DOKUZ EYLÜL UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

ON THE ESTIMATION OF POPULATION MEAN BASED ON MODIFIED RANKED SET SAMPLING

by
Büşra SEVİNÇ

August, 2016
İZMİR

ON THE ESTIMATION OF POPULATION MEAN BASED ON MODIFIED RANKED SET SAMPLING

A Thesis Submitted to the
Graduate School of Natural and Applied Sciences of Dokuz Eylül University
In Partial Fulfillment of the Requirements for the Master of Science in
Statistics, Statistics Program

by
Büşra SEVİNÇ

August, 2016
İZMİR

M.Sc THESIS EXAMINATION RESULT FORM

We have read the thesis entitled **"ON THE ESTIMATION OF POPULATION MEAN BASED ON MODIFIED RANKED SET SAMPLING"** completed by **BÜŞRA SEVİNÇ** under supervision of **ASSOC. PROF. DR. SELMA GÜRLER** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

.....

Assoc. Prof. Dr. Selma GÜRLER

Supervisor

Assoc. Prof. Dr. Neslihan Demirel

(Jury Member)

Assoc. Prof. Dr. Sevcan Demir Atalay

(Jury Member)

Prof. Dr.Ayşe OKUR

Director

Graduate School of Natural and Applied Sciences

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude and thanks to my advisor, Assoc. Prof. Dr. Selma GÜRLER for her support, direction, encouragement that she gave to me, positivity and for being cheerful all the time. I am honored to be a student of her.

This thesis is supported by Dokuz Eylül University BAP Project Grant No: 2016.KB.FEN.015. Also, I am indebted to Scientific and Technological Research Council of Turkey (TÜBİTAK-COST, Grant No: 115F300) for their support.

I would like to thank to Bekir ÇETİNTAV for his helpful and constructive comments. I also wish to thank to all of academic staff in Dokuz Eylül University Statistics Department especially, Beste BEYAZTAŞ and Ufuk BEYAZTAŞ for their support and helps.

Finally, special thanks to Lüveyza SEVİNÇ, my mother, Azmi SEVİNÇ, my father and Ayşenur SEVİNÇ, my sister for their endless support and love during my studies and thanks to Almila HACIOĞLU and Özge LİMONCU for their motiviation and supporting me in every way when I need.

Büşra SEVİNÇ

ON THE ESTIMATION OF POPULATION MEAN BASED ON MODIFIED RANKED SET SAMPLING

ABSTRACT

Ranked set sampling (RSS) is an efficient and advantageous method for estimating the population mean when the measurement of an observation is time consuming and/or expensive. RSS procedure is based upon ranking the units without the actual measurement of the variable of interest. This method come into prominence recently due to the efficiency and cost effectiveness. RSS method has been modified by various researchers. One of them is Balanced Groups RSS which takes into account measurements of the lowest, median and the largest ranked units of the sets. It is revealed that Balanced Groups RSS is more efficient than commonly used Simple Random Sampling (SRS) and it gives unbiased results when the underlying distribution is symmetric. In this study, we propose an improvement of the sampling procedure of Balanced Groups RSS which is called Adjusted-Balanced Groups RSS for even and odd set sizes in decreasing the ranking error and the cost of the method. Adjusted-Balanced Groups RSS offers a more flexible sampling plan ensuring the independence of the set and sample sizes. Also, we propose a weighted estimator with a weight formula to obtain an efficient estimation of the population mean when the underlying distribution is skewed. A Monte Carlo simulation study is conducted for comparing the performance of the mean estimators according to mean squared errors and sampling costs for the methods of Adjusted-Balanced Groups RSS, Balanced Groups RSS, Median RSS, Percentile RSS and Extreme RSS in various type of distributions. Moreover, we compare the efficiencies of the proposed weighted estimator in Balanced Groups RSS and the classical estimator for the population mean in different type of asymmetric probability distributions with different set and sample sizes.

Keywords: Ranked set sampling, mean estimation, cost efficiency, weighted estimator, balanced groups ranked set sampling.

DEĞİŞTİRİLMİŞ SIRALI KÜME ÖRNEKLEMESİNE DAYALI KİTLE ORTALAMASININ KESTİRİMİ

ÖZ

Sıralı küme örneklemesi (SKÖ) birimlerin ölçümünün zaman alıcı ve/veya pahalı olduğu durumlarda kitle ortalaması kestirimi için etkin ve faydalı bir yöntemdir. SKÖ prosedürü, ilgilenilen değişkeninin gerçek ölçümü olmadan birimlerin sıralanmasına dayanır. SKÖ yöntemi, etkinliği ve maliyet uygunluğu sebebi ile son dönemlerde önem kazanmıştır. SKÖ'de birçok araştırmacı tarafından etkinliği artırmak adına değişiklikler yapılmıştır. Değiştirilmiş SKÖ türlerinden biri olan Dengeli Gruplar SKÖ, kitle ortalaması kestirimi için sıralanmış kümedeki en küçük, ortanca ve en büyük birimleri kullanmaktadır. Dengeli Gruplar SKÖ ile yapılan kitle ortalaması kestiriminin, yaygın olarak kullanılan Basit Rastgele Örnekleme (BRÖ)'den daha etkin olduğu ve simetrik dağılımlarda yansız sonuçlar verdiği ortaya konmuştur. Bu çalışmada Dengeli Gruplar SKÖ, sıralama hatası ve yöntem maliyetini düşürebilmek adına tek ve çift örneklem genişlikleri için geliştirilerek Düzeltilmiş-Dengeli Gruplar SKÖ olarak önerilmiştir. Düzeltilmiş-Dengeli Gruplar SKÖ, küme ve örneklem genişliklerinin bağımsızlığını sağlayan daha esnek bir örneklem tasarımı sunmaktadır. Ayrıca, dağılım asimetrik olduğu durumda kitle ortalaması kestiriminde daha etkin sonuçlar veren bir ağırlıklı kestirici ile ağırlık formülü önerilmektedir. Düzeltilmiş-Dengeli Gruplar SKÖ, Dengeli Gruplar SKÖ, Ortanca SKÖ, Yüzdelik SKÖ ve Uç Değerler SKÖ yöntemlerinin performansları, hata kareler ortalaması ve örneklem maliyeti açılarından karşılaştırılması amacıyla çeşitli dağılımlar üzerinde bir Monte Carlo benzetim çalışması yapılmıştır. Buna ek olarak önerilen ağırlıklı Dengeli Gruplar SKÖ kestiricisi ile klasik kestiricinin etkinlikleri, farklı örneklem ve küme genişliklerinde çeşitli asimetrik olasılık dağılımları için karşılaştırılmıştır.

Anahtar kelimeler: Sıralı küme örneklemesi, ortalama kestirimi, maliyet etkinliği, ağırlıklı kestirici, dengeli gruplar sıralı küme örneklemesi.

CONTENTS

Page

THESIS EXAMINATION RESULT FORM ii
ACKNOWLEDGEMENTS iii
ABSTRACT iv
ÖZ v
LIST OF FIGURES viii
LIST OF TABLES ix

CHAPTER ONE - INTRODUCTION 1

CHAPTER TWO – RANKED SET SAMPLING 3

CHAPTER THREE – SOME MODIFICATIONS OF RANKED SET SAMPLING 5

3.1 Median Ranked Set Sampling 5
3.2 Extreme Ranked Set Sampling 6
3.3 Percentile Ranked Set Sampling 9
3.4 L-Ranked Set Sampling 11
3.5 Truncation-Based Ranked Set Sampling 12
3.6 Balanced Groups Ranked Set Sampling 13

CHAPTER FOUR – IMPROVEMENTS ON BALANCED GROUPS RANKED SET SAMPLING 16

4.1 Adjusted-Balanced Groups Ranked Set Sampling 16
4.2 Weighted Estimator for Population Mean 19

CHAPTER FIVE – SIMULATION STUDY 26

5.1 Cost Efficiency of Adjusted-Balanced Groups Ranked Set Sampling..... 26
5.2 Comparison with Other Modified Ranked Set Sampling Procedures..... 28

CHAPTER SIX - CONCLUSION 32

REFERENCES 33

LIST OF FIGURES

Page

Figure 4.1 RE values for weighted estimator in BGRSS for Exponential (1), where skewness of the distribution s=1.9787 21

Figure 4.2 RE values for weighted estimator in BGRSS for Exponential (10), where skewness of the distribution s=1.9475 21

Figure 4.3 RE values for weighted estimator in BGRSS for F (30, 50), where skewness of the distribution s=0.9469 22

Figure 4.4 RE values for weighted estimator in BGRSS for F (100, 100), where skewness of the distribution s=0.6223 23

Figure 4.5 RE values for weighted estimator in BGRSS for Weibull (0.5, 5), where skewness of the distribution s=6.7681 23

Figure 4.6 RE values for weighted estimator in BGRSS for Weibull (0.8, 5), where skewness of the distribution s=2.8182 24

Figure 4.7 RE values for weighted estimator in BGRSS for Weibull (1, 5), where skewness of the distribution s=2.00 24

LIST OF TABLES

Page

Table 5.1 RCE values for the cases of $m = 6$ and CR $= 2, 5, 10, 20.$ 28
Table 5.2 Relative efficiency between A-BGRSS and SRS for $m = 4$ 29
Table 5.3 Relative efficiency between A-BGRSS and SRS for $m = 6$ 29
Table 5.4 MSE values of ERSS, MRSS, PRSS, BGRSS, A-BGRSS for $m = 6.$ 29
Table 5.5 MSE values of ERSS, MRSS, PRSS, A-BGRSS for $m = 4$ 31
Table 5.6 MSE values of ERSS, MRSS, PRSS, A-BGRSS for $m = 5$ 31

CHAPTER ONE
INTRODUCTION

McIntyre (1952) has introduced the Ranked Set Sampling (RSS) method for estimation process of a population mean. This method improves the efficiency of the mean estimator when the ranking of the sets of samples is easier and inexpensive than measurement of units. Halls and Dell (1966) first introduced the name of RSS and used the method for estimation process of the forage yields. In the following years, the mathematical background of the method was developed by Takahashi and Wakimoto (1968), Takahashi (1970). Also, Dell and Clutter (1972) showed that RSS is more cost effective than commonly used Simple Random Sampling (SRS). To obtain some other results, as a particular case, in the setting of the parameter estimation see for example Stokes (1977, 1980), Sinha et al. (1996), Yu and Tam (2002), Mahdizadeh and Arghami (2012), Balcı et al. (2013). Reviews for other details and examples can be found in the works of Kaur et al. (1995), Patil (1995), Wolfe (2012), Al-Omari and Bouza (2014).

In RSS, there is a random set size of m which is selected from the population and a researcher ranks these m units based on an expert judgment, visual inspection or auxiliary variable without measuring them. After the ranking process, ranked units which has equal set number and ranked number are chosen to measure in each m set. These m ranked units are measured which corresponds to a different judgment ranking within its own set.

In the literature, there are many studies focused on the modification in the process of RSS. Samawi et al. (1996) presented extreme RSS (ERSS) in order to estimate the population mean where only minimum and maximum units are desired for estimation (one for each set). Muttlak (1997) offered median RSS (MRSS) for efficient mean estimation and suggested to use only median units of the sets. Al-Saleh and Al-Kadiri (2000) presented Double-RSS (DRSS) for estimating the population mean as the first step of multi-stage RSS studies. Al-Saleh and Al-Omari (2002) introduced Multistage RSS (MSRSS) as a generalization of RSS and it is found more efficient

than SRS for estimating the population mean. Muttlak (2003a) suggested quartile RSS (QRSS) so as to estimate the population mean. Because of using the quartiles of the sets, QRSS gives better results especially for asymmetric distributions. Similar to this method, Muttlak (2003b) proposed Percentile RSS by using percentiles instead of quartiles for more flexible selection from ranked sets. With a robust perspective, Al-Naseer (2007) investigated a generalization procedure for robust visual ranking, called L-RSS which is ground on the idea of L-statistics and an appropriate mixture of RSS, QRSS and MRSS. Jemain et al. (2008) suggested Balanced Groups RSS (BGRSS) for estimating the population mean with samples of size $m = 3k$ (where $k = 1,2,3, \ldots$) which combines ERSS and MRSS in one procedure. Al-Naseer and Mustafa (2009) proposed Robust Extreme RSS (RERSS) as a generalization of some modified methods such as ERSS, MRSS and PRSS. Al-Omari and Raqab (2013) introduced a Truncation-Based RSS (TBRSS) to estimate the population mean and median. For the examples on other modifications, see Jemain and Al-Omari (2006a, 2006b, 2006c, 2007a, 2007b), Al-Omari and Al-Saleh (2009), Al-Omari and Raqab (2013), Haq et al. (2013, 2014) and references therein.

BGRSS is more efficient for estimating the population mean than commonly used SRS and it gives unbiased results when the underlying distribution is symmetric. In BGRSS, there are two disadvantages; one of them is that; when the set size is even, two ranked units must be measured for the median to obtain a sample. And the other one is that the selection of the set size depends on the sample size. In this thesis, we give an improvement on the process of BGRSS which reduces the sampling costs. Also it provides the independence of the set size and sample size. We also propose a weighted estimator with a weight formula to obtain an efficient mean estimation for the population mean. Second chapter of the study involves introduction of RSS procedure. Third chapter includes descriptions and the mean estimators of some of the modified methods of RSS. Fourth chapter introduces the improvements on BGRSS and the mean estimators. Chapter five provides simulation results based on the several distributions and compares the performance of the estimator. Finally, the last chapter presents the conclusion remarks.

CHAPTER TWO
RANKED SET SAMPLING

RSS is an efficient and advantageous method for estimating the mean of the population especially when the exact measurement of an observation is difficult and/or expensive and visual ranking can be done easily. This method come into prominence recently due to the efficiency and cost effectiveness. It is first proposed by McIntyre (1952) without any theoretical background. It is found that RSS is more efficient than commonly used SRS for estimating the population mean and gives unbiased results. The procedure of RSS can be described as follows:

Step 1: Randomly select m^2 units from target population.

Step 2: Allocate these m^2 units into m sets, each of size m.

Step 3: Rank the units in each set by personal/expert judgement or with measurements of concomitant variable that is correlated with variable of interest.

Step 4: Select i^{th} smallest ranked unit from i^{th} set to obtain m units of a ranked set sample, where $i = 1, 2, \dots, m$.

The four steps mentioned above can be repeated r times to get a sample of size rm. It is called as r-cycled ranked set sample. The procedure of RSS for one cycle and the case of $m = 3$ can be shown as below:

$$\begin{bmatrix} \boxed{X_{1(1:3)}} \leq X_{1(2:3)} \leq X_{1(3:3)} \\ X_{2(1:3)} \leq \boxed{X_{2(2:3)}} \leq X_{2(3:3)} \\ X_{3(1:3)} \leq X_{3(2:3)} \leq \boxed{X_{3(3:3)}} \end{bmatrix}$$

$X_{1(1:3)}$ from the first set, $X_{2(2:3)}$ from the second set and lastly $X_{3(3:3)}$ from the third set are choosen to be measured. ($X_{i(j:m)}$ represents the j^{th} ranked unit in the i^{th} set for $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, m$).

If this process repeated r times;

$$\begin{bmatrix} \boxed{X_{1(1:3)1}} \leq X_{1(2:3)1} \leq X_{1(3:3)1} \\ X_{2(1:3)1} \leq \boxed{X_{2(2:3)1}} \leq X_{2(3:3)1} \\ X_{3(1:3)1} \leq X_{3(2:3)1} \leq \boxed{X_{3(3:3)1}} \end{bmatrix}$$

$$\begin{bmatrix} \boxed{X_{1(1:3)2}} \leq X_{1(2:3)2} \leq X_{1(3:3)2} \\ X_{2(1:3)2} \leq \boxed{X_{2(2:3)2}} \leq X_{2(3:3)2} \\ X_{3(1:3)2} \leq X_{3(2:3)2} \leq \boxed{X_{3(3:3)2}} \end{bmatrix}$$

...

$$\begin{bmatrix} \boxed{X_{1(1:3)r}} \leq X_{1(2:3)r} \leq X_{1(3:3)r} \\ X_{2(1:3)r} \leq \boxed{X_{2(2:3)r}} \leq X_{2(3:3)r} \\ X_{3(1:3)r} \leq X_{3(2:3)r} \leq \boxed{X_{3(3:3)r}} \end{bmatrix}$$

the ranked set sample size of $3r$ observation can be obtained. ($X_{i(j:m)h}$ represents the j^{th} ranked unit in the i^{th} set and h^{th} cycle (for $i = 1, 2, \dots m,\ \ j = 1, 2, \dots, m$ and $h = 1, 2, \dots, r$). The representation of obtained ranked set sample for set size m and cycle size r can be shown below.

$$\begin{bmatrix} X_{[1]1} & X_{[2]1} & \cdots & X_{[m]1} \\ X_{[1]2} & X_{[2]2} & \cdots & X_{[m]2} \\ \vdots & \vdots & \ddots & \vdots \\ X_{[1]r} & X_{[2]r} & \cdots & X_{[m]r} \end{bmatrix}$$

Here $X_{[i]h}$ represents the i^{th} ranked observation in the h^{th} cycle for $i = 1, 2, \dots, m$ and $h = 1, 2, \dots, r$. The estimator of the population mean and its variance of one-cycled RSS method are given in (2.1) and (2.2).

$$\overline{X}_{RSS} = \frac{1}{rm} \sum_{h=1}^{r} \sum_{i=1}^{m} X_{[i]h}\ , \tag{2.1}$$

$$V(\overline{X}_{RSS}) = \frac{1}{rm^2} \sum_{h=1}^{r} \sum_{i=1}^{m} V(X_{[i]h})\ . \tag{2.2}$$

RSS method is cost effective in some situations and more efficient than commonly used SRS method. In SRS method, there is no information about the units in which part of the population is belonging to. However, a balanced RSS procedure obtains a sample which contains one unit from each m subregions of the population. Therefore, a ranked set sample is more effective than a simple random sample on representing the population. For more information, see Wolfe (2012) and McIntyre (1952).

CHAPTER THREE
SOME MODIFICATIONS OF RANKED SET SAMPLING

There are some modifications on RSS in literature based on the characteristics of a population. In this chapter, some of the modifications of RSS which are MRSS, ERSS, PRSS, L-RSS, TBRSS and lastly BGRSS are introduced.

3.1 Median Ranked Set Sampling

The procedure of MRSS is first revealed by Muttlak (1997) as given below.

Step 1: Select m random samples of size m from the target population.

Step 2: Rank the units in each set with respect to the variable of interest by judgement or any inexpensive method.

Step 3: If the set size m is an odd number, select the $((m+1)/2)^{th}$ smallest ranked units of the sets. If m is an even number, select the $(m/2)^{th}$ smallest ranked units from the first $(m/2)$ sets and $((m+2)/2)^{th}$ smallest ranked units from the second $(m/2)$ sets.

An example of the one-cycled MRSS procedure for the case of $m = 4$ is given below.

$$\begin{bmatrix} X_{1(1:4)} \leq \boxed{X_{1(2:4)}} \leq X_{1(3:4)} \leq X_{1(4:4)} \\ X_{2(1:4)} \leq \boxed{X_{2(2:4)}} \leq X_{2(3:4)} \leq X_{2(4:4)} \\ X_{3(1:4)} \leq X_{3(2:4)} \leq \boxed{X_{3(3:4)}} \leq X_{3(4:4)} \\ X_{4(1:4)} \leq X_{4(2:4)} \leq \boxed{X_{4(3:4)}} \leq X_{4(4:4)} \end{bmatrix}$$

The set size m is an even number. Therefore, the four sets whose units are ranked with respect to a variable of interest are splitted up into two groups consisting of two sets. Then, 2^{nd} smallest ranked units from the first two sets $(X_{1(2:4)}, X_{2(2:4)})$ and 3^{rd} smallest ranked units from the second two sets $(X_{3(3:4)}, X_{4(3:4)})$ are choosen to be measured. Another example of the one-cycled MRSS procedure for the case of $m = 5$ is given below.

$$\begin{bmatrix} X_{1(1:5)} \le X_{1(2:5)} \le \boxed{X_{1(3:5)}} \le X_{1(4:5)} \le X_{1(5:5)} \\ X_{2(1:5)} \le X_{2(2:5)} \le \boxed{X_{2(3:5)}} \le X_{2(4:5)} \le X_{2(5:5)} \\ X_{3(1:5)} \le X_{3(2:5)} \le \boxed{X_{3(3:5)}} \le X_{3(4:5)} \le X_{3(5:5)} \\ X_{4(1:5)} \le X_{4(2:5)} \le \boxed{X_{4(3:5)}} \le X_{4(4:5)} \le X_{4(5:5)} \\ X_{5(1:5)} \le X_{5(2:5)} \le \boxed{X_{5(3:5)}} \le X_{5(4:5)} \le X_{5(5:5)} \end{bmatrix}$$

When the set size m is an odd number, the 3^{rd} smallest ranked units of the sets $(X_{i(3:5)}, where\ i = 1, 2, \dots, m)$ are choosen to be measured.

The mean estimator and the variance based on one-cycled MRSS for odd set size case is given below;

$$\hat{\mu}^{O}_{MRSS} = \frac{1}{m}\sum_{i=1}^{m} \mathrm{X}_{i(\frac{m+1}{2}:m)}, \tag{3.1}$$

$$V(\hat{\mu}^{O}_{MRSS}) = \frac{1}{m^2}\sum_{i=1}^{m} V\left(\mathrm{X}_{i(\frac{m+1}{2}:m)}\right), \tag{3.2}$$

where $i = 1, 2, \dots, m$ and m can be any positive odd integer. For the even set size, the mean estimator and the variance based on one-cycled MRSS is;

$$\hat{\mu}^{E}_{MRSS} = \frac{1}{m}\left(\sum_{i=1}^{m/2} \mathrm{X}_{i(\frac{m}{2}:m)} + \sum_{i=(m/2)+1}^{m} \mathrm{X}_{i(\frac{m+2}{2}:m)}\right), \tag{3.3}$$

$$V(\hat{\mu}^{E}_{MRSS}) = \frac{1}{m^2}\left(\sum_{i=1}^{m/2} V\left(\mathrm{X}_{i(\frac{m}{2}:m)}\right) + \sum_{i=(m/2)+1}^{m} V\left(\mathrm{X}_{i(\frac{m+2}{2}:m)}\right)\right), \tag{3.4}$$

where $i = 1, 2, \dots, m$ and m can be any positive even integer.

The procedure which is mentioned above can be repeated r times to obtain a sample of size rm for estimating the population mean. For more details, see Muttlak (1997).

3.2 Extreme Ranked Set Sampling

Samawi et al (1996) described the procedure of ERSS as follows:

Step 1: Randomly draw m sets of size m units from the population under study.

Step 2: Rank the units in sets from lowest to largest with any inexpensive method.

Step 3: If the set size m is even, select the 1^{st} ranked unit from the first set, m^{th} ranked unit from the second set, 1^{st} ranked unit from the third set, m^{th} ranked unit from the fourth set and so on. In other words, select the smallest ranked units from the sets which are odd numbered and select the largest ranked units from the sets which are even numbered. If the sample size is odd, selecting process for the first $m-1$ sets will be same as even set size case. There are two options for selection of the last unit from the last set:

1. For the measure of m^{th} unit, the arithmetic mean of the 1^{st} and the m^{th} unit can be calculated.
2. For the measure of m^{th} unit, $((m+1)/2)^{th}$ unit can be choosen.

Examples for the procedure of ERSS can be represented as below.

$$\begin{bmatrix} \boxed{X_{1(1:4)}} \le X_{1(2:4)} \le X_{1(3:4)} \le X_{1(4:4)} \\ X_{2(1:4)} \le X_{2(2:4)} \le X_{2(3:4)} \le \boxed{X_{2(4:4)}} \\ \boxed{X_{3(1:4)}} \le X_{3(2:4)} \le X_{3(3:4)} \le X_{3(4:4)} \\ X_{4(1:4)} \le X_{4(2:4)} \le X_{4(3:4)} \le \boxed{X_{4(4:4)}} \end{bmatrix}$$

Set size $m = 4$, after ranking the units in each set, 1^{st} ranked units from the first and third sets ($X_{1(1:4)}\ and\ X_{3(1:4)}$) and 4^{th} ranked units from the second and fourth sets ($X_{2(4:4)}\ and\ X_{4(4:4)}$) are choosen for measuring.

$$\begin{bmatrix} \boxed{X_{1(1:5)}} \le X_{1(2:5)} \le X_{1(3:5)} \le X_{1(4:5)} \le X_{1(5:5)} \\ X_{2(1:5)} \le X_{2(2:5)} \le X_{2(3:5)} \le X_{2(4:5)} \le \boxed{X_{2(5:5)}} \\ \boxed{X_{3(1:5)}} \le X_{3(2:5)} \le X_{3(3:5)} \le X_{3(4:5)} \le X_{3(5:5)} \\ X_{4(1:5)} \le X_{4(2:5)} \le X_{4(3:5)} \le X_{4(4:5)} \le \boxed{X_{4(5:5)}} \\ (X_{5(1:5)}) \le X_{5(2:5)} \le X_{5(3:5)} \le X_{5(4:5)} \le (X_{5(5:5)}) \end{bmatrix}$$

When the set size is an odd number, there are two options for selection of last unit. For the first case, it is equal to the mean of the smallest and the largest ranked units are circled above.

$$\begin{bmatrix} \boxed{X_{1(1:5)}} \le X_{1(2:5)} \le X_{1(3:5)} \le X_{1(4:5)} \le X_{1(5:5)} \\ X_{2(1:5)} \le X_{2(2:5)} \le X_{2(3:5)} \le X_{2(4:5)} \le \boxed{X_{2(5:5)}} \\ \boxed{X_{3(1:5)}} \le X_{3(2:5)} \le X_{3(3:5)} \le X_{3(4:5)} \le X_{3(5:5)} \\ X_{4(1:5)} \le X_{4(2:5)} \le X_{4(3:5)} \le X_{4(4:5)} \le \boxed{X_{4(5:5)}} \\ X_{5(1:5)} \le X_{5(2:5)} \le \boxed{X_{5(3:5)}} \le X_{5(4:5)} \le X_{5(5:5)} \end{bmatrix}$$

For the second case, last unit is the median of the last set like shown above. The mean estimator of ERSS for even set size case is defined as;

$$\hat{\mu}_{ERSS}^{E} = \frac{1}{m}\left(\sum_{i=1}^{m/2} \mathrm{X}_{2i-1(1:m)} + \sum_{i=1}^{m/2} \mathrm{X}_{2i(m:m)}\right), \tag{3.5}$$

with the variance

$$V(\hat{\mu}_{ERSS}^{E}) = \frac{1}{m^2}\left(\sum_{i=1}^{m/2} V\left(\mathrm{X}_{2i-1(1:m)}\right) + \sum_{i=1}^{m/2} V\left(\mathrm{X}_{2i(m:m)}\right)\right). \tag{3.6}$$

The mean estimators of ERSS for the odd set size case are defined in (3.7) and (3.8),

$$\hat{\mu}_{ERSS}^{O1} = \frac{1}{m}\left(\sum_{i=1}^{(m-1)/2} \mathrm{X}_{2i-1(1:m)} + \sum_{i=1}^{(m-1)/2} \mathrm{X}_{2i(m:m)} + \frac{1}{2}\left(\mathrm{X}_{m(1:m)} + \mathrm{X}_{m(m:m)}\right)\right), \tag{3.7}$$

$$\hat{\mu}_{ERSS}^{O2} = \frac{1}{m}\left(\sum_{i=1}^{(m-1)/2} \mathrm{X}_{2i-1(1:m)} + \sum_{i=1}^{(m-1)/2} \mathrm{X}_{2i(m:m)} + \mathrm{X}_{m(\frac{m+1}{2}:m)}\right), \tag{3.8}$$

and their variances are

$$V(\hat{\mu}_{ERSS}^{O1}) = \frac{1}{m^2}\begin{pmatrix} \sum_{i=1}^{(m-1)/2} V\left(\mathrm{X}_{2i-1(1:m)}\right) + \sum_{i=1}^{(m-1)/2} V\left(\mathrm{X}_{2i(m:m)}\right) \\ + \frac{1}{4}\left(V\left(\mathrm{X}_{m(1:m)}\right) + V\left(\mathrm{X}_{m(m:m)}\right)\right) \end{pmatrix}, \tag{3.9}$$

$$V(\hat{\mu}_{ERSS}^{O2}) = \frac{1}{m^2}\left(\sum_{i=1}^{(m-1)/2} V\left(X_{2i-1(1:m)}\right) + \sum_{i=1}^{(m-1)/2} V\left(X_{2i(m:m)}\right) + V\left(X_{m\left(\frac{m+1}{2}:m\right)}\right)\right). \quad (3.10)$$

The procedure which is mentioned above can be repeated r times to obtain a sample of size rm for estimating the population mean. See Samawi (1996), Muttlak and Abu-Dayyeh (2004) for more details.

3.3 Percentile Ranked Set Sampling

Muttlak (2003b) is proposed the PRSS for estimating the mean of target population. Following steps are described to obtain a sample by using PRSS method.

Step 1: Randomly select m sets of size m units from the population of interest.

Step 2: Rank the units in each set by judgement or any inexpensive method.

Step 3: If m is an even number, select the $[p(m+1)]^{th}$ranked units from the first $(m/2)$ sets and select the $[(1-p)(m+1)]^{th}$ ranked units from the second $(m/2)$ sets, where $0 \le p \le 1$. If m is an odd number, select the $[p(m+1)]^{th}$ ranked units from the first $((m-1)/2)$ sets, $[(1-p)(m+1)]^{th}$ ranked units from the second $((m-1)/2)$ sets and median unit of the m^{th} set, where $0 \le p \le 1$ and $[x]$ is the nearest integer value to x.

Examples for the procedure of PRSS can be represented as below for even and odd set sizes respectively.

$$\begin{bmatrix} \boxed{X_{1(1:4)}} \le X_{1(2:4)} \le X_{1(3:4)} \le X_{1(4:4)} \\ \boxed{X_{2(1:4)}} \le X_{2(2:4)} \le X_{2(3:4)} \le X_{2(4:4)} \\ X_{3(1:4)} \le X_{3(2:4)} \le X_{3(3:4)} \le \boxed{X_{3(4:4)}} \\ X_{4(1:4)} \le X_{4(2:4)} \le X_{4(3:4)} \le \boxed{X_{4(4:4)}} \end{bmatrix}$$

Suppose that the set size $m = 4$ and $p = 0.25$. After ranking the units, 1^{st} ranked units from the first half of sets ($X_{1(1:4)}$ and $X_{2(1:4)}$) and 4^{th} ranked units from the second half of sets ($X_{3(4:4)}$, and $X_{4(4:4)}$) are choosen for measuring.

$$\begin{bmatrix} X_{1(1:5)} \le \boxed{X_{1(2:5)}} \le X_{1(3:5)} \le X_{1(4:5)} \le X_{1(5:5)} \\ X_{2(1:5)} \le \boxed{X_{2(2:5)}} \le X_{2(3:5)} \le X_{2(4:5)} \le X_{2(5:5)} \\ X_{3(1:5)} \le X_{3(2:5)} \le X_{3(3:5)} \le \boxed{X_{3(4:5)}} \le X_{3(5:5)} \\ X_{4(1:5)} \le X_{4(2:5)} \le X_{4(3:5)} \le \boxed{X_{4(4:5)}} \le X_{4(5:5)} \\ X_{5(1:5)} \le X_{5(2:5)} \le \boxed{X_{5(3:5)}} \le X_{5(4:5)} \le X_{5(5:5)} \end{bmatrix}$$

Suppose that the set size $m = 5$ and $p = 0.3$. After ranking the units, 2^{nd} ranked units from the first $(m-1)/2$ of sets ($X_{1(2:5)}$ and $X_{2(2:5)}$), 4^{th} ranked units from the second $(m-1)/2$ of sets ($X_{3(4:5)}$ and $X_{4(4:5)}$) and lastly median of the last set ($X_{5(3:5)}$) are choosen to be measured. Mean estimators based on PRSS method are given in (3.11) and (3.12) for even and odd set sizes respectively,

$$\hat{\mu}^{E}_{PRSS} = \frac{1}{m}\left(\sum_{i=1}^{m/2} \mathrm{X}_{i(a:m)} + \sum_{i=(m/2)+1}^{m} \mathrm{X}_{i(b:m)}\right), \tag{3.11}$$

$$\hat{\mu}^{O}_{PRSS} = \frac{1}{m}\left(\sum_{i=1}^{(m-1)/2} \mathrm{X}_{i(a:m)} + \sum_{i=((m-1)/2)+1}^{m-1} \mathrm{X}_{i(b:m)} + \mathrm{X}_{i(\frac{m+1}{2}:m)}\right), \tag{3.12}$$

with variances

$$V(\hat{\mu}^{E}_{PRSS}) = \frac{1}{m^2}\left(\sum_{i=1}^{m/2} V\left(\mathrm{X}_{i(a:m)}\right) + \sum_{i=(m/2)+1}^{m} V\left(\mathrm{X}_{i(b:m)}\right)\right), \tag{3.13}$$

$$V(\hat{\mu}^{O}_{PRSS}) = \frac{1}{m^2}\left(\sum_{i=1}^{(m-1)/2} V\left(\mathrm{X}_{i(a:m)}\right) + \sum_{i=((m-1)/2)+1}^{m-1} V\left(\mathrm{X}_{i(b:m)}\right) + V\left(\mathrm{X}_{i(\frac{m+1}{2}:m)}\right)\right), \tag{3.14}$$

where, $a = [p(m+1)]$ and $b = [(1-p)(m+1)]$. The procedure which is mentioned above completes one cycle of PRSS and it can be repeated r times to obtain a sample of size rm units. See Muttlak (2003b) for more details.

3.4 L-Ranked Set Sampling

Al-Naseer (2007) introduced L-RSS as a generalization of different type of RSS methods. In order to obtain a sample by using L-RSS, the steps mentioned below should be followed.

Step 1: Randomly identify m sets, where each set composed of m units.

Step 2: Rank the units by any inexpensive method with respect to variable of interest.

Step 3: Choose the L-RSS coefficient $k = \lfloor m\alpha \rfloor$, where $0 \leq \alpha < 0.5$, and $\lfloor x \rfloor$ is the largest integer value, less than or equal to x.

Step 4: Select the $(k+1)^{th}$ ranked units from the first $(k+1)$ of sets, $(m-k)^{th}$ ranked units from the last $(k+1)$ of sets and i^{th} ranked units from the remaining sets which are numbered with i, where $i = k+2, \ldots, m-k-1$.

The prodecure of L-RSS for $m = 6$ and $k = 1$ is shown below.

$$\begin{bmatrix} X_{1(1:6)} \leq \boxed{X_{1(2:6)}} \leq X_{1(3:6)} \leq X_{1(4:6)} \leq X_{1(5:6)} \leq X_{1(6:6)} \\ X_{2(1:6)} \leq \boxed{X_{2(2:6)}} \leq X_{2(3:6)} \leq X_{2(4:6)} \leq X_{2(5:6)} \leq X_{2(6:6)} \\ X_{3(1:6)} \leq X_{3(2:6)} \leq \boxed{X_{3(3:6)}} \leq X_{3(4:6)} \leq X_{3(5:6)} \leq X_{3(6:6)} \\ X_{4(1:6)} \leq X_{4(2:6)} \leq X_{4(3:6)} \leq \boxed{X_{4(4:6)}} \leq X_{4(5:6)} \leq X_{4(6:6)} \\ X_{5(1:6)} \leq X_{5(2:6)} \leq X_{5(3:6)} \leq X_{5(4:6)} \leq \boxed{X_{5(5:6)}} \leq X_{5(6:6)} \\ X_{6(1:6)} \leq X_{6(2:6)} \leq X_{6(3:6)} \leq X_{6(4:6)} \leq \boxed{X_{6(5:6)}} \leq X_{6(6:6)} \end{bmatrix}$$

After ranking the units, 2^{nd} smallest ranked units from the first $k+1$ of sets ($X_{1(2:6)}$ and $X_{2(2:6)}$), 5^{th} smallest ranked units from the last $k+1$ of sets ($X_{5(5:6)}$ and $X_{6(5:6)}$) and from remaining sets, the units which have equal ranked number and set number ($X_{3(3:6)}$ and $X_{4(4:6)}$) are selecting for measuring.

It is noted that when $k = 0$, the procedure of L-RSS is same with classical RSS method and when $k = [(m-1)/2]$, both L-RSS and MRSS schemes are equivalent. Also, PRSS could be the special case of this method. Mean estimator for the population is given in (3.15),

$$\hat{\mu}_{LRSS} = \frac{1}{m}\left(\sum_{i=1}^{k} X_{i((k+1):m)} + \sum_{i=k+1}^{m-k} X_{i(i:m)} + \sum_{i=m-k+1}^{m} X_{i((m-k):m)}\right), \tag{3.15}$$

and its variance is given by;

$$V(\hat{\mu}_{LRSS}) = \frac{1}{m^2}\left(\sum_{i=1}^{k} V(X_{i((k+1):m)}) + \sum_{i=k+1}^{m-k} V(X_{i(i:m)}) + \sum_{i=m-k+1}^{m} V(X_{i((m-k):m)})\right). \tag{3.16}$$

The procedure which is mentioned above completes one cycle of LRSS and it can be repeated r times to obtain a sample of size rm units. For detailed information, see Al-Naseer (2007).

3.5 Truncation-Based Ranked Set Sampling

The TBRSS scheme is presented by Al-Omari and Raqab (2013) as follows:

Step 1: Randomly draw m sets of size m units from the target population.

Step 2: Rank the units in each set by any inexpensive method with respect to variable of interest.

Step 3: Choose a TBRSS coefficient $k = \lfloor m\alpha \rfloor$ where, $0 \leq \alpha < 0.5$, and $\lfloor x \rfloor$ is the largest integer value, less than or equal to x.

Step 4: Select the minimums of the first k sets, select the maximums of the last k sets and select the units with equal set and rank number from the remaining $m - 2k$ sets.

An example for the procedure of TBRSS when $m = 8$, $\alpha = 0.35$ and $k = 2$.

$$\begin{bmatrix} \boxed{X_{1(1:8)}} \le X_{1(2:8)} \le X_{1(3:8)} \le X_{1(4:8)} \le X_{1(5:8)} \le X_{1(6:8)} \le X_{1(7:8)} \le X_{1(8:8)} \\ \boxed{X_{2(1:8)}} \le X_{2(2:8)} \le X_{2(3:8)} \le X_{2(4:8)} \le X_{2(5:8)} \le X_{2(6:8)} \le X_{2(7:8)} \le X_{2(8:8)} \\ X_{3(1:8)} \le X_{3(2:8)} \le \boxed{X_{3(3:8)}} \le X_{3(4:8)} \le X_{3(5:8)} \le X_{3(6:8)} \le X_{3(7:8)} \le X_{3(8:8)} \\ X_{4(1:8)} \le X_{4(2:8)} \le X_{4(3:8)} \le \boxed{X_{4(4:8)}} \le X_{4(5:8)} \le X_{4(6:8)} \le X_{4(7:8)} \le X_{4(8:8)} \\ X_{5(1:8)} \le X_{5(2:8)} \le X_{5(3:8)} \le X_{5(4:8)} \le \boxed{X_{5(5:8)}} \le X_{5(6:8)} \le X_{5(7:8)} \le X_{5(8:8)} \\ X_{6(1:8)} \le X_{6(2:8)} \le X_{6(3:8)} \le X_{6(4:8)} \le X_{6(5:8)} \le \boxed{X_{6(6:8)}} \le X_{6(7:8)} \le X_{6(8:8)} \\ X_{7(1:8)} \le X_{7(2:8)} \le X_{7(3:8)} \le X_{7(4:8)} \le X_{7(5:8)} \le X_{7(6:8)} \le X_{7(7:8)} \le \boxed{X_{7(8:8)}} \\ X_{8(1:8)} \le X_{8(2:8)} \le X_{8(3:8)} \le X_{8(4:8)} \le X_{8(5:8)} \le X_{8(6:8)} \le X_{8(7:8)} \le \boxed{X_{8(8:8)}} \end{bmatrix}$$

Depending on the procedure of TBRSS defined above, after ranking the units, minimums of the first two sets ($X_{1(1:8)}$, $X_{2(1:8)}$), maximums of the last two sets ($X_{7(8:8)}$, $X_{8(8:8)}$) and the units which has the equal set and rank number ($X_{i(i:8)}$, for $i = 3, 4, 5, 6$) of the remaining sets are selected.

When $k = 0$ or $k = 1$, TBRSS scheme is equivalent to classical RSS scheme. The mean estimator of TBRSS for size of m is defined by

$$\hat{\mu}_{TBRSS} = \frac{1}{m}\left(\sum_{i=1}^{k} \mathrm{X}_{i(1:m)} + \sum_{i=k+1}^{m-k} \mathrm{X}_{i(i:m)} + \sum_{i=m-k+1}^{m} \mathrm{X}_{i(m:m)} \right), \tag{3.17}$$

with variance

$$V(\hat{\mu}_{TBRSS}) = \frac{1}{m^2}\left(\sum_{i=1}^{k} V(\mathrm{X}_{i(1:m)}) + \sum_{i=k+1}^{m-k} V(\mathrm{X}_{i(i:m)}) + \sum_{i=m-k+1}^{m} V(\mathrm{X}_{i(m:m)}) \right). \tag{3.18}$$

The procedure which is mentioned above completes one cycle of TBRSS and it can be repeated r times to obtain a sample of size rm units. For more detailed information, see Al-Omari and Raqab (2012).

3.6 Balanced Groups Ranked Set Sampling

BGRSS is first suggested by Jemain et al. (2008). The method of BGRSS can be described as follows:

Step 1: Randomly select $m = 3k$ $(k = 1, 2, 3, \ldots)$ sets each of size m from the target population,

Step 2: Sets are randomly allocating into three groups. Units in each set are ranked by judgement or any inexpensive method.

Step 3: The *k* smallest units from the first group, *k* median units from the second group and *k* largest units from the third group of ranked sets are chosen. For the second group, if the set size is odd, the median is the $((m+1)/2)^{th}$ ranked unit in the set and if the set size is even the median is defined as the mean of $(m/2)^{th}$ and $((m+2)/2)^{th}$ ranked units.

BGRSS scheme when $k = 3$ can be shown as follows:

$$\begin{bmatrix}
\boxed{X_{1(1:9)}} \le X_{1(2:9)} \le X_{1(3:9)} \le X_{1(4:9)} \le X_{1(5:9)} \le X_{1(6:9)} \le X_{1(7:9)} \le X_{1(8:9)} \le X_{1(9:9)} \\
\boxed{X_{2(1:9)}} \le X_{2(2:9)} \le X_{2(3:9)} \le X_{2(4:9)} \le X_{2(5:9)} \le X_{2(6:9)} \le X_{2(7:9)} \le X_{2(8:9)} \le X_{2(9:9)} \\
\boxed{X_{3(1:9)}} \le X_{3(2:9)} \le X_{3(3:9)} \le X_{3(4:9)} \le X_{3(5:9)} \le X_{3(6:9)} \le X_{3(7:9)} \le X_{3(8:9)} \le X_{3(9:9)} \\
X_{4(1:9)} \le X_{4(2:9)} \le X_{4(3:9)} \le X_{4(4:9)} \le \boxed{X_{4(5:9)}} \le X_{4(6:9)} \le X_{4(7:9)} \le X_{4(8:9)} \le X_{4(9:9)} \\
X_{5(1:9)} \le X_{5(2:9)} \le X_{5(3:9)} \le X_{5(4:9)} \le \boxed{X_{5(5:9)}} \le X_{5(6:9)} \le X_{5(7:9)} \le X_{5(8:9)} \le X_{5(9:9)} \\
X_{6(1:9)} \le X_{6(2:9)} \le X_{6(3:9)} \le X_{6(4:9)} \le \boxed{X_{6(5:9)}} \le X_{6(6:9)} \le X_{6(7:9)} \le X_{6(8:9)} \le X_{6(9:9)} \\
X_{7(1:9)} \le X_{7(2:9)} \le X_{7(3:9)} \le X_{7(4:9)} \le X_{7(5:9)} \le X_{7(6:9)} \le X_{7(7:9)} \le X_{7(8:9)} \le \boxed{X_{7(9:9)}} \\
X_{8(1:9)} \le X_{8(2:9)} \le X_{8(3:9)} \le X_{8(4:9)} \le X_{8(5:9)} \le X_{8(6:9)} \le X_{8(7:9)} \le X_{8(8:9)} \le \boxed{X_{8(9:9)}} \\
X_{9(1:9)} \le X_{9(2:9)} \le X_{9(3:9)} \le X_{9(4:9)} \le X_{9(5:9)} \le X_{9(6:9)} \le X_{9(7:9)} \le X_{9(8:9)} \le \boxed{X_{9(9:9)}}
\end{bmatrix}$$

The units selecting for measuring are the minimums from the first 3 sets $(X_{1(1:9)}, X_{2(1:9)}, X_{3(1:9)})$, median from the second 3 sets $(X_{4(5:9)}, X_{5(5:9)}, X_{6(5:9)})$ and maximums from the last 3 sets $(X_{7(9:9)}, X_{8(9:9)}, X_{9(9:9)})$. After measuring the units, estimation of the population mean can be calculated by the arithmetic means of the selected units. BGRSS process for $k = 2$ can be described as follows:

$$\begin{bmatrix}
\boxed{X_{1(1:6)}} \le X_{1(2:6)} \le X_{1(3:6)} \le X_{1(4:6)} \le X_{1(5:6)} \le X_{1(6:6)} \\
\boxed{X_{2(1:6)}} \le X_{2(2:6)} \le X_{2(3:6)} \le X_{2(4:6)} \le X_{2(5:6)} \le X_{2(6:6)} \\
X_{3(1:6)} \le X_{3(2:6)} \le \boxed{X_{3(3:6)}} \le \boxed{X_{3(4:6)}} \le X_{3(5:6)} \le X_{3(6:6)} \\
X_{4(1:6)} \le X_{4(2:6)} \le \boxed{X_{4(3:6)}} \le \boxed{X_{4(4:6)}} \le X_{4(5:6)} \le X_{4(6:6)} \\
X_{5(1:6)} \le X_{5(2:6)} \le X_{5(3:6)} \le X_{5(4:6)} \le X_{5(5:6)} \le \boxed{X_{5(6:6)}} \\
X_{6(1:6)} \le X_{6(2:6)} \le X_{6(3:6)} \le X_{6(4:6)} \le X_{6(5:6)} \le \boxed{X_{6(6:6)}}
\end{bmatrix}$$

Estimation of the population mean is the arithmetic mean of the minimums, ($X_{1(1:6)}$, $X_{2(1:6)}$), medians ($\frac{1}{2}(X_{3(3:6)} + X_{3(4:6)})$, $\frac{1}{2}(X_{4(3:6)} + X_{4(4:6)})$) and the maximums ($X_{5(6:6)}$, $X_{6(6:6)}$) of the groups. The arithmetic mean of the sample equals to the estimation of the population mean. It is noted that BGRSS method equals to classical RSS method in the case of $k = 1$. The mean estimator and its variance for odd set size are given as follows:

$$\hat{\mu}^{O}_{BGRSS} = \frac{1}{3k}\left(\sum_{i=1}^{k} \mathrm{X}_{i(1:m)} + \sum_{i=k+1}^{2k} \mathrm{X}_{i(\frac{m+1}{2}:m)} + \sum_{i=2k+1}^{3k} \mathrm{X}_{i(m:m)}\right), \tag{3.19}$$

$$V(\hat{\mu}^{O}_{BGRSS}) = \frac{1}{9k^2}\left(\sum_{i=1}^{k} V\left(\mathrm{X}_{i(1:m)}\right) + \sum_{i=k+1}^{2k} V\left(\mathrm{X}_{i(\frac{m+1}{2}:m)}\right) + \sum_{i=2k+1}^{3k} V\left(\mathrm{X}_{i(m:m)}\right)\right). \tag{3.20}$$

For the even set size, the mean estimator and its variance of BGRSS are as below:

$$\hat{\mu}^{E}_{BGRSS} = \frac{1}{3k}\left(\sum_{i=1}^{k} \mathrm{X}_{i(1:m)} + \frac{1}{2}\sum_{i=k+1}^{2k}\left(\mathrm{X}_{i(\frac{m}{2}:m)} + \mathrm{X}_{i(\frac{m+2}{2}:m)}\right) + \sum_{i=2k+1}^{3k} \mathrm{X}_{i(m:m)}\right), \tag{3.21}$$

$$V(\hat{\mu}^{E}_{BGRSS}) = \frac{1}{9k^2}\left(\begin{array}{l} \sum_{i=1}^{k} V(\mathrm{X}_{i(1:m)}) + \frac{1}{4}\sum_{i=k+1}^{2k}\left(\begin{array}{l} V(\mathrm{X}_{i(\frac{m}{2}:m)}) + V(\mathrm{X}_{i(\frac{m+2}{2}:m)}) \\ + 2Cov(\mathrm{X}_{i(\frac{m}{2}:m)}, \mathrm{X}_{i(\frac{m+2}{2}:m)}) \end{array}\right) \\ + \sum_{i=2k+1}^{3k} V(\mathrm{X}_{i(m:m)}) \end{array}\right). \tag{3.22}$$

The procedure which is mentioned above completes one cycle of BGRSS and it can be repeated r times to obtain a sample of size rm units. For more details, see Jemain et al. (2008).

CHAPTER FOUR
IMPROVEMENTS ON BALANCED GROUPS RANKED SET SAMPLING

Balanced Groups RSS (BGRSS) is proposed by Jemain et al. (2008) as a mixture of Extreme RSS (ERSS) and Median RSS (MRSS) sampling schemes. They showed that BGRSS is unbiased for symmetric distributions, biased for asymmetric distributions and more efficient than Simple Random Sampling (SRS) for both type of distributions. Because of the lowest ranked units, medians and the largest ranked units are being chosen to measure in BGRSS, it is more advantageous method comparing to other modified methods. However, the sampling procedure of BGRSS has some disadvantages given as follows:

i) set sizes must be chosen multiples of three. The other set sizes (i.e. 4, 5, 7 or etc.) cannot be used.

ii) for even set size case, the median units are determined by the mean of two units $((m/2)^{th}$ and $((m+2)/2)^{th}$ units). It means that we must measure k more units for $k = 2, 4, 6, \ldots$ and rk units for r cycle as total number of measurements. This may cause a serious increase in sampling cost in practice.

4.1 Adjusted-Balanced Groups Ranked Set Sampling

In BGRSS, set size and sample size have to be equal, that is, there are $m = 3k$ independent sets and each set has *m* units. Under this circumstance, it is not possible to study with the sample size which is not multiple of three. In this study we propose an adjustment for BGRSS which is called Adjusted-BGRSS. Advantages of Adjusted-BGRSS are described as below:

i. It decreases the quantity of measurements for even set size. In BGRSS, we need to measure 8 units to get a sample of size 6 in a cycle, while Adjusted-BGRSS decrease these units to 6.

ii. It provides flexibility on the set size *m* where *m* can be any positive integer and it does not have to equal to 3*k* as in BGRSS for both even and odd set size cases.

iii. It provides independency on quantity of set size and sample size. Studying with smaller sample size may give rise to decreasing on ranking error and ranking cost.

The procedure of Adjusted-BGRSS is described as follows:

Step 1: Select $6k$ sample of size of m by SRS. ($k = 1, 2, \dots, m = 3, 4, 5, \dots$)

Step 2: Allocate the samples into 4 groups of size $2k, k, k, 2k$ respectively and rank the units in each set by judgement or any inexpensive method.

Step 3: If the set size is an odd number; select the minimums of the first $2k$ group, $((m+1)/2)^{th}$ ranked units of the second and third k group and the maximums of the last $2k$ group. If the set size is an even number; select the minimums of the first $2k$ group, $(m/2)^{th}$ ranked units of the second k group, $((m+2)/2)^{th}$ ranked units of third k group and the maximums of the last $2k$ group.

An example for the case of $m = 4$ and $k = 1$ is as follows:

$$\begin{bmatrix} \boxed{X_{1(1:6)}} \le X_{1(2:6)} \le X_{1(3:6)} \le X_{1(4:6)} \\ \boxed{X_{2(1:6)}} \le X_{2(2:6)} \le X_{2(3:6)} \le X_{2(4:6)} \\ X_{3(1:6)} \le \boxed{X_{3(2:6)}} \le X_{3(3:6)} \le X_{3(4:6)} \\ X_{4(1:6)} \le X_{4(2:6)} \le \boxed{X_{4(3:6)}} \le X_{4(4:6)} \\ X_{5(1:6)} \le X_{5(2:6)} \le X_{5(3:6)} \le \boxed{X_{5(4:6)}} \\ X_{6(1:6)} \le X_{6(2:6)} \le X_{6(3:6)} \le \boxed{X_{6(4:6)}} \end{bmatrix}$$

In the above example of Adjusted-BGRSS, the minimums from the first group ($X_{1(1:4)}$ and $X_{2(1:4)}$), the $(m/2)^{th} = (2)^{nd}$ ranked unit from the second group ($X_{3(2:4)}$), the $((m+2)/2)^{th} = (3)^{rd}$ ranked unit from the third group ($X_{4(3:4)}$), and the maximums from the last group ($X_{5(4:4)}$ and $X_{6(4:4)}$) are chosen to be measured.

When the set size is equal to six, then Adjusted-BGRSS is the same as TBRSS for the case of $k = 2$ and $m = 6$. When the set size is three, selection is the same with two cycled classical RSS. When $m = 6k$, where m is an odd number, Adjusted-BGRSS is equalivalent to two-cycled BGRSS method for fixed set size m.

Let $X_1, X_2, \dots, X_m$ be a random sample with probability density function $f(x)$, with mean μ and variance σ^2 and let $X_{1(1:m)}, X_{1(2:m)}, \dots, X_{1(m:m)}$; $X_{2(1:m)}, X_{2(2:m)}, \dots, X_{2(m:m)}$; $\dots$; $X_{6k(1:m)}, X_{6k(2:m)}, \dots, X_{6k(m:m)}$ be independent simple random samples with the same cumulative distribution function $F(x)$. Estimator of the population mean in Adjusted-BGRSS is as given below when the set size is odd and even, respectively:

$$\hat{\mu}^{O}_{A-BGRSS} = \frac{1}{6k}\left(\sum_{i=1}^{2k} \mathrm{X}_{i(1:m)} + \sum_{i=2k+1}^{4k} \mathrm{X}_{i(\frac{m+1}{2}:m)} + \sum_{i=4k+1}^{6k} \mathrm{X}_{i(m:m)}\right), \tag{4.1}$$

$$\hat{\mu}^{E}_{A-BGRSS} = \frac{1}{6k}\left(\sum_{i=1}^{2k} \mathrm{X}_{i(1:m)} + \sum_{i=2k+1}^{3k} \mathrm{X}_{i(\frac{m}{2}:m)} + \sum_{i=3k+1}^{4k} \mathrm{X}_{i(\frac{m+2}{2}:m)} + \sum_{i=4k+1}^{6k} \mathrm{X}_{i(m:m)}\right). \tag{4.2}$$

Their corresponding variances can be given as follows:

$$V(\hat{\mu}^{O}_{A-BGRSS}) = \frac{1}{36k^2}\left(\sum_{i=1}^{2k} V(\mathrm{X}_{i(1:m)}) + \sum_{i=2k+1}^{4k} V(\mathrm{X}_{i(\frac{m+1}{2}:m)}) + \sum_{i=4k+1}^{6k} V(\mathrm{X}_{i(m:m)})\right), \tag{4.3}$$

$$V(\hat{\mu}^{E}_{A-BGRSS}) = \frac{1}{36k^2}\left(\begin{array}{l}\sum_{i=1}^{2k} V(\mathrm{X}_{i(1:m)}) + \sum_{i=2k+1}^{3k} V(\mathrm{X}_{i(\frac{m}{2}:m)}) \\ + \sum_{i=3k+1}^{4k} V(\mathrm{X}_{i(\frac{m+2}{2}:m)}) + \sum_{i=4k+1}^{6k} V(\mathrm{X}_{i(m:m)})\end{array}\right). \tag{4.4}$$

Since we know that if the underlying distribution is symmetric about $\mu = 0$, then $\mu_{i:m} = -\mu_{m-i+1:m}$ for all i, $(i = 1,2, \dots, m)$, and it can be shown that the mean

estimators in (4.1) and (4.2) based on Adjusted-BGRSS are unbiased for symmetric distributions (see David and Nagaraja (2003) and Jemain et al. (2008)).

4.2 Weighted Estimator for Population Mean

Abu-Dayyeh et al. (2002) have proposed a weighted mean estimator in ERSS and they have been found that weighted mean is unbiased and more efficient estimator in comparison with SRS when the underlying distribution is skewed. Also, Muttlak and Abu-Dayyeh (2004) have suggested the weighted mean for PRSS, MRSS and ERSS to increase the efficiency of estimation on population mean for skewed distributions. Al-Naseer (2007) proposed a weighted mean estimator based on entropy measure from information theory to improve the L-RSS scheme in estimating the population mean, when the underlying distribution is skewed. Al-Naseer and Mustafa (2009) introduced weighted estimator based on entropy measure for RERSS when the underlying distribution is asymmetric. In this study, we propose a weighted mean estimator for the skewed-distributed populations based on BGRSS under the idea that the minimum and the maximum units of the observations have the same weight equals to 1 and the median units should have equal or greater weight (say $w \geq 1$) than the weight of minimum and maximum units. Let $\tilde{\mu}_{BGRSS}$ represent the weighted mean estimator for the population mean. The estimators for odd and even set sizes are given as below respectively:

$$\tilde{\mu}^{o}_{BGRSS} = \frac{1}{k(w+2)}\left(\sum_{i=1}^{k} X_{i(1:m)} + w\sum_{i=k+1}^{2k} X_{i(\frac{m+1}{2}:m)} + \sum_{i=2k+1}^{3k} X_{i(m:m)}\right), \tag{4.5}$$

$$\tilde{\mu}^{E}_{BGRSS} = \frac{1}{k(w+2)}\left(\sum_{i=1}^{k} X_{i(1:m)} + \frac{w}{2}\sum_{i=k+1}^{2k}\left(X_{i(\frac{m}{2}:m)} + X_{i(\frac{m+2}{2}:m)}\right) + \sum_{i=2k+1}^{3k} X_{i(m:m)}\right). \tag{4.6}$$

After some numerical studies, we formulate the weight value w as

$$w = \sqrt{\left(\frac{|s|}{2} + 1\right)m}\,, \tag{4.7}$$

where m represents set size and s represents skewness coefficient of the distribution. It can be seen on the weight formula, the weight increase while the skewness of the distribution and set size increase, reasonably.

For the comparison of efficiencies of the mean estimators in (4.5) and (4.6) in terms of mean squared errors (MSEs), we design a simulation study in R software using the value of the proposed weight formula and the value of the optimal weight which have the smallest MSE of weighted estimator. Figure (4.1-4.7) shows the relative efficiency (RE) values which are the ratios of the MSE of the classical estimator (MSE $(\hat{\mu}_{BGRSS})$) and the MSE of the weighted estimator $(MSE(\tilde{\mu}_{BGRSS}))$. In the following, we define the relative efficiencies RE_1 and RE_2 using the MSE of the mean estimator for formulated weight value, $MSE(\tilde{\mu}^1_{BGRSS})$, and the MSE of the mean estimator for optimal weight value, $MSE(\tilde{\mu}^2_{BGRSS})$, respectively,

$$RE_1 = \frac{MSE(\hat{\mu}_{BGRSS})}{MSE(\tilde{\mu}^1_{BGRSS})} \text{ and } RE_2 = \frac{MSE(\hat{\mu}_{BGRSS})}{MSE(\tilde{\mu}^2_{BGRSS})},$$

where $MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{\mu} - \mu)^2$ and n is the number of estimations of population mean in this simulation study.

Lines in figures represent the *RE* values for the case of $m = 12$ while dashed and dotted lines represent the *RE* values for the case of $m = 9$ and $m = 6$, respectively. In figures, triangular marks on the lines show the weights which are calculated from the proposed weight formula in (4.7) and corresponding RE values (RE_1). The top of the lines are circled as the optimal weights which give the highest relative efficiency (RE_2). It is seen from the figures that; the proposed weighted estimator is much more efficient than the classical estimator of BGRSS for all cases of m when the underlying distribution is asymmetric. It can be said that weighted estimator is more

efficient when the skewness of distribution is greater. Also, Wo, Wf, RE_1 and RE_2 values are increasing while m is increasing.

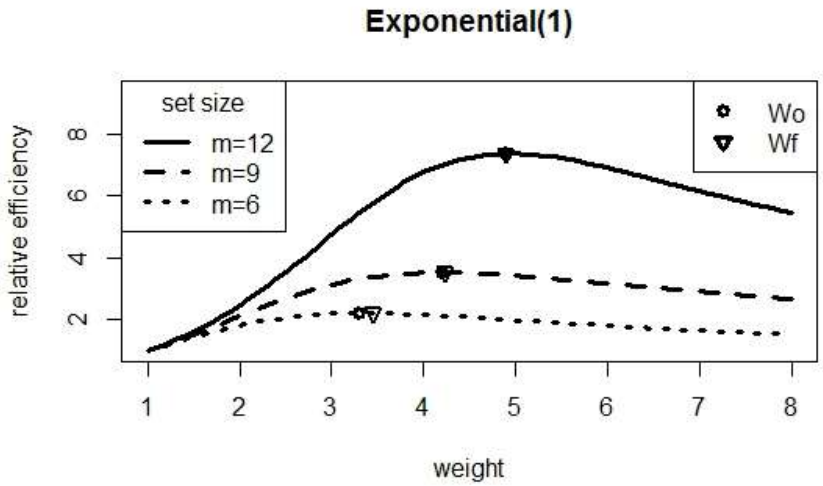


Figure 4.1 RE values for weighted estimator in BGRSS for Exponential (1), where skewness of the distribution $s = 1.9787$.

In the Figure (4.1), optimal weights and formulated weights are overlapped for the case of $m = 12$ ($RE_1 = 7.3494$, $RE_2 = 7.3501$) and $m = 9$ ($RE_1 = 3.5197$, $RE_2 = 3.5200$). When the set size is equals to 6, formulated weight (Wf) and optimal weight (Wo) are not so close which are 3.45 and 3.30, respectively. However, relative efficiencies are almost equal ($RE_1 = 2.2110$ and $RE_2 = 2.2138$). Also, the formulated weight is getting closer to the optimal weight point when m is increasing.

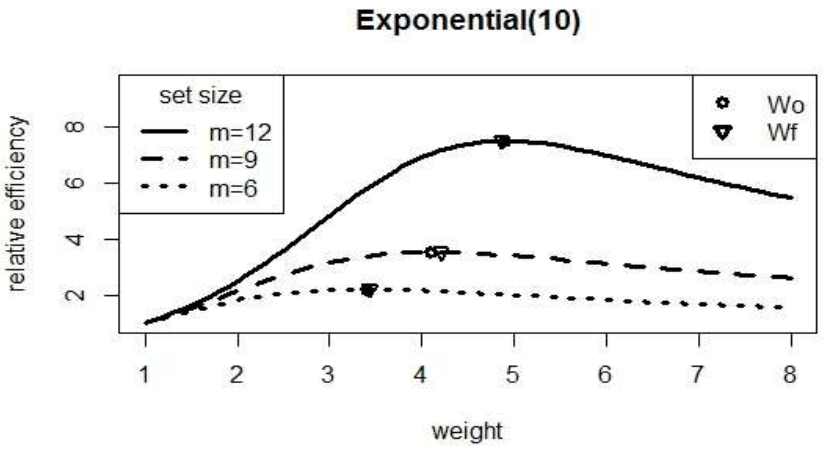


Figure 4.2 RE values for weighted estimator in BGRSS for Exponential (10) where skewness of the distribution $s = 1.9475$.

For the Exponential (10) distribution in Figure (4.2), cases of $m = 12$ and $m = 6$ give the same results for Wf and Wo. When the set size is 9, weights are not equal (Wf=4.21, Wo=4.10) contrary RE_1 and RE_2 are close to each other ($RE_1 = 3.5197$, $RE_2 = 3.5219$).

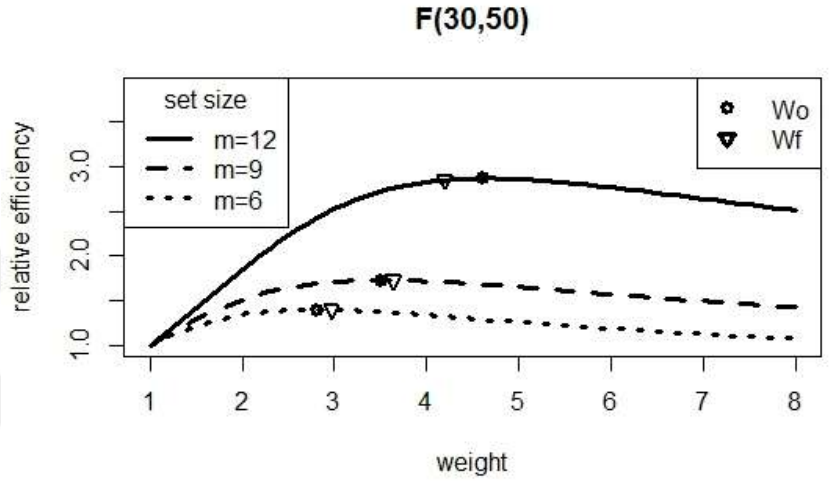


Figure 4.3 RE values for weighted estimator in BGRSS for F (30, 50), where skewness of the distribution $s = 0.9469$.

In Figure (4.3), Wo and Wf are not overlapped. Even though the optimal weight and formulated weight are not close, relative efficiencies are highly close to each other. Weight and relative efficiency values for $m = 12$ are $Wf = 4.20$, $Wo = 4.60$ and $RE_1 = 2.8478$, $RE_2 = 2.8655$, respectively. For $m = 9$, Wf=3.64, Wo=3.50 and $RE_1 = 1.7300$, $RE_2 = 1.7310$. And for the case of $m = 6$, Wf=2.97, Wo=2.80 and $RE_1 = 1.3985$, $RE_2 = 1.4004$.

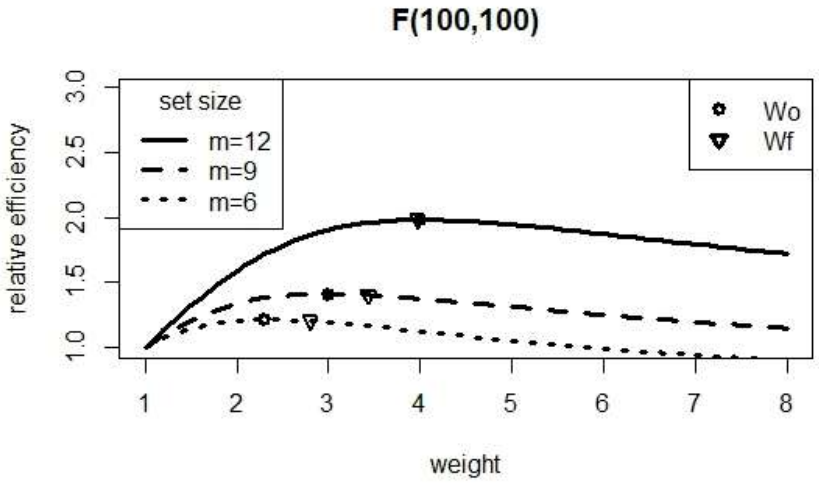


Figure 4.4 RE values for weighted estimator in BGRSS for F (100, 100), where skewness of the distribution $s = 0.6223$.

In Figure (4.4), Wo and Wf values are almost equal for $m = 12$ ($RE_1 = 1.9798$, $RE_2 = 1.9800$).

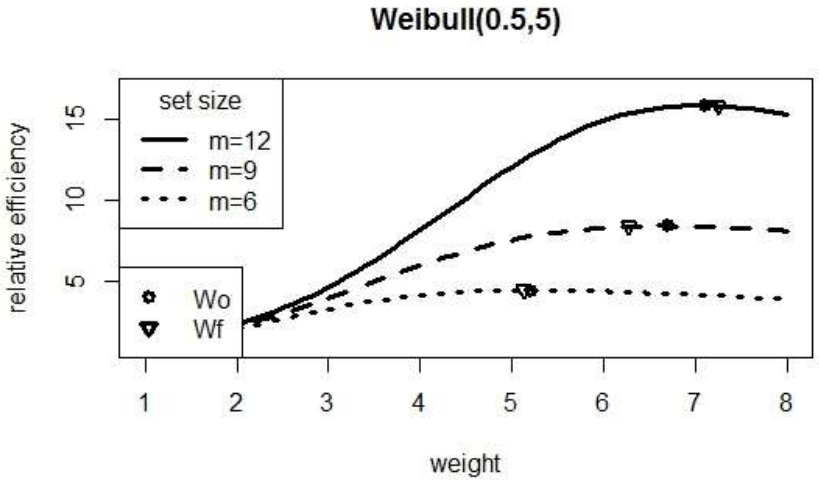


Figure 4.5 RE values for weighted estimator in BGRSS for Weibull (0.5, 5), where skewness of the distribution $s = 6.7681$.

In Figure (4.5), for the case of $m = 12$, formulated and optimal weights are 7.00 and 7.25, respectively ($RE_1 = 15.7589$, $RE_2 = 15.7850$). When $m = 9$, formulated weight is 6.28 and optimal weight is 6.70. In spite of 0.42 difference between the weights, relative efficiencies are affected only 3.5% ($RE_1 = 8.3708$ and $RE_2 =$

8.4053). Another point is that for the case of $m = 12$, the proposed weighted estimator is 15.76 times more efficient than the classical estimator of BGRSS.

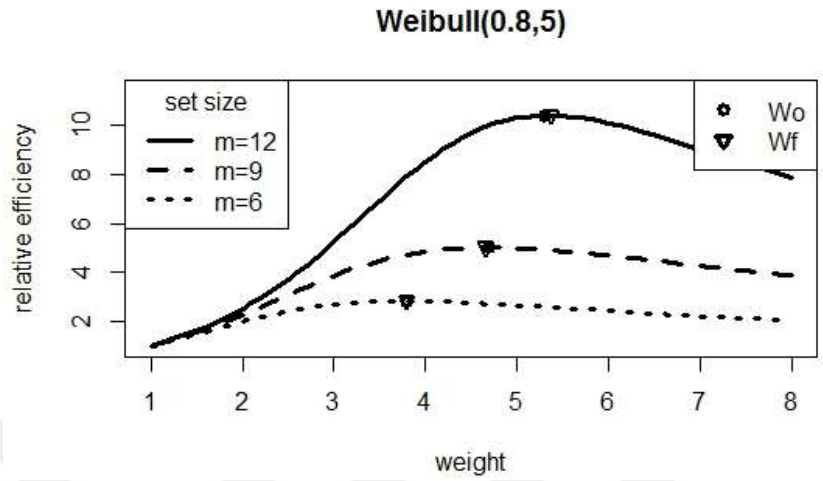


Figure 4.6 RE values for weighted estimator in BGRSS for Weibull (0.8, 5), where skewness of the distribution $s = 2.8182$.

Weight formula displayed the best performance on the Weibull (0.8, 5) distribution with skewness coefficient 2.8182. As it seen on Figure (4.6), weight and RE values are almost equal for the cases of $m = 12$ and $m = 9$ (Wf=5.38, Wo=5.30 and $RE_1 = 10.4069$, $RE_2 = 10.4067$ for $m = 12$, Wf=4.66, Wo=4.70 and $RE_1 = 5.0232$ $RE_2 = 5.0234$) for $m = 9$). Weight and RE values are equal when $m = 6$ $RE_1 = 2.8436$ and $RE_2 = 2.8437$) (Wf and Wo values are equal to 3.80).

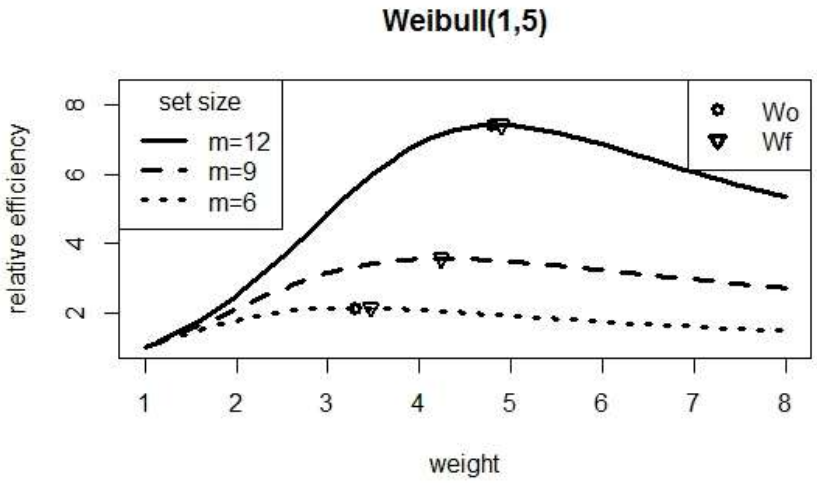


Figure 4.7 RE values for weighted estimator in BGRSS for Weibull (1, 5), where skewness of the distribution $s = 2.00$.

Figure (4.7) shows that Wo and Wf values are approximate values for $m = 12$ and $m = 9$. For the case of $m = 6$, it seen that weighted estimator is considerably more efficient than classical estimator (where $\mathrm{Wf} = 3.47$ and $RE_1 = 2.1353$). Moreover, for the case of $m = 12$, weighted estimator gives the similar result ($\mathrm{Wf} = 4.90$ and $RE_1 = 7.4100$). When m equals to 9, optimal weight is 4.20 and formulated weight is 4.24 while RE_2 equals to 3.5603 and RE_1 equals to 3.5608.

CHAPTER FIVE
SIMULATION STUDY

In this section, we design a simulation study with R software for analyzing the performance of the mean estimator of Adjusted-BGRSS (A-BGRSS) method. General comparison of the MSE's of the mean estimators based on some modified methods such as ERSS, MRSS, PRSS, BGRSS and A-BGRSS is presented. In order to compare the methods, we simulate data sets from some symmetric and asymmetric distributions, namely, exponential, Weibull, chi-square, normal, gamma, beta, F, uniform. The estimations for population mean of 10,000 sample are compared for set sizes $m = 4, 5, 6$ and $k = 1, 2, 3, 4, 5, 10$. Set size m and k are chosen specifically to obtain equal sample sizes from methods.

5.1 Cost Efficiency of Adjusted-Balanced Groups Ranked Set Sampling

Classical and modified RSS methods are useful when the measurement the concerning feature of the units is expensive or time consuming but ranking the units with expert judgment or a concomitant variable is more cheap and easy. Dell and Clutter (1972) introduced relative cost efficiency (RCE) formula to compare the sampling methods for given measurement costs (cost of measurement the concerning variable X, C_X) and ranking costs (C_Y) which is the cost of using expert or measuring the concomitant variable Y. The RCE formula can be described as the multiplication of relative costs (RC) and relative efficiencies (RE) as,

$$RCE = RC \cdot RE\,, \tag{5.1}$$

where $RC = \dfrac{Cost(\hat{\mu}_{A-BGRSS})}{Cost(\hat{\mu}_{BGRSS})}$ and $RE = \dfrac{MSE(\hat{\mu}_{A-BGRSS})}{MSE(\hat{\mu}_{BGRSS})}$.

Suppose that we select $6k_1$ sample by SRS of size m_1 in A-BGRSS, where k_1 is a positive integer ($k_1 \geq 1$). Then, we have $6k_1m_1$ sample units in our sampling scheme. In BGRSS, to provide for working with even set size, we need to chose k_2 is even. Under these assumptions, we have the following RC formula as

$$RC = \frac{Cost(\hat{\mu}^{E}_{A-BGRSS})}{Cost(\hat{\mu}^{E}_{BGRSS})} = \frac{m_1 6k_1 * C_Y + 6k_1 * C_X}{m_2 3k_2 * C_Y + (m_2 + k_2) * C_X} \leq 1 \tag{5.2}$$

Especially, when the measurement of interested variable X is expensive or time consuming, A-BGRSS keep down the cost in comparison with BGRSS for even set size. For even set size, it is clear that, the cost of A-BGRSS method is equal or smaller than the cost of BGRSS method for all values of measuring and ranking costs (C_X and C_Y). Also, remember that adjusted model has set size flexibility.

For example, when we want to obtain the sample of size 6, the set size must be 6 in BGRSS, and then we need 36 units for ranking in a cycle. However, in adjusted model for the same sample size, we can achieve this with 24 units instead of 36 by selecting the set size is 4. That is, the ranking cost could be reduced as follows;

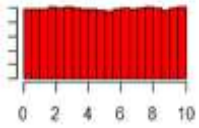
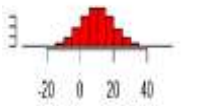
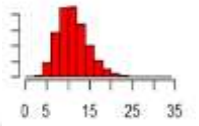
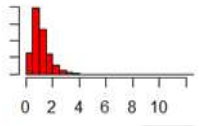
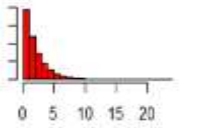
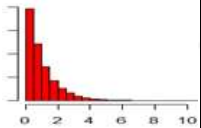
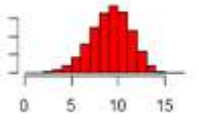
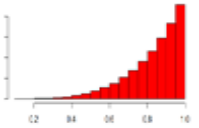
$$RC = \frac{Cost(\hat{\mu}^{E}_{A-BGRSS})}{Cost(\hat{\mu}^{E}_{BGRSS})} = \frac{24C_Y + 6C_X}{36C_Y + 8C_X}.$$

In comparison of the A-BGRSS and BGRSS with respect to RCE, we construct a simulation study for fixed and even set size for each sample design. In the simulation study, we define the cost ratio (*CR*) of the interested variable and concomitant variable as

$$CR = \frac{C_X}{C_Y}, \tag{5.3}$$

and we compare the cost efficiencies of BGRSS and A-BGRSS using the *RCE*. As it is seen on Table 5.1, A-BGRSS is more cost-efficient than BGRSS method for all cases of the *CR* and sample size. The cost efficiency RCE is increasing when the cost ratio is increasing. For example, if the measuring of the variable of interest is 20 times expensive than the concomitant variable, estimator of adjusted model is 1.24501 times more cost effective than estimator of BGRSS for chi-Square (2) distribution in the case of sample size 6.

Table 5.1 RCE values for the cases of $m = 6$ and $CR = 2, 5, 10, 20$.

Distribution		Cost ratio	Sample size					
			6	12	18	24	30	60
	Uniform (0, 10)	2	1.01586	1.01460	1.01500	1.02487	1.02470	1.02406
		5	1.07980	1.07845	1.07888	1.08937	1.08919	1.08851
		10	1.13307	1.13167	1.13211	1.14312	1.14293	1.14222
		20	1.17816	1.17669	1.17716	1.18861	1.18841	1.18767
	Normal (10, 10)	2	1.05568	1.06266	1.05419	1.05055	1.05969	1.04838
		5	1.12213	1.12954	1.12054	1.11667	1.12638	1.11436
		10	1.17749	1.18528	1.17583	1.17177	1.18196	1.16935
		20	1.22434	1.23244	1.22261	1.21839	1.22899	1.21587
	Gamma (10, 0.9)	2	1.05801	1.05161	1.05667	1.06018	1.05513	1.06381
		5	1.12460	1.11780	1.12317	1.12691	1.12154	1.13077
		10	1.18009	1.17295	1.17859	1.18251	1.17688	1.18656
		20	1.22704	1.21962	1.22548	1.22956	1.22370	1.23377
	F (10, 20)	2	1.06816	1.06473	1.07269	1.07491	1.07613	1.07221
		5	1.13539	1.13174	1.14020	1.14256	1.14386	1.13970
		10	1.19141	1.18759	1.19646	1.19894	1.20030	1.19593
		20	1.23882	1.23484	1.24407	1.24664	1.24806	1.24351
	Chi-Square (2)	2	1.07350	1.07202	1.07289	1.07460	1.07453	1.07975
		5	1.14107	1.13949	1.14041	1.14223	1.14216	1.14770
		10	1.19737	1.19571	1.19668	1.19859	1.19851	1.20433
		20	1.24501	1.24329	1.24430	1.24628	1.24620	1.25225
	Exponential (1)	2	1.06670	1.07121	1.07533	1.07388	1.07545	1.07486
		5	1.13383	1.13863	1.14301	1.14147	1.14314	1.14251
		10	1.18978	1.19481	1.19941	1.19779	1.19954	1.19888
		20	1.23712	1.24235	1.24713	1.24545	1.24727	1.24658
	Weibull (5, 10)	2	1.05628	1.04798	1.05071	1.05396	1.05245	1.05266
		5	1.12276	1.11393	1.11684	1.12029	1.11868	1.11891
		10	1.17816	1.16890	1.17195	1.17557	1.17388	1.17412
		20	1.22504	1.21541	1.21858	1.22234	1.22059	1.22084
	Beta (5, 1)	2	1.05619	1.05952	1.06402	1.06584	1.06500	1.06814
		5	1.12266	1.12620	1.13098	1.13292	1.13203	1.13536
		10	1.17806	1.18177	1.18679	1.18882	1.18788	1.19138
		20	1.22493	1.22879	1.23401	1.23612	1.23515	1.23879

5.2 Comparison with Other Modified Ranked Set Sampling Procedures

In this subsection, we compare the efficiency of A-BGRSS and SRS for the cases of $m = 4, 6$ and $k = 1, 2, 3, 4, 5, 10$ in Table 5.2 and Table 5.3. It can be said that A-BGRSS is more efficient than SRS for the cases of $m = 4$. When the underlying distribution is symmetric and $m = 6$, A-BGRSS is more efficient than SRS. When

the underlying distribution is asymmetric, efficiency of A-BGRSS is decreasing while k is increasing.

Table 5.2 Relative efficiency between A-BGRSS and SRS for $m = 4$.

Distribution	k=2	k=4	k=6
Uniform (0, 10)	2.672677	2.715108	2.789075
Normal (10, 10)	2.253815	2.249663	2.232386
Gamma (10, 0.9)	2.175418	2.116074	2.092769
F (10, 20)	1.676763	1.671505	1.531280
Chi-square (2)	1.514796	1.535250	1.389911
Exponential (1)	1.578313	1.536399	1.363532
Weibull (5, 10)	2.192157	2.315184	2.243014
Beta (5, 1)	1.892861	1.862145	1.781022

Table 5.3 Relative efficiency between A-BGRSS and SRS for $m = 6$.

Distribution	k=1	k=2	k=3	k=4	k=5	k=10
Uniform (0, 10)	4.501722	4.165675	4.402158	4.298143	4.171871	4.324987
Normal (10, 10)	2.758471	2.741217	2.803662	2.808038	2.773932	2.740627
Gamma (10, 0.9)	2.502156	2.496517	2.400500	2.376692	2.259123	1.977107
F (10, 20)	1.694404	1.463444	1.350615	1.269565	1.222680	0.879060
Chi-square (2)	1.481031	1.277589	1.157388	0.944049	0.919952	0.605265
Exponential (1)	1.542563	1.333179	1.123413	0.942737	0.940057	0.606477
Weibull (5, 10)	2.832444	2.746319	2.777904	2.850799	2.670887	2.673251
Beta (5, 1)	2.154033	1.812618	1.641196	1.472876	1.312453	0.956431

Also, MSEs of the mean estimators based on ERSS, MRSS, PRSS, BGRSS and A-BGRSS are compared. In Table 5.4, A-BGRSS give relatively good results in all cases. As it is expected, MSE results of BGRSS are smaller than adjusted model. However, BGRSS is more efficient in terms of MSE; we have shown previously that adjusted model is more efficient in terms of *RCE*. Because in BGRSS there are extra k units are measured for the same sample size.

Table 5.4 MSE values of ERSS, MRSS, PRSS, BGRSS, A-BGRSS for $m = 6$.

Distribution	Sample size	ERSS	MRSS	PRSS (p=0.3)	BGRSS	A-BGRSS
Uniform (0, 10)	6	0.25574	0.50174	0.42067	0.32407	0.34560
	12	0.12566	0.24537	0.20689	0.15515	0.16566
	18	0.08268	0.16699	0.13869	0.10597	0.11311
	24	0.06429	0.12912	0.10749	0.08054	0.08514
	30	0.05276	0.10349	0.08528	0.06594	0.06998
	60	0.02573	0.05087	0.04251	0.03178	0.03362

Table 5.4 MSE values of ERSS, MRSS, PRSS, BGRSS, A-BGRSS for $m = 6$. (continue)

Distribution	Sample size	ERSS	MRSS	PRSS (p=0.3)	BGRSS	A-BGRSS
Normal (10, 10)	6	7.08953	4.10726	4.62456	5.90080	6.05535
	12	3.57541	2.05599	2.41806	2.95085	3.00825
	18	2.32608	1.34743	1.53770	1.94085	1.99450
	24	1.73839	1.02589	1.16718	1.44318	1.48822
	30	1.37652	0.81887	0.94322	1.16434	1.19032
	60	0.67624	0.41677	0.46580	0.58370	0.60316
Gamma (10, 0.9)	6	1.04320	0.54792	0.59269	0.79699	0.81606
	12	0.61695	0.32432	0.30770	0.41265	0.42510
	18	0.45715	0.22713	0.20331	0.28129	0.28839
	24	0.36033	0.19591	0.15726	0.21545	0.22016
	30	0.32123	0.16435	0.12835	0.17890	0.18368
	60	0.23156	0.11750	0.07145	0.10246	0.10434
F (10, 20)	6	0.07121	0.02445	0.02083	0.04209	0.04269
	12	0.04866	0.01739	0.01163	0.02383	0.02424
	18	0.04141	0.01570	0.00871	0.01788	0.01806
	24	0.03638	0.01436	0.00695	0.01403	0.01414
	30	0.03318	0.01353	0.00598	0.01172	0.01180
	60	0.02782	0.01224	0.00417	0.00803	0.00811
Chi-square (2)	6	0.84339	0.30218	0.21315	0.42648	0.43038
	12	0.64707	0.25541	0.12855	0.25944	0.26218
	18	0.55219	0.23171	0.09602	0.19044	0.19230
	24	0.53230	0.21943	0.08096	0.17498	0.17640
	30	0.48301	0.21089	0.07072	0.14509	0.14627
	60	0.43761	0.20221	0.05329	0.11190	0.11227
Exponential (1)	6	0.22188	0.07827	0.05429	0.10935	0.11106
	12	0.15683	0.06223	0.03170	0.06412	0.06485
	18	0.13802	0.05714	0.02406	0.04964	0.05001
	24	0.13212	0.05519	0.02058	0.04368	0.04406
	30	0.12083	0.05330	0.01785	0.03609	0.03635
	60	0.10745	0.04965	0.01288	0.02671	0.02693
Weibull (5, 10)	6	0.29935	0.19159	0.21042	0.25119	0.25762
	12	0.15967	0.10095	0.10717	0.13115	0.13558
	18	0.10653	0.06754	0.07181	0.08436	0.08698
	24	0.08463	0.05165	0.05320	0.06268	0.06442
	30	0.07218	0.04321	0.04509	0.05233	0.05386
	60	0.03949	0.02433	0.02189	0.02668	0.02745
Beta (5, 1)	6	0.00289	0.00145	0.00110	0.00149	0.00153
	12	0.00202	0.00105	0.00060	0.00087	0.00089
	18	0.00178	0.00094	0.00043	0.00067	0.00068
	24	0.00165	0.00085	0.00034	0.00056	0.00057
	30	0.00153	0.00081	0.00030	0.00049	0.00049
	60	0.00138	0.00072	0.00019	0.00035	0.00035

We note that, there is no comparison for BGRSS in Table 5.5 and Table 5.6 because BGRSS cannot be used with the set sizes such as 4, 5, 7, etc. We can see in Table 5.5 and Table 5.6, A-BGRSS method gives the best results in some cases such as, F (10, 20), exponential (1), chi-square (2), beta (5, 1) and relatively good results in other cases. It is note that in Table 5.6, sample size is 30 for all modified methods.

Table 5.5 MSE values of ERSS, MRSS, PRSS, A-BGRSS for $m = 4$.

Distribution	sample size	ERSS	MRSS	PRSS (p=0.3)	A-BGRSS
Uniform (0, 10)	12	0.22604	0.33575	0.34160	0.26246
	24	0.11071	0.16436	0.16137	0.12792
	36	0.07277	0.11002	0.10993	0.08521
Normal (10, 10)	12	4.12555	3.02702	3.01268	3.78122
	24	2.03086	1.52684	1.53911	1.89078
	36	1.36685	0.99062	0.99479	1.24090
Gamma (10, 0.9)	12	0.57217	0.39813	0.40202	0.47812
	24	0.31306	0.22523	0.22542	0.24212
	36	0.22310	0.16247	0.16240	0.16614
F (10, 20)	12	0.03231	0.01718	0.01727	0.02130
	24	0.01885	0.01204	0.01203	0.01077
	36	0.01553	0.01039	0.01046	0.00782
Chi-Square (2)	12	0.35734	0.21420	0.21256	0.21596
	24	0.22845	0.16148	0.16028	0.10678
	36	0.19636	0.14451	0.14403	0.07979
Exponential (1)	12	0.08535	0.05167	0.05162	0.05202
	24	0.05943	0.04029	0.04034	0.02767
	36	0.04871	0.03562	0.03546	0.02000
Weibull (5, 10)	12	0.17663	0.14032	0.14013	0.16598
	24	0.08977	0.06939	0.07019	0.07940
	36	0.06133	0.04886	0.04855	0.05449
Beta (5, 1)	12	0.00126	0.00101	0.00100	0.00086
	24	0.00082	0.00066	0.00066	0.00045
	36	0.00066	0.00057	0.00056	0.00031

Table 5.6 MSE values of ERSS, MRSS, PRSS, A-BGRSS for $m = 5$.

Distribution	ERSS	MRSS	PRSS (p=0.3)	A-BGRSS
Uniform (0, 10)	0.06507	0.11763	0.10549	0.08350
Normal (10, 10)	1.48158	0.95744	1.04390	1.32430
Gamma (10, 0.9)	0.28482	0.18168	0.15164	0.18118
F (10, 20)	0.02445	0.01405	0.00838	0.01010
Chi-square (2)	0.33574	0.21607	0.10843	0.10874
Exponential (1)	0.08439	0.05562	0.02807	0.02761
Weibull (5, 10)	0.06993	0.04796	0.04860	0.05639
Beta (5, 1)	0.00110	0.00084	0.00044	0.00039

CHAPTER SIX
CONCLUSION

In this thesis, we introduce two main improvements on the BGRSS which is an efficient modified RSS method for estimating the population mean. The first improvement is on the procedure of BGRSS, which is called Adjusted-BGRSS. This method provides flexibility on choosing the set size and reduces the sampling cost. In the simulation study, it is shown that Adjusted-BGRSS is cost-effective than BGRSS for all cases of sample size and cost ratio on various distributions. It is seen that relative cost efficieny between Adjusted-BGRSS and BGRSS is increasing when the cost ratio is increasing. Also we compared Adjusted-BGRSS with other modified RSS methods, namely, ERSS, MRSS, PRSS and BGRSS with a Monte-Carlo simulation study on R software. In the simulation study, 10,000 samples are compared for some symmetric and asymmetric distributions such as exponential, Weibull, chi-square, normal, etc. For all methods in comparison, we study with fix set size and sample size. Cycle sizes in each method are chosen specifically to obtain equal sample size. Results are showed that Adjusted-BGRSS gives best results for small set size cases. It can be said that new method could be an alternative for modified RSS methods.

As a second improvement, we propose a weighted mean estimator of the population mean for asymmetric distributions. In order to improve the performance of mean estimation in asymmetric distributed population, we propose a new weighted estimator with a general formula to calculate the weight values for different set sizes and skewness coefficients of different populations. It is shown numerically that the weights calculated via the given formula converges the optimum weights values providing best efficiency. Also the simulation results show that, the relative efficiency between weighted estimator with the weight formula and classical BGRSS estimator is increasing while set size and the skewness coefficient is increasing. Numerical results for Weibull (0.5, 5) prove that weighted estimator with the proposed weight formula is 15.7589 times more efficient than classical estimator of BGRSS.

REFERENCES

Abu-Dayyeh, W., Samawi, H. M. & Elies, K. (2002). Weighted extreme ranked set sample for skewed population. *Calcutta Statistical Association Bulletin*, 53, 53-63.

Al-Nasser, A. D. (2007). L ranked set sampling: a generalization procedure for robust visual sampling. *Communications in Statistics - Simulation and Computation*, 36 (1), 33-43.

Al-Naseer, A. D. & Mustafa, A. B. (2009). Robust extreme ranked set sampling. *Journal of Statistical Computation and Simulation*, 79 (7), 859-867.

Al-Omari, A. I., & Raqab, M. Z. (2013). Estimation of the population mean and median using truncation-based ranked set samples. *Journal of Statistical Computation and Simulation*, 83 (8), 1453-1471.

Al-Omari, A. I., & Al-Saleh, M. F. (2009). Quartile double ranked set sampling for estimating the population mean. *Economic Quality* Control, 24 (2), 243-253.

Al-Omari, A. I. & Bouza, C. N. (2014). Review of ranked set sampling: modifications and applications. *Revista Investigación Operacional*, 3, 215-240.

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

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

Balci, S., Akkaya, A. D. & Ulgen, B. E. (2013). Modified maximum likelihood estimators using ranked set sampling. *Journal of Computational and Applied Mathematics*, 238, 171-179.

David, H. A., & Nagaraja, H. N. (2003). *Order statistics* (3rd ed.) New York: John Wiley.

Dell, T. R., & Clutter, J. L. (1972). Ranked set sampling theory with order statistics background. *Biometrics*, 28, 545-553.

Halls, L. S., & Dell, T. R. (1966). Trial of ranked set sampling for forage yields. *Forest Science*, 12 (1), 22-26.

Haq, A., Brown, J., Moltchanova, E. & Al-Omari, A. I. (2013). Partial ranked set samplin. *Environmetrics*, 24 (3), 201-207.

Haq, A., Brown, J., Moltchanova, E. & Al-Omari, A. I. (2014). Mixed ranked set sampling design. *Journal of Applied Statistics*, 41 (10), 2141-2156.

Jemain, A. A., & Al-Omari, A. I. (2006a). Double quartile ranked set samples. *Pakistan Journal of Statistics,* 22 (3), 217-228.

Jemain, A. A., & Al-Omari, A. I. (2006b). Double percentile ranked set samples for estimating the population mean. *Advances and Applications in Statistics,* 6 (3), 261-276.

Jemain, A. A., & Al-Omari, A. I. (2006c). Multistage median ranked set samples for estimating the population mean. *Pakistan Journal of Statistics,* 22 (3), 195-207.

Jemain, A. A., & Al-Omari, A. I. (2007a). Multistage percentile ranked set samples. *Advances and Applications in Statistics,* 7 (1), 127-139.

Jemain, A. A., & Al-Omari, A. I. (2007b). Multistage quartile ranked set samples. *Pakistan Journal of Statistics,* 23 (1), 11-22.

Jemain, A. A., Al-Omari, A. & Ibrahim, K. (2008). Some variations of ranked set sampling. *Electronic Journal of Applied Statistical Analysis*, 1, 1-15.

Kaur, A., Patil, G. P., Sinha, A. K. & Taillie, C. (1995). Ranked set sampling: an annotated bibliography. *Environmental and Ecological Statistics*, 2, 25–54.

Mahdizadeh, M., & Arghami, N. R. (2012). Quantile estimation using ranked set samples from a population with known mean. *Communications in Statistics-Simulation and Computation*, 41 (10), 1872-1881.

McIntyre, G. A. (1952). A method for unbiased selective sampling, using ranked sets. *Australian Journal of Agricultural Research*, 3, 385-390.

Muttlak, H. A. (1997). Median ranked set sampling. *Journal of Applied Statistical Sciences*, 6 (4), 577-586.

Muttlak, H. A. (2003a). Investigating the use of quartile ranked set samples for estimating the population mean. *Applied Mathematics and Computation*, 146, 437-443.

Muttlak, H. A. (2003b). Modified ranked set sampling methods. *Pakistan Journal of Statistics*, 19 (3), 315-323.

Muttlak, H. A. & Abu-Dayyeh, W. A. (2004). Weighted modified ranked set sampling. *Applied Mathematics and Computation*, 151, 645-657.

Patil, G. P. (1995). Editorial: Ranked set sampling. *Environmental and Ecological Statistics*, 2, 271-285.

Samawi, H. M., Ahmed, M. S. & Abu-Dayyeh, W. (1996). Estimating the population mean using extreme ranked set sampling. *Biometrical Journal*, 38, 577-586.

Sinha, Bimal K., Sinha, Bikas K. & Purkayastha, S. (1996). On some aspects of ranked set sampling for estimation of normal and exponential parameters. *Statistics and Decisions*, 14, 223-240.

Stokes, S. L. (1977). Ranked set sampling with concomitant variables. *Communications in Statistics - Theory and Methods*, 6 (12), 1207–1211.

Stokes, S. L. (1980). Inferences on the correlation coef- ficient in bivariate normal populations from ranked set samples. *Journal of the American Statistical Association*, 75, 989–995.

Takahashi, K., & Wakimoto, K. (1968). On unbiased estimates of the population mean based on the sample stratified by means of ordering. *Annals of the Institute of Statistical Mathematics*, 20, 1-31.

Takahashi, K. (1970). Practical note on estimation of population means based on samples stratified by means of ordering. *Annals of the Institute of Statistical Mathematics*, 22, 421-428.

Wolfe, D. A. (2012). Ranked set sampling: Its relevance and impact on statistical inference. *International Scholarly Research Network, Probability and Statistics*, 2012, 1-32.

Yu, P. L. H., & Tam, C. Y. C. (2002). Ranked set sampling in the presence of censored data. *Environmetrics*, 13 (4), 379-396.