recruited and several offspring cohort studies have been added since then. The dataset contains risk factors for CHD such as systolic blood pressure (SBP), diastolic blood pressure (DBP), cholesterol, body mass index (BMI), glucose level, heart rate, etc.

For the purpose of this thesis, we would focus on the two groups; those who have 10 year risk of coronary heart disease (y) and those who do not have 10 year risk of coronary heart disease (x). The SBP would be used as the predictor for diabetes with an estimated AUC value of 0.62 and the concomitant variable chosen is DBP whose Pearson correlation coefficient with SBP is ($\rho_x = 0.785, \rho_y = 0.760$). Our chosen set set is $m = n = 2$ and sample sizes are $n_x = n_y = 20, 40, 80$, the number of cycle is determined by $k = nx/m, l = ny/n$. The 95% confidence intervals and interval lengths for the AUC using BRSS and MRSS summarized in Table 4.2 indicates that the KERNEL-BRSS has the shortest interval lengths, the kernel-based methods have shorter interval lengths than the MW statistic methods, and all other results are consistent with the simulation study.

Table (4.2) 95% Confidence Intervals and Interval Lengths using the CHD data

	(20,20)			(40,40)			(80,80)		
Method	Length	LB	UB	Length	LB	UB	Length	LB	UB
KERNEL-BRSS	0.2234	0.5156	0.7390	0.1620	0.5982	0.7602	0.1159	0.5394	0.6553
BRSS-EL	0.3157	0.4692	0.7849	0.2021	0.5836	0.7856	0.1582	0.5255	0.6837
KERNEL-EL	0.2776	0.4866	0.7642	0.1897	0.5748	0.7645	0.1454	0.5236	0.6690
BRSS-JEL	0.3524	0.4547	0.8071	0.2342	0.5725	0.8066	0.1757	0.5176	0.6933
KERNEL-JEL	0.3061	0.4747	0.7808	0.2346	0.5725	0.8066	0.1620	0.5141	0.6761
KERNEL-AJEL	0.3487	0.4244	0.7731	0.2562	0.5478	0.8039	0.1698	0.5009	0.6707
Two-stage MSRSS-EL	0.2645	0.4972	0.7617	0.1753	0.5267	0.7020	0.1695	0.5841	0.7202
Two-stage KERNEL-MSRSS-EL	0.2461	0.5059	0.7520	0.1661	0.5231	0.6892	0.1286	0.5749	0.7035
Two-stage MSRSS-JEL	0.3504	0.4553	0.8057	0.2476	0.4904	0.7380	0.1695	0.5685	0.7381
Two-stage KERNEL-MSRSS-JEL	0.3150	0.4680	0.7830	0.2328	0.4923	0.7250	0.1626	0.5631	0.7258
Three-stage MSRSS-EL	0.2848	0.4565	0.7413	0.1956	0.5128	0.7085	0.1388	0.5908	0.7296
Three-stage KERNEL-MSRSS-EL	0.2535	0.4745	0.7281	0.1747	0.5189	0.6936	0.1309	0.5805	0.7114
Three-stage MSRSS-JEL	0.3640	0.4173	0.7812	0.2505	0.4856	0.7362	0.1679	0.5771	0.7450
Three-stage KERNEL-MSRSS-JEL	0.3286	0.4244	0.7731	0.2254	0.4951	0.7205	0.1596	0.5698	0.7293

CHAPTER 5

CONCLUSIONS

In this thesis, we focused on the estimation of the area under the ROC curve (AUC) using ranked set samples which is an important measure in diagnostic medicine. We proposed EL, JEL, and AJEL methods to obtain the average length of the confidence interval and coverage probability for AUC and compared with some methods previously proposed.

We conducted simulation studies under Normal, Uniform, and Log-normal distribution and assessed the coverage and average lengths of the 95% confidence interval. From the results, we concluded that kernel-based methods tend to give shorter lengths than those calculated using the Mann-Whitney statistic. The AJEL confidence intervals have longer average lengths and better coverage probability of the confidence interval than JEL, while JEL confidence intervals have longer average lengths and better coverage probability of the confidence intervals than the EL methods. The MSRSS-EL have longer average length and better coverage probability of the confidence interval than the KERNEL-MSRSS-EL. However, the KERNEL-MRSS-JEL have longer average length and better coverage probability of the confidence interval than the MSRSS-EL in most cases.

Furthermore, we illustrated the methods using two real data sets with the help of the confidence intervals and confidence interval lengths and we confirmed that in some cases, the two-stage KERNEL-MSRSS-EL has shorter interval lengths than the KERNEL-BRSS. The KERNEL methods have shorter interval lengths than the methods based on the MW statistic and the results were consistent with the results from the simulation studies. In the future, multistage ranked set sampling technique can be applied to the Mann-Whitney statistics for estimating the AUC using some empirical methods to study the effect of increasing the stages of RSS on the statistic especially when the AUC value is close to 1.

REFERENCES

"About the Framingham Heart Study," Framingham Heart Study, Boston University and the National Heart, Lung, and Blood Institute, https://framinghamheartstudy.org/fhs-about/, Accessed 15 April 2021.

Al-Saleh, M.F. and Al-Kadiri, M. (2000), "Double ranked set sampling," Statistics and Probability Letters 48, 205–212.

Al-Saleh, M.F.and Al-Omari, A.I. (2002), "Multistage ranked set sampling," Journal of Statistical Planning and Inference, 102, 273–286

Al-Saleh, M. F. and Samuh, M. H. (2008), "On multistage ranked set sampling for distribution and median estimation," Computational Statistics and Data Analysis, 52, 2066–2078.

Bamber, D. (1975), "The area above the ordinal dominance graph and the area below the receiver operating characteristic graph," Journal of Mathematical Psychology, 12, 387–415.

Berger, Y. G. and De La Riva Torres, O. (2016), "Empirical likelihood confidence intervals for complex sampling designs," Journal of the Royal Statistical Society: Series B (Statistical Methodology), 78, 319–341.

Bohn, L. L. and Wolfe, D. A. (1992), "Nonparametric two-sample procedures for ranked set samples data," Journal of the American Statistical Association, 87, 552–561.

Chang, J., Tang, C. Y., and Wu, T. T. (2018), "A new scope of penalized empirical likelihood with high-dimensional estimating equations," Annals of Statistics, 46, 3185–3216.

Chen, J., Variyath, A., and Abraham, B. (2008). "Adjusted empirical likelihood and its properties," Journal of Computational and Graphical Statistics, 17:426-443.

Chen, Z., Bai, Z., and Sinha, B. K. (2004), "Ranked Set Sampling: Theory and Applications," Springer, New York, NY.

Cheng, C., Liu, Y., Liu, Z., and Zhou, Z. (2018). "Balanced augmented jackknife empirical likelihood for two sample U-statistics," Science China Mathematics 61, 1129–1138.

Devroye, L.P. and Wagner, T. (1979), "The L1 convergence of kernel density estimates," Annals of Statistics, 7, 1136–1139.

DiCiccio, T. J., Hall, P., and Romano, J. P. (1991), "Empirical likelihood is Bartlett correctable," Annals of Statistics, 19, 1053–1061.

Dümbgen, L. and Zamanzade, E. (2020), "Inference on a distribution function from ranked set samples," Annals of the Institute of Statistical Mathematics, 72, 157–185.

Frey, J. (2019), "Bounds on the efficiency of unbalanced ranked-set sampling," Communications in Statistics - Theory and Methods, 49, 1–14.

Frey, J. and Zhang, Y. (2017), "Testing perfect rankings in ranked-set sampling with binary data," Canadian Journal of Statistics, 45, 326–339.

Frey, J. and Zhang, Y. (2019), "An omnibus two-sample test for ranked-set sampling data," Journal of the Korean Statistical Society, 48, 106 – 116.

Ghosh, S., Chatterjee, A., and Balakrishnan, N. (2017), "Nonparametric confidence intervals for ranked set samples," Computational Statistics, 32, 1689–1725.

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

Hatefi, A., Jozani, M. J., and Ozturk, O. (2015), "Mixture model analysis of partially rank-ordered set samples: age groups of Fish from length-frequency Data," Scandinavian Journal of Statistics, 42, 848–871.

Hatefi, A., Reid, N., and Jozani, M. J. (2020), "Finite mixture modeling, classification and statistical learning with order statistics," Statistica Sinica, 30, 1881–1903.

Kvam, P. H. and Samaniego, F. J. (1994), “Nonparametric maximum likelihood estimation based on ranked set samples,” Journal of the American Statistical Association, 89, 526–537.

Li, T., Balakrishnan, N., Ng, H. K. T., Lu, Y., and An, L. (2019), “Precedence tests for equality of two distributions based on early failures of ranked set samples,” Journal of Statistical Computation and Simulation, 89, 2328–2353.

Liu, T., Lin, N., and Zhang, B. (2009), ”Empirical likelihood for balanced ranked-set sampled data,” Science in China, 6, 1351-1364.

Lloyd, C. J. (1998), “Using smoothed receiver operating characteristic curves to summarize and compare diagnostic systems,” Journal of the American Statistical Association, 93, 1356–1364.

Lloyd, C. J. and Yong, Z. (1999), “Kernel estimators of the ROC curve are better than empirical,” Statistics and Probability Letters, 44, 221-228.

Lora L. Bohn and Douglas A. Wolfe (1994), “The effect of imperfect judgment rankings on properties of procedures based on the ranked-set samples analog of the Mann-Whitney-Wilcoxon statistic,” Journal of the American Statistical Association, 89, 168–176.

Mahdizadeh, M. and Zamanzade, E. (2016), “Kernel-based estimation of $\mathrm{P}(X > Y)$ in ranked set sampling,” SORT, 40, 243–266.

Mahdizadeh, M. and Zamanzade, E. (2017a), ”Estimation of a symmetric distribution function in multistage ranked set sampling,” Statistical Paper, 61, 851-867.

Mahdizadeh, M. and Zamanzade, E. (2018), “Smooth estimation of the area under the ROC curve in multistage ranked set sampling,” Statistical Papers, 52, 750-768.

Moon, C., Wang, W., and Lim, J. (2020), ”Empirical likelihood inference for area under the ROC curve using ranked set samples” arXiv:2010.12185.

Omidvar, S., Jafari Jozani, M., and Nematollahi, N. (2018), "Judgment post-stratification in finite mixture modeling: An example in estimating the prevalence of osteoporosis," Statistics in Medicine, 37, 4823–4836.

Owen, A. B. (1988), "Empirical likelihood ratio confidence intervals for a single functional," Biometrika, 75, 237–249.

Owen, A. (1990), "Empirical likelihood ratio confidence regions," Annals of Statistics, 18, 90–120.

Owen, A. (1991), "Empirical likelihood for linear models," Annals of Statistics, 19, 1725–1747.

Ozturk, O. (2002), "Rank regression in ranked-set samples," Journal of the American Statistical Association, 97, 1180–1191.

Ozturk, O. (2012), "Quantile inference based on partially rank-ordered set samples," Journal of Statistical Planning and Inference, 142, 2116 – 2127.

Ozturk, O. (2018), "Statistical inference using rank-based post-stratified samples in a finite population," TEST, 28, 1113–1143.

Ozturk, O. (2019), "Two-stage cluster samples with ranked set sampling designs," Annals of the Institute of Statistical Mathematics, 71, 63–91.

Pepe, M. S. and Cai, T. (2004), "The analysis of placement values for evaluating discriminatory measures," Biometrics, 60, 528–535.

Presnell, B. and Bohn, L. L. (1999), "U-Statistics and imperfect ranking in ranked set sampling," Journal of Nonparametric Statistics, 10, 111–126.

Pruim, R. (2015), NHANES: Data from the US National Health and Nutrition Examination Study, R package version 2.1.0.

Qin, G. and Zhou, X.-H. (2006), "Empirical likelihood inference for the area under the ROC Curve," Biometrics, 62, 613–622.

Qin, J. and Lawless, J. (1994), "Empirical likelihood and general estimating equations," Annals of Statistics, 22, 300–325

Qin, J., Zhang, B., and Leung, D. H. Y. (2009), "Empirical likelihood in missing data problems," Journal of the American Statistical Association, 104, 1492–1503.

Sengupta, S. and Mukhuti, S. (2008), "Unbiased estimation of P$(X > Y)$ using ranked set sample data," Statistics, 42, 223–230.

Silverman, B. W. (1986), Density Estimation for Statistics and Data Analysis, Chapman and Hall/CRC, London.

Stokes, S. L. and Sager, T. W. (1988), "Characterization of a ranked-set sample with application to estimating distribution functions," Journal of the American Statistical Association, 83, 374–381.

Wang, D., Wu, T. T., and Zhao, Y. (2019), "Penalized empirical likelihood for the sparse Cox regression model," Journal of Statistical Planning and Inference, 201, 71 – 85.

Wang, Q., Linton, O., and Härdle, W. (2004), "Semi-parametric regression analysis with missing response at random," Journal of the American Statistical Association, 99, 334–345.

Wang, Q. and Rao, J. (2002a), "Empirical likelihood-based inference in linear errors-in-covariables models with validation data," Biometrika, 89, 345–358.

— (2002b), "Empirical likelihood-based inference under imputation for missing response data," Annals of Statistics, 30, 896–924.

Wang, X., Lim, J., and Stokes, L. (2016), "Using ranked set sampling with cluster randomized designs for improved inference on treatment effects," Journal of the American Statistical Association, 111, 1576–1590.

Wang, X., Wang, M., Lim, J., and Ahn, S. (2020), “Using ranked set sampling with binary outcomes in cluster randomized designs,” Canadian Journal of Statistics, 48, 342–365.

Wolfe, D. A. (2012), “Ranked set sampling: its relevance and impact on statistical inference,” ISRN Probability and Statistics, 2012, 1–32.

Yang, H. and Zhao, Y. (2012), “Smoothed empirical likelihood for ROC curves with censored data,” Journal of Multivariate Analysis, 109, 254 – 263.

Yin, J., Hao, Y., Samawi, H., and Rochani, H. (2016), “Rank-based kernel estimation of the area under the ROC curve,” Statistical Methodology, 32, 91 – 106.

Zamanzade, E. and Mahdizadeh, M. (2020), “Using ranked set sampling with extreme ranks in estimating the population proportion,” Statistical Methods in Medical Research, 29, 165–177.

Zhang, S., Han, P., and Wu, C. (2019), “Empirical likelihood inference for non-randomized pretest-posttest studies with missing data,” Electronic Journal of Statistics, 13, 2012–2042.

Zhang, Z., Liu, T., and Zhang, B. (2016), ”Jackknife empirical likelihood inference for the population mean with ranked set samples,” Statistics and Probability Letters, 108, 16-22.

Zhao, P., Ghosh, M., Rao, J. N. K., and Wu, C. (2020), “Bayesian empirical likelihood inference with complex survey data,” Journal of the Royal Statistical Society: Series B (Statistical Methodology), 82, 155–174.

Zou, K. H., Hall, W. J., and Shapiro, D. E. (1997), “Smooth non-parametric receiver operating characteristic (ROC) curves for continuous diagnostic tests,” Statistics in Medicine, 16, 2143–2156.