# QUALITY CONTROL CHARTS BASED ON RANKED SET SAMPLING UNDER VARIOUS SYMMETRIC DISTRIBUTIONS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

DİDEM EGEMEN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
STATISTICS

JULY 2013

Approval of the thesis:

**QUALITY CONTROL CHARTS BASED ON RANKED SET SAMPLING UNDER VARIOUS SYMMETRIC DISTRIBUTIONS**

submitted by **DİDEM EGEMEN** in partial fulfillment of the requirements for the degree of **Master of Science in Statistics Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen
Dean, Graduate School of **Natural and Applied Sciences** ________________

Prof. Dr. İnci Batmaz
Head of Department, **Statistics** ________________

Assoc. Prof. Dr. Barış Sürücü
Supervisor, **Statistics Dept., METU** ________________

**Examining Committee Members:**

Prof. Dr. Olcay Arslan
Statistics Dept., Ankara Üniversitesi ________________

Assoc. Prof. Dr. Barış Sürücü
Statistics Dept., METU ________________

Prof. Dr. Birdal Şenoğlu
Statistics Dept., Ankara Üniversitesi ________________

Assist. Prof. Dr. Berna Burçak Başbuğ Erkan
Statistics Dept., METU ________________

Assist. Prof. Dr. Ceylan Talu Yozgatlıgil
Statistics Dept., METU ________________

**Date:** ________________

**I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.**

Name, Last Name: DİDEM EGEMEN

Signature :

# ABSTRACT

QUALITY CONTROL CHARTS BASED ON RANKED SET SAMPLING UNDER VARIOUS SYMMETRIC DISTRIBUTIONS

Egemen, Didem
M.S., Department of Statistics
Supervisor: Assoc. Prof. Dr. Barış Sürücü

July 2013, 86 pages

Quality of a product is a very important property for producers and service providers, and can be maintained by statistical quality control methods. One of these methods is Shewhart control chart, which is the most frequently used in the literature. Construction of accurate control limits is the main issue in this method and it strictly depends on the efficiency of the estimators used to construct control limits. In this thesis, a new sampling method ranked set sampling (RSS), which is more efficient than simple random sampling (SRS), is proposed in the estimation part of Shewhart control charts. With this aim, quality control charts constructed by using SRS and RSS are compared under three different symmetric distributions which are normal, long-tailed symmetric and short-tailed symmetric. The simulation results show that RSS provides more efficient estimators compared to SRS. Moreover, it is demonstrated by simulations that type I errors of the charts obtained by RSS technique are very close to the predefined error level. To give a better illustration of the methods, a real life example is provided at the end of the study.

Keywords: Extreme Ranked Set Sampling, Ranked Set Sampling, Shewhart Control Charts, Symmetric Distributions

# ÖZ


SİMETRİK DAĞILIMLAR ALTINDA SIRALI KÜME ÖRNEKLEMESİ YÖNTEMİYLE KALİTE KONTROL GRAFİKLERİ



Egemen, Didem

Yüksek Lisans, İstatistik Bölümü

Tez Yöneticisi: Doç. Dr. Barış Sürücü



Temmuz 2013, 86 sayfa



Bir ürünün kalitesi üreticiler ve servis sağlayıcılar için çok önemli bir özelliktir ve istatistiksel kalite kontrol metotlarıyla sağlanabilmektedir. Bu metotlardan literatürde en çok kullanılanlardan biri Shewhart kontrol grafiğidir. Bu metotdaki temel konu doğru kontrol limitlerini oluşturabilmektir ve bu da kontrol limitleri kurmak için kullanılan tahmin edicilerin etkinliğine sıkı sıkıya bağlıdır. Bu tez çalışmasında yeni bir örneklem metodu olan ve basit rastgele örneklem metodundan daha etkin sonuçlar veren sıralı küme örneklemesi önerilmiştir. Bu amaçla, basit rastgele örneklem ve sıralı küme örneklemesi metotları kullanılarak üretilen kalite kontrol grafikleri normal, uzun kuyruklu simetrik ve kısa kuyruklu simetrik dağılımlar kullanılarak karşılaştırılmıştır. Simülasyon sonuçları sıralı küme örneklemesi metodunun basit rastgele örneklem metoduna kıyasla daha etkin tahmin ediciler sağladığını göstermiştir. Dahası, sıralı küme örneklemesi metodu altında kurulan kontrol grafiklerinin birincil tip hatası, bu hatanın daha önceden belirlenen değerine daha yakın çıkmıştır. Bu metotların daha açık ifade edilebilmesi için çalışmanın sonunda gerçek bir veriyle örnek sunulmuştur.



Anahtar Kelimeler: Shewhart Kontrol Grafikleri, Sıralı Küme Örneklemesi, Simetrik Dağılımlar, Uç Sıralı Küme Örneklemesi

*To my family,*

*Behice, Mehmet and Öykü Egemen, for their ever lasting love and support.*

# ACKNOWLEDGMENTS

My deepest and heartfelt appreciation goes to Assoc. Prof. Dr. Barış Sürücü. Without his invaluable guidance and persistent help this thesis would not have been possible. I have always found a chance to take benefit from his precious knowledge and professional experience in academic life. He has always encouraged me to research new ideas and leaded me to critical thinking. It has been a great honor for me to be a student of him and work with him. With his friendly manner, he always has been more than a supervisor to me. Apart from the thesis work, I owe a very important debt to him in leading me to an academic career and always supporting me. He has always been and will be a role model for me in academic life.

Advice and comments given by thesis committee members, Prof. Dr. Olcay Arslan, Prof. Dr. Birdal Şenoğlu, Assist. Prof. Dr. B. Burçak Başbuğ Erkan, and Assist. Prof. Dr. Ceylan Talu Yozgatlıgil have been a great help in developing this thesis. Therefore, I am deeply grateful to them for their valuable time to review my thesis and their suggestions.

The numerical calculations reported in this thesis were partially performed at TUBITAK ULAKBIM, High Performance and Grid Computing Center (TRUBA resources), and partially performed at HPC resources provided by the Department of Computer Engineering, Middle East Technical University.

I would like to offer my special thanks to all members of the Department of Statistics for providing me with an excellent working environment. Furthermore, I would like to express my gratitude to all of my instructors for providing me with deep understanding and knowledge in this science. Also, special thanks to my friends in the department for their everlasting support and companionship.

Finally, I want to thank to my dear family for their endless support, encouragement, love, ... everything. And my friends who are my ray of sunshine; I love you all.

# TABLE OF CONTENTS

APPENDICES

# LIST OF TABLES

TABLES

# LIST OF FIGURES

FIGURES

# LIST OF ABBREVIATIONS

| | |
|---|---|
| ARL | Average Run Length |
| BLUE | Best Linear Unbiased Estimation/Estimator |
| CDF | Cumulative Density Function |
| CL | Central Limit |
| CLT | Central Limit Theorem |
| DRSS | Double Ranked Set Sampling/Sample |
| ERSS | Extreme Ranked Set Sampling/Sample |
| LCL | Lower Control Limit |
| LRSS | L Ranked Set Sampling |
| MDRSS | Median Double Ranked Set Sampling |
| MERSS | Moving Extreme Ranked Set Sampling |
| MLE | Maximum Likelihood Estimation/Estimator |
| MMLE | Modified Maximum Likelihood Estimation/Estimator |
| MRSS | Median Ranked Set Sampling/Sample |
| MSE | Mean Squared Error |
| ORSS | Ordered Ranked Set Sampling |
| PDF | Probability Density Function |
| RERSS | Robust Extreme Ranked Set Sampling |
| RSS | Ranked Set Sampling/Sample |
| SRS | Simple Random Sampling/Sample |
| UCL | Upper Control Limit |

# CHAPTER 1

# INTRODUCTION

Quality control is not just a bunch of techniques, it is, as stated by Caulcutt [20], a way of thinking. Montgomery [57] emphasizes that although most people define quality as some desired characteristics that a product should have, this is just the beginning of understanding this concept. First of all, what is meant by a "product"? Actually, it can be manufactured goods, such as automobiles, electronic devices, furniture, processed food, or can be services, such as health care, banking, transportation. The basic aim is to improve the quality of these products since quality makes the consumers prefer that product among the others. In today's industrialized world, there is a strong competition among factories, companies, service providers, etc. In order to have an enhanced position in this competitive business world, the importance of an improved quality should not be underestimated (Montgomery [57]).

There are many different ways to assess the quality, one of which is quality control charts. This method was first developed by Walter A. Shewhart in his research at Bell Telephone Laboratories. He was the first to apply statistical methods to quality control (Duncan [33]), that is why this chart is also named after Shewhart. Quality control charts set standards or goals, which can be called as quality, for a product and come to a decision whether or not that goal has been achieved (Duncan [33]). There may occur two types of variation in quality of a product: assignable and chance causes. Montgomery [57] defines chance causes as the inherent part of the process. Moreover, Duncan [33] mentions that chance variations behave in a random manner, and they do not show any defined pattern. However, they follow statistical laws. If only chance causes affect the process, the process is said to be under-control. On the contrary, the variation occurred by assignable causes is much larger than the one by chance causes. Therefore, the existence of assignable causes leads to a process out-of-control. Due to the fact that the variation by assignable causes is not an acceptable one, some actions should be taken in order to eliminate it. Control charts are extremely useful tools to detect this undesirable shift in the process, and to determine when to take corrective actions.

$\bar{X}$ charts (or Shewhart control charts) consist of two phases. In these phases, the control limits are set and the process is monitored, respectively. The limits can be set by using the known population parameters. In case of unknown parameters, $M$ samples of size $n$ are selected randomly from the in-control process. Then the mean of the process and the variance of the process mean are estimated by using these samples.

The most frequent assumption made while constructing control charts is normality. However, if it is desired to detect the shift in the population mean, then this assumption is no more a must. As stated by Shewhart [78], the skewness and kurtosis of the distribution of sample means approach to those of normal regardless of the underlying population distribution (by central limit theorem (CLT)). Since $\bar{X}$ chart is studied in this thesis, with the help of CLT two different non normal distributions are used in addition to the normal.

When it is aimed to construct $\bar{X}$ chart, samples of size $n$ are taken very often (Duncan [33]). Considering the cost efficiency, these frequently taken samples should not be very large. Moreover, Page [66] stated that if it is desired to detect large changes in the mean, samples of smaller sizes are taken frequently. Emerging from these points of view, if it is desired to take small samples to construct quality control charts (more specifically $\bar{X}$ charts); ranked set sampling (RSS) scheme may provide more efficient and unbiased samples.

The correctness of the information obtained for the population strictly depends on the unbiasedness and efficiency of the sample. Therefore, it is of crucial importance to use efficient sampling methods so as to achieve better statistical inference. Simple random sampling (SRS), systematic sampling, and stratified sampling are the most and very well known sampling techniques used in practice. Ranked set sampling (RSS) is among one of the popular sampling methods and known to be more efficient than SRS.

Ranked set sampling (RSS) is a sampling method which provides great observational economy especially when the quantification of sample units is quite expensive (Patil et al. [67]-[68]). This sampling method was first introduced by McIntyre [55] and developed further by Halls and Dell [38]. RSS is best for the cases when measurement of the units are costly and time consuming. In the basic theory of RSS, it is assumed that the population of interest is infinite and there exist easily and cheaply ranked units without actual measurement (Chen et al. [27]). RSS is applicable either when the latent values of the units are easy to rank by visual inspection, or when there exist concomitant variables which can easily be ranked (Chen et al. [27]). Concomitant variables are not the variables of interest but have a strong correlation with the variable which is the major concern of the study.

It must be noted that the samples obtained by RSS hold the information regarding both measurements on the variable of interest and ranks of the variables. This proves that RSS contains more information than simple random sampling (SRS). RSS estimators are unbiased, and they have lower mean squared errors (MSE) compared to SRS estimators.

Since RSS depends on visual inspection, there may occur some imperfect ranking errors. To overcome this problem, lots of modifications have been developed, which enabled RSS to be applicable to a wider range of areas than originally intended (Chen et al. [27]). Extreme ranked set sampling (ERSS) is one of these modifications which is more robust to imperfect ranking errors since it only identifies first and last ordered units for quantification (see McIntyre [55], Samawi et al. [74], and Shaibu and Muttlak [76]). On the other hand, ERSS has two major drawbacks: First, there may be outliers in the sample. Second, distribution of the population of interest may be skewed. In these cases, RSS is more efficient than ERSS (Patil et al. [67]). However, in this thesis, symmetric distributions are taken into consideration which makes the second drawback inconsiderable.

Estimation of location-scale parameters for RSS and ERSS has been discussed in many studies; see for example Ozturk [64]-[65] for RSS and Shaibu and Muttlak [76]-[77]) for ERSS. However, in the estimation processes of RSS and ERSS, there arise some problems which may be overcome by using the modified maximum likelihood estimation (MMLE) method proposed by Tiku [92] and Tiku and Suresh [97]. In this thesis, the modified maximum likelihood estimators for location and scale parameters for normal, long tailed symmetric and short tailed symmetric distributions under RSS scheme are obtained. Moreover, location and scale parameters are obtained by using both the MMLE method and the best linear unbiased estimation (BLUE).

In this thesis, efficiencies of quality control charts under SRS and RSS and its modification, ERSS are compared under three symmetric distributions; normal, long-tailed symmetric and short-tailed symmetric distributions. In Chapter 2, RSS and ERSS methods are compared with their SRS counterparts for constructing Shewhart control charts under normal distribution. Chapter 3 considers the long-tailed symmetric distribution. In this chapter, parameter estimators are obtained by using MMLE method under both the SRS and RSS schemes, and construction of $\bar{X}$ charts are discussed. Since normality assumption is not satisfied, three-moment t-approximation is used to construct control limits. Chapter 4 discusses construction of control limits for Shewhart control charts under short-tailed symmetric distribution. Moment approximation is also considered when the normality is not satisfied. The parameters under both SRS and RSS are obtained through MMLE. Chapter 5 gives one real data application and one simulated example on the methods covered in previous chapters. Finally, Chapter 6 gives concluding remarks and discusses possible future works.

## 1.1 LITERATURE REVIEW

### 1.1.1 QUALITY CONTROL CHARTS

Although quality control dates back to 1700s, statistical quality control can be considered as a newly developed method. It was first proposed by Shewhart in 1931 at Bell Telephone laboratories. At that time, the importance of this method could not grasped by the industry. However, it was no later than the World War II, the statistical quality control became very popular (Montgomery [57]). This is due to the fact that controlling and improving quality of products issues gained more importance during the war time. After the war, the American Society for Quality Control was established in 1948 in order to foster the use of statistical quality control in all kinds of products.

Since then, various forms of Shewhart control charts have been developed. For instance, in the classical Shewhart control chart, the control limits are constructed under normality assumption. However, in the case of violation of this assumption, Yourstone and Zimmer [106] showed the improperness of this method. Therefore, a great number of studies have been conducted to enable the use of control charting procedure for non-normal populations. Chou et al. [30] and Spedding and Rawlings [81] proposed the use of transformation to satisfy normality assumption. Ferrel [36] studied the lognormal distribution whereas Nelson [61] proposed a method for the Weibull distribution. Tagaras [88], Elsayed and Chen [34], and Wu [105] worked on developing asymmetric control limits for skewed distributions. Choobineh and Ballard [29] proposed the weighted variance method for skewed distributions; see also Bai and Choi [13], and Amhemad [45]. Chan and Cui [23] suggested a new method for eliminating the skewness of a distribution. Tadikamalla and Popescu [87] worked on distributions having kurtosis greater than that of normal. For more information on distributional studies, see Montgomery [56], and Ho and Case [43]. For a general review on quality control readers can refer to Stoumbos et al. [85].

While constructing control limits, the choice of sample size and the number of samples are important issues and have been discussed by many authors. The construction of control limits depends on whether population parameters are known or unknown. In case of known parameters, events of being out of limits are consecutive Bernoulli trials. For $3\sigma$ control limits this probability equals to 0.0027. On the other hand, when the parameters are unknown, the limits are estimated by the sample statistics. In this case, the event of false alarm is no more independent of limits, hence the false alarm probability for $3\sigma$ limits is greater than 0.0027. Actually, probability of a false alarm depends only on number of samples ($M$) and sample size ($n$) (Quesenberry [70]). Although most of the researchers (see, Montgomery [57], Ryan [72]) claim that 20 to 30 samples of size 4 or 5 are enough to construct control limits, Quesenberry [70] proposed to use $M$ samples such that $M$ is equal to as $400/(n-1)$; $n$ being the sample size.

Some modifications of Shewhart control chart, known as exponentially weighted moving average charts (EWMA) and cumulative sum charts (CUSUM) were proposed to improve the ability of this chart in detecting small shifts in the process. CUSUM charts were first introduced by Page [66] as one of the modifications. Since then, many studies have been conducted on this topic; see for example Taylor [90], Ewan [35], Chiu [28], Hawkins and Olwell [42], Nishina and Peng-Hsiung [62]. For EWMA charts, see Lowry et al. [50], Lu et al. [51], and Lucas and Saccucci [52].

Several Bayesian approaches were also implemented in constructing quality control methods. Girshick and Rubin [37], Shiryaev [79], and Roberts [71] were the first to apply Bayesian procedures. There are also studies to construct control limits for multivariate distributions; see for example, Hotelling [44], Hawkins [41], Liu [49], and Mason et al. [54]. One can refer to Chakraborti et al. [22] for the literature on quality control procedures based on nonparametric methods.

### 1.1.2 RANKED SET SAMPLING

Ranked Set Sampling (RSS) is a newly developed method which was first proposed by McIntyre [55]. He developed this method for his research on the yield of pastures when he realized that inspection of the hay with an experienced eye is possible for the pastures from which the sampling and measurement is costly and time consuming. He proposed the procedure of RSS, and showed the gain in efficiency in estimating the population mean when compared to simple random sampling (SRS). McIntyre [55], also, demonstrated that the efficiency gain is high when the distribution of the population is symmetric, and the gain decreases as the distribution becomes skewed. Moreover, he mentioned some possible problems that may be faced with RSS; such as imperfect ranking and correlation among the sample units.

Halls and Dell [38] applied RSS in their study on forage yields in a pine hardwood forest. They are the ones to coin the name "ranked set sampling" for this method. Dell and Clutter [32] proved that the RSS estimator of the population mean is unbiased even in the case of errors in judgment ranking. They stated that the efficiency of the estimators obtained through RSS depends on the underlying distribution and the accuracy of the judgment ranking.

Takahasi and Wakimoto [89] were the first to obtain some theoretical results on RSS. Following them, Dell and Clutter [32] and David and Levine [31] proved the unbiasedness of the RSS estimator and its variance being smaller than that of the SRS estimator under both perfect and imperfect ranking. After that, important contributions were made by Stokes [83]-[82] who studied the efficiency of the variance estimator of RSS and set an estimator for correlation coefficient for bivariate normal distribution under RSS. Later, inferences about distribution function for RSS were made by Stokes and Sager [84], Kvam and Samaniego [48], and Chen [24].

Since 1990s there have been remarkable developments in the topic of RSS, which consist of its modifications, as well as some parametric and nonparametric applications. The modifications of RSS are extreme ranked set sampling (ERSS) (Shaibu and Muttlak, [76]-[77], Al-Naser and Mustafa [6], L ranked set sampling (Al-Nasser [3]), double ranked set sampling (DRSS) (Al-Saleh and Al-Kadiri [11], Al-Omari and Jaber [8], moving extreme ranked set sampling (MERSS) (Al-Saleh and Al-Hadrami [9]-[10], and ordered ranked set sampling (ORSS) (Balakrishnan and Li [15]-[16]-[17]-[18]). For the estimation of parameters in RSS, the readers can refer to Ozturk [63]-[64]-[65], Shadid et al. [75], and MacEachern et al. [53]. For the optimal design issues in RSS, see Chen and Bai [26], Nahhas et al. [60], and Wang et al. [102]. For some general information on this sampling method, the following references are useful: Chen et al. [27], Chen [25], Bai and Chen [14], Bouza [19], Sinha [80], Wolfe [103]-[104].

### 1.1.3 QUALITY CONTROL CHARTS BASED ON RANKED SET SAMPLING

In the literature, there are very few studies combining quality control and RSS. The first study suggesting the use of RSS method in quality control charts for the process mean is due to Salazar and Sinha [73]. They found that control charts based on RSS gave better results than its SRS counterparts. Later, Muttlak and Al-Sabah [59] compared quality control charts obtained under some modifications of RSS (ERSS, median ranked set sampling (MRSS)) with those based on SRS. They constructed all of the charts under the normality assumption, and use average run length (ARL) for the comparison when the process is both in control and in need of improvement for the process mean. Moreover, they compared the performances of the charts based on RSS, ERSS, and MRSS under perfect and imperfect ranking conditions. According to their research, RSS and two of its modifications were found to be more superior than the classical control chart based on SRS. Furthermore, MRSS was found to be the best in terms of detecting an out-of control process when compared to RSS and ERSS, and they suggested to use MRSS and ERSS rather than RSS for control charting procedure since they are more robust to imperfect ranking errors and more practical in real life cases.

Another research on quality control based on RSS and its modifications was conducted by Abujiya and Muttlak [1]. In their study, they developed quality control charts based on double ranked set sampling (DRSS), median double ranked set sampling (MDRSS), double median ranked set sampling (DMRSS), and compared their efficiencies with those based on SRS, RSS, and MRSS. They concluded that DRSS is better than SRS and RSS, and most efficient results are obtained by DMRSS method.

Al-Omari and Al-Naser [7], constructed control charts using robust extreme ranked set sampling (RERSS). In this study, they reached the conclusion that the efficiencies of RSS and RERSS, become more evident as the process gets out-of-control. Al-Nasser and Al-Rawwash [5] used robust L-ranked set sampling (LRSS) in order to construct quality control chart for the mean of a process. Folded ranked set sampling (FRSS) which reduces the waste of units in RSS was applied to control charting procedure and was found to be more efficient than SRS but less efficient than RSS (Al-Nasser et al. [4]).

# CHAPTER 2

# QUALITY CONTROL CHARTS UNDER NORMAL DISTRIBUTION

In this thesis, control charts for the process mean are studied under three different symmetrical distributions. Therefore, when "quality control chart" is mentioned in the text, it means quality control chart for the mean of a process or simply $\bar{X}$ chart. Quality control charts are constructed mostly under the assumption of normality; therefore, the first distribution which will be covered in this chapter is normal distribution.

As mentioned in Chapter 1, control chart procedure consists of two phases. In the first phase, control limits are constructed. If the population parameters are known, the limits are set by directly using these values. Since the interest is the population mean ($\mu$), control limits are set around $\mu$. This means the center line for the control chart is $\mu$, and the lower and upper limits of the control chart are set apart 3 standard deviations from the center line $\mu$, as shown in Equation 2.1. In the literature, the most frequently used limit is $3\sigma$ for Shewhart control charts; $1.5\sigma$ and $2\sigma$ limits also being of use. On the other hand, in the case of unknown population parameters, $M$ samples of size $n$ are selected from the process under control, and the parameters are estimated by using these samples.

$$\begin{aligned} UCL &= \mu + 3\frac{\sigma}{\sqrt{n}} \ , \\ CL &= \mu \ , \\ LCL &= \mu - 3\frac{\sigma}{\sqrt{n}} \ . \end{aligned} \tag{2.1}$$

In the second phase of the control charting procedure, the process is monitored. For the quality of a product, the sample mean is used to estimate the central tendency of a grouped data (Shewhart [78]). At some points of time, one sample of size $n$ is selected, and the mean of this sample is plotted on the chart. The state of the process, whether being under control or not, is judged according to the graph of the sample means. If any sample statistic is out-of the control limits or there is a pattern in the sample statistics lying within the control limits, this indicates that the process is out-of-control.

Most researchers have been interested in the distribution of the events of observations being out-of control limits. In the known parameter case, these events, which are also called false signal, are consecutive and independent Bernoulli trials and the probability is equal to 0.0027 for $3\sigma$ limits as stated in Quesenberry [70]. However, when the parameters are unknown, observations and estimated control limits are dependent to each other, and the event of false signal is distributed as a linear function of normal distribution (see Quesenberry [70]). The probability of false alarm, in this case, is greater than 0.0027. As the number of samples $M$ increases, the false alarm probability approaches to 0.0027.

## 2.1 Quality Control Charts with Simple Random Sampling

In the literature, the most frequently used sampling method for control charts is simple random sampling (SRS). To obtain control limits, MLE of $\mu$ (which is $\bar{X}$) is utilized as $\bar{X}$ is an unbiased and efficient estimator of $\mu$.

$$E(\bar{X}) = \mu \ , \tag{2.2}$$

$$Var(\bar{X}) = \sigma^2/n \ . \tag{2.3}$$

In the estimation part of the control chart, $M$ random samples of size $n$ are selected by using SRS, which leads to the use of $\bar{\bar{X}}$ to estimate $\mu$. On the other hand, the variance of $\mu$ is estimated by $\frac{\bar{S}}{c_4\sqrt{n}}$ (see Equation 2.5) where $\bar{S}$ is the mean of $M$ sample standard deviations. The constant $c_4$ is used to make $S$ an unbiased estimator of $\sigma$. The constant $c_4$ can easily be obtained and is given approximately by

$$c_4 = \sqrt{\frac{2}{n-1}}\left(\frac{\Gamma(n/2)}{\Gamma(n-1/2)}\right) \ ; \tag{2.4}$$

see Duncan [33] for details.

Thus, $3\sigma$ control limits for normal distribution under SRS can be written as follows:

$$\begin{aligned} UCL &= \bar{\bar{X}} + 3\frac{\bar{S}}{c_4\sqrt{n}} \ , \\ CL &= \bar{\bar{X}} \ , \\ LCL &= \bar{\bar{X}} - 3\frac{\bar{S}}{c_4\sqrt{n}} \ . \end{aligned} \tag{2.5}$$

## 2.2 Quality Control Charts with Ranked Set Sampling

Shewhart [78] emphasized the importance of the sampling method in a quality control chart procedure. In the literature, due to the ease of application, mostly SRS is used when constructing control charts. However, small samples are preferred in the monitoring stage of Shewhart control charts (as mentioned in Chapter 1), and RSS has been shown to be superior over SRS in that case. Due to this fact, it is worth seeing whether RSS also performs better under Shewhart control chart process.

In this sampling method, one firstly identifies a large sample from the population of interest, and then chooses the units to be quantified according to a certain process. That is to say, if it is desired to select a random sample of size $n$ from a population, initially a random sample of size $n^2$ is selected in order for identification. Then, $n^2$ sample units are randomly allocated into $n$ samples of size $n$. Finally, each of the $n$ samples is ranked within itself according to visual inspection without actually measuring the units, and the first ranked one is selected from the first sample for measurement on the variable of interest. Following the first measurement, the second ranked unit from the second set is picked and measured. This continues until the $n^{th}$ ranked unit in the $n^{th}$ set is selected for actual measurement. If it is desired to obtain a larger sample, say a sample of size $N = n * r$, the process explained above is repeated $r$ times. When considering the sample selection scheme of RSS, there is a similarity between this method and the stratified sampling in the sense that both processes divide the population into many subpopulations. In the RSS case, each of these subpopulations has a different distribution (Patil et al. [67]-[68], Chen et al. [27]). The RSS scheme is illustrated in Figure 2.1 for $n = 4$ and $r$ cycles . The observations in bold are those which are chosen for actual measurement.

| | **Judgment Rank** | | | |
|---|---|---|---|---|
| **Cycle** | **1** | **2** | **3** | **4** |
| **1** | $\mathbf{X_{(1)1}}$ | $X_{(2)1}$ | $X_{(3)1}$ | $X_{(4)1}$ |
| | $X_{(1)2}$ | $\mathbf{X_{(2)2}}$ | $X_{(3)2}$ | $X_{(4)2}$ |
| | $X_{(1)3}$ | $X_{(2)3}$ | $\mathbf{X_{(3)3}}$ | $X_{(4)3}$ |
| | $X_{(1)4}$ | $X_{(2)4}$ | $X_{(3)4}$ | $\mathbf{X_{(4)4}}$ |
| **...** | ... | ... | ... | ... |
| | ... | ... | ... | ... |
| | ... | ... | ... | ... |
| | ... | ... | ... | ... |
| **r** | $\mathbf{X_{(1)1}}$ | $X_{(2)1}$ | $X_{(3)1}$ | $X_{(4)1}$ |
| | $X_{(1)2}$ | $\mathbf{X_{(2)2}}$ | $X_{(3)2}$ | $X_{(4)2}$ |
| | $X_{(1)3}$ | $X_{(2)3}$ | $\mathbf{X_{(3)3}}$ | $X_{(4)3}$ |
| | $X_{(1)4}$ | $X_{(2)4}$ | $X_{(3)4}$ | $\mathbf{X_{(4)4}}$ |

Figure 2.1: RSS scheme for $n = 4$ and $r$ cycles

In the literature, there are a few applications of RSS in control charts, see Salazar and

Sinha [73], Muttlak and Al-Sabah [59], Abujiya and Muttlak [1], Al-Omari and Al-Naser [7]. However, no closed form solutions of the parameter estimators are available in those studies. This is due to the fact that the likelihood function consists of some intractable functions.

Let the distribution of the random variable of interest $X$ be denoted by $(1/\sigma)\, f\left(\frac{x-\mu}{\sigma}\right)$,

$$\frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}; \quad -\infty < x < \infty,\ -\infty < \mu < \infty,\ \sigma > 0\,. \tag{2.6}$$

Since each order statistic is taken from a different sample, the order statistics are independent of each other. Therefore, the likelihood function can be written as the product of the density function of each order statistic

$$L = \prod_{i=1}^{n} f_{i:n}(x_{(i)})\,, \tag{2.7}$$

where

$$f_{i:n}(x_{(i)}) = \frac{n!}{(n-i)!(i-1)!} \frac{1}{\sigma} f\left(\frac{x_{(i)}-\mu}{\sigma}\right) \left[F\left(\frac{x_{(i)}-\mu}{\sigma}\right)\right]^{i-1} \left[1 - F\left(\frac{x_{(i)}-\mu}{\sigma}\right)\right]^{n-i}; \tag{2.8}$$

$$1 \le i \le n,\ -\infty < x_{(i)} < \infty,\ -\infty < \mu < \infty,\ \sigma > 0\,.$$

Further, the likelihood can be written as

$$L = \prod_{i=1}^{n} \frac{n!}{(n-i)!(i-1)!} \frac{1}{\sigma^n} \prod_{i=1}^{n} f(z_{(i)}) \prod_{i=1}^{n} \left[F(z_{(i)})\right]^{i-1} \prod_{i=1}^{n} \left[1 - F(z_{(i)})\right]^{n-i}\,, \tag{2.9}$$

where $F(z) = \int_{-\infty}^{z} f(t)\,\mathrm{d}t$ is the cdf of the normal distribution; $z_{(i)} = (x_{(i)} - \mu)/\sigma$.

To estimate $\mu$ and $\sigma$, the following partial derivatives of likelihood equations are obtained:

$$\frac{\partial lnL}{\partial \mu} = \left(\frac{1}{\sigma}\right) \left[\sum_{i=1}^{n} z_{(i)} - \sum_{i=1}^{n} (i-1) \frac{f(z_{(i)})}{F(z_{(i)})} + \sum_{i=1}^{n} (n-i) \frac{f(z_{(i)})}{1 - F(z_{(i)})}\right], \tag{2.10}$$

$$\frac{\partial lnL}{\partial \sigma} = \left(\frac{1}{\sigma}\right) \left[\sum_{i=1}^{n} z_{(i)}^2 - \sum_{i=1}^{n} (i-1) z_{(i)} \frac{f(z_{(i)})}{F(z_{(i)})} + \sum_{i=1}^{n} (n-i) z_{(i)} \frac{f(z_{(i)})}{1 - F(z_{(i)})} - n\right]. \tag{2.11}$$

The MLEs of these parameters can not be obtained explicitly from Equations 2.10 and 2.11 due to the nonlinear functions

$$g_1(z_{(i)}) = \frac{f(z_{(i)})}{F(z_{(i)})} , \qquad g_2(z_{(i)}) = \frac{f(z_{(i)})}{1-F(z_{(i)})} ; \quad 1 \le i \le n . \tag{2.12}$$

### 2.2.1 Modified Maximum Likelihood Estimation

Tiku [92], and Tiku and Suresh [97] proposed modified maximum likelihood to overcome the difficulties mentioned above. There is only one assumption for this model: Te first two derivatives of ln$f$ must exist ($f$ is the pdf of the random variable of interest). In this method, nonlinear terms in the likelihood functions are linearized by Taylor series expansion. Consider the nonlinear terms in the likelihood functions 2.10 and 2.11,

$$g_1(z_{(i)}) \cong \alpha_{1i} - \beta_{1i} z_{(i)} , \qquad g_2(z_{(i)}) \cong \alpha_{2i} + \beta_{2i} z_{(i)} ; \quad 1 \le i \le n . \tag{2.13}$$

The linearized forms of the nonlinear terms (2.13) are obtained by using the first two terms of Taylor series expansions of the functions $g_1(z_{(i)})$ and $g_2(z_{(i)})$ around the expected values of $z_{(i)}$. The expected values of the standardized normal order statistics are given by Harter [40]. The coefficients $\alpha_{1i}$, $\beta_{1i}$, $\alpha_{2i}$, and $\beta_{2i}$ are given in 2.14 - 2.15,

$$\beta_{1i} = \frac{t_{(i)} f(t_{(i)})}{F(t_{(i)})} + \left(\frac{f(t_{(i)})}{F(t_{(i)})}\right)^2 , \qquad \alpha_{1i} = \frac{f(t_{(i)})}{F(t_{(i)})} + t_{(i)} \beta_{1i} , \tag{2.14}$$

$$\beta_{2i} = \left(\frac{f(t_{(i)})}{1-F(t_{(i)})}\right)^2 - \frac{t_{(i)} f(t_{(i)})}{1-F(t_{(i)})} , \qquad \alpha_{2i} = \frac{f(t_{(i)})}{1-F(t_{(i)})} - t_{(i)} \beta_{2i} , \tag{2.15}$$

where $E(Z_{(i)}) = t_{(i)}$.

Substituting the linearized forms in the partial derivatives of likelihood function (2.10 - 2.11), the estimators of parameters $\mu$ and $\sigma$ can be obtained as

$$\hat{\mu} = \sum_{i=1}^{n} a_i x_{(i)} ; \qquad a_i = \frac{u_i}{\sum\limits_{i=1}^{n} u_i} \tag{2.16}$$

and

$$\hat{\sigma} = \frac{-B + \sqrt{B^2 + 4nC}}{2n} \tag{2.17}$$

where

$$B = \sum_{i=1}^{n} w_i x_{(i)} , \qquad C = \sum_{i=1}^{n} u_i \left(x_{(i)} - \hat{\mu}\right)^2 , \tag{2.18}$$

$$u_i = (i-1)\beta_{1i} + (n-i)\beta_{2i} + 1 \,, \qquad w_i = (i-1)\alpha_{1i} - (n-i)\alpha_{2i} \,. \tag{2.19}$$

The modified likelihood equation is asymptotically equivalent to the likelihood equation; therefore, MML estimators are asymptotically fully efficient and unbiased (Tiku and Akkaya [95]).

**Quality control limits for Shewhart control chart under RSS:** After obtaining the closed forms of the estimators of unknown parameters, the control limits are set by using these estimators. The mean of the process $\mu$ is estimated by $\bar{\hat{\mu}}$, where $\hat{\mu}$, is an unbiased estimator. To construct control limits, variance of $\hat{\mu}$ is needed, and it can be obtained as

$$Var(\hat{\mu}) = \frac{\sigma^2 \sum_{i=1}^{n} \left(u_i^2 Var(Z_{(i)})\right)}{\left(\sum_{i=1}^{n} u_i\right)^2} \,. \tag{2.20}$$

Then the control limits are

$$\begin{aligned} UCL &= \bar{\hat{\mu}} + 3\frac{c_1 \bar{\hat{\sigma}} \sqrt{\sum_{i=1}^{n} \left(u_i^2 Var(Z_{(i)})\right)}}{\sum_{i=1}^{n} u_i} \,, \\ CL &= \bar{\hat{\mu}} \,, \\ LCL &= \bar{\hat{\mu}} - 3\frac{c_1 \bar{\hat{\sigma}} \sqrt{\sum_{i=1}^{n} \left(u_i^2 Var(Z_{(i)})\right)}}{\sum_{i=1}^{n} u_i} \,, \end{aligned} \tag{2.21}$$

where $c_1$ is a constant used to make $\hat{\sigma}$ an unbiased estimator of $\sigma$, and is obtained by simulation.

### 2.2.2 Simulation Results

To see the superiority of RSS in the estimation of parameters, Monte Carlo simulation with 100,000 repetitions is conducted under normal distribution with parameters $\mu = 0$ and $\sigma = 1$. Table 2.1 shows the bias of the estimators and the mean squared errors (MSE) under the corresponding sample size and sampling method.

According to the simulation results in Table 2.1, the bias of $\hat{\mu}$ obtained by both SRS and RSS are negligible. RSS yields smaller MSE for all sample sizes (3,...,10). When

Table 2.1: Comparison of RSS and SRS in the estimation of Normal distribution parameters

| | $\hat{\mu}$ | | | | $\hat{\sigma}$ | | | |
|---|---|---|---|---|---|---|---|---|
| | **Bias** | | **MSE** | | **Bias** | | **MSE** | |
| **n** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** |
| 3 | 0.0007 | 0.0009 | 0.3343 | 0.1733 | -0.1132 | -0.1879 | 0.2272 | 0.1623 |
| 4 | -0.0010 | 0.0007 | 0.2487 | 0.1048 | -0.0794 | -0.1127 | 0.1575 | 0.1026 |
| 5 | 0.0001 | 0.0016 | 0.1998 | 0.0699 | -0.0604 | -0.0658 | 0.1205 | 0.0711 |
| 6 | 0.0013 | 0.0001 | 0.1670 | 0.0498 | -0.0485 | -0.0417 | 0.0964 | 0.0538 |
| 7 | -0.0001 | -0.0002 | 0.1423 | 0.0377 | -0.0414 | -0.0251 | 0.0814 | 0.0426 |
| 8 | 0.0018 | 0.0003 | 0.1250 | 0.0292 | -0.0343 | -0.0133 | 0.0699 | 0.0348 |
| 9 | -0.0005 | 0.0003 | 0.1119 | 0.0231 | -0.0311 | -0.0056 | 0.0613 | 0.0287 |
| 10 | 0.0005 | -0.0003 | 0.0994 | 0.0190 | -0.0263 | 0.0005 | 0.0546 | 0.0242 |

the scale parameter is of concern, RSS estimators have smaller bias and MSEs for most of the sample sizes. To more clearly demonstrate the efficiency of RSS over SRS, relative efficiency (RE) table is constructed by dividing the MSE of the estimators obtained under SRS by those under RSS. Obtaining RE greater than 1 means RSS is more efficient, and vice versa.

Table 2.2: Relative efficiency of $\hat{\mu}$ and $\hat{\sigma}$ under SRS and RSS

| n | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|
| $\hat{\mu}$ | 1.93 | 2.37 | 2.86 | 3.35 | 3.78 | 4.29 | 4.84 | 5.22 |
| $\hat{\sigma}$ | 1.40 | 1.53 | 1.69 | 1.79 | 1.91 | 2.01 | 2.13 | 2.25 |

According to Table 2.2, it can be said that relative efficiency of RSS is always greater than 1 meaning that RSS is always more efficient than SRS. Moreover, as the sample size increases, RE also increases.

Another Monte Carlo simulation is conducted to compare the performances of Shewhart control charts under SRS and RSS with 100,000 repetitions. The control chart simulations are conducted under Normal distribution with parameters $\mu = 0$, $\sigma = 1$, and $3\sigma$ control limits are set such that the type I error level is 0.0027. Four different sets of samples ($M = 20, 30, 50, 100$) are used in the simulations for estimating the control limits. The results are given in Table 2.3.

The results in Table 2.3 show that with fewer number of samples RSS yields smaller type I errors for smaller number of samples. On the other hand, as the number of samples increases, both sampling methods perform almost the same, and the type I error levels get closer to the target value 0.0027

Table 2.3: Type I error comparison for Shewhart control charts constructed under SRS and RSS

| M | 20 | | 30 | | 50 | | 100 | |
|---|---|---|---|---|---|---|---|---|
| **n** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** |
| 3 | 0.0055 | 0.0050 | 0.0045 | 0.0042 | 0.0037 | 0.0037 | 0.0032 | 0.0033 |
| 4 | 0.0047 | 0.0044 | 0.0040 | 0.0038 | 0.0034 | 0.0034 | 0.0031 | 0.0031 |
| 5 | 0.0044 | 0.0041 | 0.0038 | 0.0036 | 0.0033 | 0.0033 | 0.0030 | 0.0030 |
| 6 | 0.0042 | 0.0039 | 0.0037 | 0.0035 | 0.0032 | 0.0032 | 0.0030 | 0.0030 |
| 7 | 0.0041 | 0.0038 | 0.0035 | 0.0034 | 0.0032 | 0.0031 | 0.0029 | 0.0030 |
| 8 | 0.0040 | 0.0037 | 0.0035 | 0.0034 | 0.0032 | 0.0031 | 0.0029 | 0.0029 |
| 9 | 0.0039 | 0.0037 | 0.0035 | 0.0034 | 0.0031 | 0.0031 | 0.0029 | 0.0029 |
| 10 | 0.0039 | 0.0036 | 0.0034 | 0.0033 | 0.0031 | 0.0031 | 0.0029 | 0.0029 |

## 2.3 Quality Control Charts with Extreme Ranked Set Sampling

The efficiency of the sampling method RSS depends thoroughly on the accuracy of judgment ranks. In the worst case (completely wrong judgment ranks), the RSS is at least as efficient as SRS; however, RSS loses its superiority over SRS. To overcome the judgment ranking errors, a more robust version called Extreme Ranked Set Sampling (ERSS) has been developed.

ERSS scheme is very similar to that of RSS except that only the first and last ordered units are selected for measurement. As in the first and second stages of RSS, $n^2$ units are selected for identification and are randomly allocated into $n$ sets of size $n$. Next, each set is ranked within itself, and if $n$ is even, the first ranked units in the first $n/2$ sets and the last ranked units in the last $n/2$ sets are selected for actual measurement. However, if the sample size $n$ is odd, the first ranked units in the first $(n-1)/2$ sets, the last ranked units in the following $(n-1)/2$ sets, and the median of the last set are selected for quantification. Since identifying the extreme values in a set is an easier process than identifying the $i^{th}, (i = 2, 3, ..., n-1)$ units, ERSS is more robust to imperfect ranking errors. In Figures 2.2 and 2.3, ERSS scheme for even and odd sized samples are illustrated. The observations written in bold are selected for actual measurement.

For the estimation process in ERSS, let the distribution of the random variable $X$ be denoted by

$$\frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}; \quad -\infty < x < \infty, \ -\infty < \mu < \infty, \ \sigma > 0\,. \tag{2.22}$$

The likelihood function can easily be obtained through the multiplication of the pdfs of the first and last order statistics (or median, if odd sized sample) since each order

| | Judgment Rank | | | |
|---|---|---|---|---|
| **Cycle** | **1** | **2** | **3** | **4** |
| **1** | $\boldsymbol{X_{(1)1}}$ | $X_{(2)1}$ | $X_{(3)1}$ | $X_{(4)1}$ |
| | $\boldsymbol{X_{(1)2}}$ | $X_{(2)2}$ | $X_{(3)2}$ | $X_{(4)2}$ |
| | $X_{(1)3}$ | $X_{(2)3}$ | $X_{(3)3}$ | $\boldsymbol{X_{(4)3}}$ |
| | $X_{(1)4}$ | $X_{(2)4}$ | $X_{(3)4}$ | $\boldsymbol{X_{(4)4}}$ |
| ... | ... | ... | ... | ... |
| | ... | ... | ... | ... |
| | ... | ... | ... | ... |
| | ... | ... | ... | ... |
| **r** | $\boldsymbol{X_{(1)1}}$ | $X_{(2)1}$ | $X_{(3)1}$ | $X_{(4)1}$ |
| | $\boldsymbol{X_{(1)2}}$ | $X_{(2)2}$ | $X_{(3)2}$ | $X_{(4)2}$ |
| | $X_{(1)3}$ | $X_{(2)3}$ | $X_{(3)3}$ | $\boldsymbol{X_{(4)3}}$ |
| | $X_{(1)4}$ | $X_{(2)4}$ | $X_{(3)4}$ | $\boldsymbol{X_{(4)4}}$ |

Figure 2.2: ERSS scheme for $n = 4$ and $r$ cycles

statistic is obtained from a different sample and hence independent of each other. For an even sized sample, the likelihood function can be written as

$$L = \prod_{i=1}^{n/2} f_{1:n}(x_{(1)i}) \prod_{i=n/2+1}^{n} f_{n:n}(x_{(n)i}) \,, \tag{2.23}$$

where

$$f_{1:n}(x_{(1)}) = n\frac{1}{\sigma} f\left(\frac{x_{(1)} - \mu}{\sigma}\right)\left[1 - F\left(\frac{x_{(1)} - \mu}{\sigma}\right)\right]^{n-1} \quad ; \quad -\infty < x_{(1)} < \infty \,, \tag{2.24}$$

$$f_{n:n}(x_{(n)}) = n\frac{1}{\sigma} f\left(\frac{x_{(n)} - \mu}{\sigma}\right)\left[F\left(\frac{x_{(n)} - \mu}{\sigma}\right)\right]^{n-1} \quad ; \quad -\infty < x_{(n)} < \infty \,. \tag{2.25}$$

The likelihood function can also be written as

$$L = \left(\frac{n}{\sigma}\right)^n \prod_{i=1}^{n/2} \left[f(z_{(1)i})\left(1 - F(z_{(1)i})\right)^{n-1}\right] \prod_{i=n/2+1}^{n} \left[f(z_{(n)i}) F(z_{(n)i})^{n-1}\right] \,, \tag{2.26}$$

where $F(z) = \int_{-\infty}^{z} f(t)\,\mathrm{d}t$ is the cdf of the normal distribution, and $z_{(i)} = (x_{(i)} - \mu)/\sigma$.

Likewise, the likelihood function for an odd sized sample is

$$L = f_{\frac{n+1}{2}:n}(x_{(\frac{n+1}{2})}) \prod_{i=1}^{n/2} f_{1:n}(x_{(1)i}) \prod_{i=n/2+1}^{n} f_{n:n}(x_{(n)i}) \,, \tag{2.27}$$

| | Judgment Rank | | | | |
|---|---|---|---|---|---|
| **Cycle** | **1** | **2** | **3** | **4** | **5** |
| **1** | $\boldsymbol{X_{(1)1}}$ | $X_{(2)1}$ | $X_{(3)1}$ | $X_{(4)1}$ | $X_{(5)1}$ |
| | $\boldsymbol{X_{(1)2}}$ | $X_{(2)2}$ | $X_{(3)2}$ | $X_{(4)2}$ | $X_{(5)2}$ |
| | $X_{(1)3}$ | $X_{(2)3}$ | $X_{(3)3}$ | $X_{(4)3}$ | $\boldsymbol{X_{(5)3}}$ |
| | $X_{(1)4}$ | $X_{(2)4}$ | $X_{(3)4}$ | $X_{(4)4}$ | $\boldsymbol{X_{(5)4}}$ |
| | $X_{(1)5}$ | $X_{(2)5}$ | $\boldsymbol{X_{(3)5}}$ | $X_{(4)5}$ | $X_{(5)5}$ |
| | ... | ... | ... | ... | ... |
| | ... | ... | ... | ... | ... |
| **...** | ... | ... | ... | ... | ... |
| | ... | ... | ... | ... | ... |
| | ... | ... | ... | ... | ... |
| | ... | ... | ... | ... | ... |
| **r** | $\boldsymbol{X_{(1)1}}$ | $X_{(2)1}$ | $X_{(3)1}$ | $X_{(4)1}$ | $X_{(5)1}$ |
| | $\boldsymbol{X_{(1)2}}$ | $X_{(2)2}$ | $X_{(3)2}$ | $X_{(4)2}$ | $X_{(5)2}$ |
| | $X_{(1)3}$ | $X_{(2)3}$ | $X_{(3)3}$ | $X_{(4)3}$ | $\boldsymbol{X_{(5)3}}$ |
| | $X_{(1)4}$ | $X_{(2)4}$ | $X_{(3)4}$ | $X_{(4)4}$ | $\boldsymbol{X_{(5)4}}$ |
| | $X_{(1)5}$ | $X_{(2)5}$ | $\boldsymbol{X_{(3)5}}$ | $X_{(4)5}$ | $X_{(5)5}$ |

Figure 2.3: ERSS scheme for $n = 5$ and $r$ cycles

where

$$f_{\frac{n+1}{2}:n}(x_{(\frac{n+1}{2})}) = \frac{n!}{\left(\frac{n-1}{2}\right)!\left(\frac{n-1}{2}\right)!}\left[F(x_{(\frac{n+1}{2})})\right]^{(n-1)/2} f(x_{(\frac{n+1}{2})}) \left[1 - F(x_{(\frac{n+1}{2})})\right]^{(n-1)/2} \tag{2.28}$$

is the pdf of the median. More specifically, the likelihood for odd sized samples can also be written as

$$L = C\left(\frac{1}{\sigma}\right)^{n}\prod_{i=1}^{(n-1)/2}\left[f(z_{(1)i})\left(1-F(z_{(1)i})\right)^{n-1}\right]\prod_{i=(n+1)/2}^{n-1}\left[f(z_{(n)i})F(z_{(n)i})^{n-1}\right] \left[F(z_{(\frac{n+1}{2})})^{(n-1)/2} f(z_{(\frac{n+1}{2})})\left(1-F(z_{(\frac{n+1}{2})})\right)^{(n-1)/2}\right], \tag{2.29}$$

where $C = \frac{n!n^{n-1}}{(\frac{n-1}{2})!(\frac{n-1}{2})!}$.

Then, one can take the partial derivatives of the likelihood function (even sized samples) as follows:

$$
\begin{aligned}
\frac{\partial lnL}{\partial \mu} &= \left(\frac{1}{\sigma}\right)\Bigg\{\sum_{i=1}^{n/2} z_{(1)i} + \sum_{i=n/2+1}^{n} z_{(n)i} \\
&\quad +(n-1)\left(\sum_{i=1}^{n/2} g_1(z_{(1)i}) - \sum_{i=n/2+1}^{n} g_2(z_{(n)i})\right)\Bigg\} , \qquad (2.30) \\
\frac{\partial lnL}{\partial \sigma} &= \left(\frac{1}{\sigma}\right)\Bigg\{\sum_{i=1}^{n/2} z_{(1)i}^2 + \sum_{i=n/2+1}^{n} z_{(n)i}^2 - n \\
&\quad +(n-1)\left(\sum_{i=1}^{n/2} g_1(z_{(1)i}) - \sum_{i=n/2+1}^{n} g_2(z_{(n)i})\right)\Bigg\} . \qquad (2.31)
\end{aligned}
$$

For odd sized samples, partial derivative equations are obtained as

$$
\begin{aligned}
\frac{\partial lnL}{\partial \mu} &= \left(\frac{1}{\sigma}\right)\Bigg\{\sum_{i=1}^{n/2} z_{(1)i} + \sum_{i=n/2+1}^{n} z_{(n)i} + z_{(\frac{n+1}{2})} \\
&\quad + (n-1)\left[\sum_{i=1}^{n/2} g_1(z_{(1)i}) - \sum_{i=n/2+1}^{n} g_2(z_{(n)i})\right] \\
&\quad + \left(\frac{n-1}{2}\right)\left[g_1(z_{(\frac{n+1}{2})}) - g_2(z_{(\frac{n+1}{2})})\right]\Bigg\} , \qquad (2.32) \\
\frac{\partial lnL}{\partial \sigma} &= \left(\frac{1}{\sigma}\right)\Bigg\{\sum_{i=1}^{n/2} z_{(1)i}^2 + \sum_{i=n/2+1}^{n} z_{(n)i}^2 + z_{(\frac{n+1}{2})}^2 - n \\
&\quad + (n-1)\left[\sum_{i=1}^{n/2} g_1(z_{(1)i}) - \sum_{i=n/2+1}^{n} g_2(z_{(n)i})\right] \\
&\quad + \left(\frac{n-1}{2}\right) z_{(\frac{n+1}{2})}\left[g_1(z_{(\frac{n+1}{2})}) - g_2(z_{(\frac{n+1}{2})})\right]\Bigg\} . \qquad (2.33)
\end{aligned}
$$

As in the estimation process of RSS, nonlinear terms (see 2.34) exist in the partial derivatives of the likelihood function; therefore, it is not possible to directly obtain closed form solutions of MLEs for parameters $\mu$ and $\sigma$.

$$
g_1(z_{(i)}) = \frac{f(z_{(i)})}{1-F(z_{(i)})} , \qquad g_2(z_{(i)}) = \frac{f(z_{(i)})}{F(z_{(i)})} ; \qquad i = 1, \frac{n+1}{2}, n . \qquad (2.34)
$$

### 2.3.1 Modified Maximum Likelihood Estimation

As proposed by Tiku [92], and Tiku and Suresh [97], the nonlinear terms (2.34) can be linearized by Taylor series expansion around the expected values of standardized order statistics ($E(z_{(i)}) = t_{(i)}$). They are given by

$$g_1(z_{(i)}) \cong \alpha_{1i} + \beta_{1i} z_{(i)} \ , \qquad g_2(z_{(i)}) \cong \alpha_{2i} - \beta_{2i} z_{(i)} \ ; \qquad i = 1, \frac{n+1}{2}, n \ . \tag{2.35}$$

where

$$\beta_{1i} = \left( \frac{f(t_{(i)})}{1 - F(t_{(i)})} \right)^2 - \frac{t_{(i)} f(t_{(i)})}{1 - F(t_{(i)})} \ , \qquad \alpha_{1i} = \frac{f(t_{(i)})}{1 - F(t_{(i)})} - t_{(i)} \beta_{1i} \ , \tag{2.36}$$

$$\beta_{2i} = \left( \frac{f(t_{(i)})}{F(t_{(i)})} \right)^2 + \frac{t_{(i)} f(t_{(i)})}{F(t_{(i)})} \quad \text{and} \qquad \alpha_{2i} = \frac{f(t_{(i)})}{F(t_{(i)})} - t_{(i)} \beta_{2i} \ . \tag{2.37}$$

Substituting the linearized forms of the terms (2.35s) in the partial derivatives of likelihood function, the estimators of unknown parameters $\mu$ and $\sigma$ can be obtained as

$$\hat{\mu} = \frac{\sum\limits_{i=1}^{n/2} x_{(1)i} + \sum\limits_{i=n/2+1}^{n} x_{(n)i}}{n} \tag{2.38}$$

and

$$\hat{\sigma} = \frac{-B + \sqrt{B^2 + 4nC}}{2n} \tag{2.39}$$

where

$$B = (n-1)\alpha_{11} \left( \sum_{i=n/2+1}^{n} x_{(n)i} - \sum_{i=1}^{n/2} x_{(1)i} \right) , \tag{2.40}$$

$$C = (1 + (n-1)\beta_{11}) \left[ \sum_{i=1}^{n/2} (x_{(1)i} - \hat{\mu})^2 + \sum_{i=n/2+1}^{n} (x_{(n)i} - \hat{\mu})^2 \right] \tag{2.41}$$

and $n$ is even.

For odd sized samples, the parameter estimators are

$$\hat{\mu} = \frac{(1 + (n-1)\beta_{11}) \left( \sum\limits_{i=1}^{(n-1)/2} x_{(1)i} + \sum\limits_{i=(n+1)/2}^{n-1} x_{(n)i} \right) + \left( 1 + (n-1)\beta_{1\frac{n+1}{2}} \right) x_{(\frac{n+1}{2})}}{(n-1)\left(1 + (n-1)\beta_{11}\right) + \left( 1 + (n-1)\beta_{1\frac{n+1}{2}} \right)} \tag{2.42}$$

and

$$\hat{\sigma} = \frac{-B + \sqrt{B^2 + 4nC}}{2n} \tag{2.43}$$

where

$$B = (n-1)\alpha_{11}\left(\sum_{i=n/2+1}^{n} x_{(n)i} - \sum_{i=1}^{n/2} x_{(1)i}\right), \tag{2.44}$$

$$\begin{aligned} C = (1+(n-1)\beta_{11})\left[\sum_{i=1}^{n/2}(x_{(1)i}-\hat{\mu})^2 + \sum_{i=n/2+1}^{n}(x_{(n)i}-\hat{\mu})^2\right] \\ + \left(1+(n-1)\beta_{1\frac{n+1}{2}}\right)\left(x_{(\frac{n+1}{2})}-\hat{\mu}\right)^2 . \end{aligned} \tag{2.45}$$

As mentioned in Section 2.2.1, the MML estimators are asymptotically unbiased and fully efficient.

**Quality control limits for Shewhart control chart under ERSS:** In order to construct control limits, MML estimators under ERSS can be used. The mean of the process is estimated by $\bar{\hat{\mu}}$ which is an unbiased estimator of $\mu$ for both even and odd sized samples and the variance of the process mean is estimated by

$$Var(\hat{\mu}_{mml,even}) = \frac{1}{n}\sigma^2 Var(Z_{(1)}) \tag{2.46}$$

$$\begin{aligned} Var(\hat{\mu}_{mml,odd}) = \sigma^2 \Bigg\{ & \frac{(n-1)\left(1+(n-1)\beta_{11}\right)^2 Var(Z_{(1)})}{\left[(n-1)\left(1+(n-1)\beta_{11}\right)+\left(1+(n-1)\beta_{1\frac{n+1}{2}}\right)\right]^2} \\ & + \frac{\left(1+(n-1)\beta_{1\frac{n+1}{2}}\right)^2 Var\left(Z_{(\frac{n+1}{2})}\right)}{\left[(n-1)\left(1+(n-1)\beta_{11}\right)+\left(1+(n-1)\beta_{1\frac{n+1}{2}}\right)\right]^2}\Bigg\} . \end{aligned} \tag{2.47}$$

Then, the $3\sigma$ control limits are constructed as

$$\begin{aligned} UCL &= \bar{\hat{\mu}} + 3\sqrt{Var(\hat{\mu})} , \\ CL &= \bar{\hat{\mu}} , \\ LCL &= \bar{\hat{\mu}} - 3\sqrt{Var(\hat{\mu})} . \end{aligned} \tag{2.48}$$

The unknown parameter $\sigma$ in the variance of $\hat{\mu}$ is estimated by $c_1\bar{\hat{\sigma}}$ where $c_1$ is a constant to make the $\hat{\sigma}$ an unbiased estimator of $\sigma$. The constant $c_1$ can easily be obtained by simulation for both even and odd sized samples.

### 2.3.2 Best Linear Unbiased Estimation Method

Parameter estimators under ERSS scheme can also be obtained by best linear unbiased estimation (BLUE) method (Tiku and Akkaya [95]) as

$$\hat{\mu}_{blue,even} = \frac{\sum\limits_{i=1}^{n/2} x_{(1)i} + \sum\limits_{i=n/2+1}^{n} x_{(n)i}}{n}, \tag{2.49}$$

$$\hat{\sigma}_{blue,even} = \frac{\sum\limits_{i=n/2+1}^{n} x_{(n)i} - \sum\limits_{i=1}^{n/2} x_{(1)i}}{nt_{(n)}} \tag{2.50}$$

and

$$\hat{\mu}_{blue,odd} = \frac{\left(\sum\limits_{i=1}^{(n-1)/2} x_{(1)i} + \sum\limits_{i=n/2+1}^{n} x_{(n)i}\right) Var\left(Z_{(\frac{n+1}{2})}\right) + x_{(\frac{n+1}{2})} Var(Z_{(1)})}{Var(Z_{(1)}) + (n-1)Var\left(Z_{(\frac{n+1}{2})}\right)}, \tag{2.51}$$

$$\hat{\sigma}_{blue,odd} = \frac{\sum\limits_{i=(n+1)/2}^{n} x_{(n)i} - \sum\limits_{i=1}^{(n-1)/2} x_{(1)i}}{(n-1)t_{(n)}}. \tag{2.52}$$

Since $\hat{\mu}_{mml,even}$ and $\hat{\mu}_{blue,even}$ are in the same forms, $Var(\hat{\mu}_{blue,even})$ equals to $Var(\hat{\mu}_{mml,even})$ as in Equation 2.46. On the other hand, one can obtain the variance of $\hat{\mu}$ for odd sized samples as

$$Var(\hat{\mu}_{blue,odd}) = \sigma^2 \frac{Var(Z_{(1)})Var\left(Z_{(\frac{n+1}{2})}\right)}{Var(Z_{(1)}) + (n-1)Var\left(Z_{(\frac{n+1}{2})}\right)}. \tag{2.53}$$

The $3\sigma$ control limits by using BLU estimators can then be written as

$$\begin{aligned} UCL &= \bar{\hat{\mu}} + 3\sqrt{Var(\hat{\mu})}, \\ CL &= \bar{\hat{\mu}}, \\ LCL &= \bar{\hat{\mu}} - 3\sqrt{Var(\hat{\mu})}. \end{aligned} \tag{2.54}$$

### 2.3.3 Simulation Results

In this part of the simulations, ERSS estimators obtained by MMLE and BLUE methods are compared with SRS estimators. For this comparison, Monte Carlo simulations with 100,000 iterations are used under normal distribution with parameters $\mu = 0$ and $\sigma = 1$. Since the likelihood function differs for even and odd sized samples, the simulation results are given in distinct tables.

Table 2.4: Bias and MSE comparison for $\hat{\mu}_{srs}$, $\hat{\mu}_{erss}$ MMLE (denoted by ERSS*), and $\hat{\mu}_{erss}$ BLUE (denoted by ERSS**) for even sized samples

| | $\hat{\mu}$ | | | | | |
|---|---|---|---|---|---|---|
| | **Bias** | | | **MSE** | | |
| **n** | **SRS** | **ERSS*** | **ERSS**** | **SRS** | **ERSS*** | **ERSS**** |
| 4 | -0.0010 | 0.0011 | 0.0004 | 0.2487 | 0.1217 | 0.1233 |
| 6 | 0.0013 | 0.0012 | -0.0005 | 0.1670 | 0.0690 | 0.0695 |
| 8 | 0.0017 | 0.0005 | 0 | 0.1250 | 0.0465 | 0.0463 |
| 10 | 0.0005 | -0.0002 | -0.0006 | 0.0994 | 0.0342 | 0.0343 |

Table 2.5: Bias and MSE comparison for $\hat{\mu}_{srs}$, $\hat{\mu}_{erss}$ MMLE (denoted by ERSS*), and $\hat{\mu}_{erss}$ BLUE (denoted by ERSS**) for odd sized samples

| | $\hat{\mu}$ | | | | | |
|---|---|---|---|---|---|---|
| | **Bias** | | | **MSE** | | |
| **n** | **SRS** | **ERSS*** | **ERSS**** | **SRS** | **ERSS*** | **ERSS**** |
| 3 | 0.0007 | 0.0008 | -0.0021 | 0.3343 | 0.1718 | 0.1732 |
| 5 | 0.0001 | -0.0008 | 0.0008 | 0.1998 | 0.0807 | 0.0803 |
| 7 | 0 | 0.0006 | 0.0009 | 0.1423 | 0.0501 | 0.0500 |
| 9 | -0.0005 | -0.0002 | -0.0007 | 0.1119 | 0.0353 | 0.0349 |

According to Tables 2.4 and 2.5, the bias of $\hat{\mu}$ obtained by SRS and ERSS are all negligible. MSEs of $\hat{\mu}_{erss}$ are almost equal for MML and BLU estimators and both are less than the SRS counterparts.

Table 2.6: Bias and MSE comparison for $\hat{\sigma}_{srs}$, $\hat{\sigma}_{erss}$ MMLE (denoted by ERSS*), and $\hat{\sigma}_{erss}$ BLUE (denoted by ERSS**) for even sized samples

| | $\hat{\sigma}$ | | | | | |
|---|---|---|---|---|---|---|
| | **Bias** | | | **MSE** | | |
| **n** | **SRS** | **ERSS*** | **ERSS**** | **SRS** | **ERSS*** | **ERSS**** |
| 4 | -0.0794 | -0.0680 | -0.0006 | 0.1575 | 0.0796 | 0.1162 |
| 6 | -0.0485 | 0.0059 | -0.0004 | 0.0964 | 0.0376 | 0.0432 |
| 8 | -0.0343 | 0.0313 | 0 | 0.0699 | 0.0240 | 0.0230 |
| 10 | -0.0262 | 0.0448 | -0.0002 | 0.0546 | 0.0178 | 0.0145 |

The results in Tables 2.6 and 2.7 demonstrate that bias of $\hat{\sigma}_{erss}$ for both MMLE and BLUE methods are mostly less than $\hat{\sigma}_{srs}$. While comparing MSE, it is obvious that the ERSS estimator obtained by MMLE method yields the smallest MSE for small sample sizes. As the sample size increases, the MSEs of ERSS estimators for both MMLE and BLUE methods become almost equal but still less than that of $\hat{\sigma}_{srs}$. Relative efficiency table (see Table 2.8) shows the efficiency of the estimators under ERSS more clearly where the relative efficiency is defined as $RE = \frac{MSE_{srs}}{MSE_{erss}}$. Values of RE greater than 1 mean that ERSS estimators are more efficient. The results in Table 2.8 show that estimators under ERSS scheme for both MMLE and BLUE are

Table 2.7: Bias and MSE comparison for $\hat{\sigma}_{srs}$, $\hat{\sigma}_{erss}$ MMLE (denoted by ERSS*), and $\hat{\sigma}_{erss}$ BLUE (denoted by ERSS**) for odd sized samples

| | $\hat{\sigma}$ | | | | | |
|---|---|---|---|---|---|---|
| | **Bias** | | | **MSE** | | |
| **n** | **SRS** | **ERSS*** | **ERSS**** | **SRS** | **ERSS*** | **ERSS**** |
| 3 | -0.1132 | -0.1881 | 0.0006 | 0.2272 | 0.1621 | 0.3925 |
| 5 | -0.0604 | -0.0353 | 0 | 0.1204 | 0.0585 | 0.0828 |
| 7 | -0.0414 | 0.0133 | -0.0002 | 0.0814 | 0.0317 | 0.0355 |
| 9 | -0.0310 | 0.0344 | 0.0007 | 0.0613 | 0.0215 | 0.0203 |

more efficient than SRS. For small sample sizes, REs of ERSS estimators obtained by MMLE are slightly greater than that of BLUE and become almost equal when sample size increases. Moreover, all RE values increase as the sample size increases, which means that ERSS estimators become more efficient as the sample size gets larger.

Table 2.8: Relative efficiency of $\hat{\mu}$ and $\hat{\sigma}$ under SRS and ERSS

| | n | 4 | 6 | 8 | 10 | 3 | 5 | 7 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| **MMLE** | $\hat{\mu}$ | 2.043 | 2.42 | 2.69 | 2.91 | 1.94 | 2.48 | 2.84 | 3.17 |
| | $\hat{\sigma}$ | 1.98 | 2.56 | 2.91 | 3.07 | 1.40 | 2.06 | 2.56 | 2.85 |
| **BLUE** | $\hat{\mu}$ | 2.02 | 2.40 | 2.70 | 2.90 | 1.93 | 2.49 | 2.85 | 3.21 |
| | $\hat{\sigma}$ | 1.36 | 2.23 | 3.03 | 3.76 | 0.58 | 1.45 | 2.29 | 3.01 |

Tables 2.9- 2.10 give results of simulations for type I error levels of Shewhart control charts under SRS, ERSS by using MMLE and BLUE. In the simulations, normal distribution is used with parameters $\mu = 0$ and $\sigma = 1$, and control limits are constructed under $3\sigma$ limits with the target type I error level 0.0027. For both odd and even sized samples, control charts under ERSS obtained by MML estimators yield the smallest type I error when fewer samples of small sizes are considered. As the number of samples in sets and corresponding sample sizes increase, the type I errors for three different control charts become almost equal to each other and converge to the true type I error level 0.0027. It should be noted that the type I errors of control charts under ERSS scheme are continuously decreasing for even sized samples; however, for odd sized samples, there is not a monotonic decrease in the type I error. That is due to the design of odd sized samples.

Table 2.9: Type I error comparison for Shewhart control charts constructed under SRS and ERSS for even sized samples (ERSS* denotes control charts under ERSS and MMLE, ERSS** denotes charts under ERSS and BLUE)

| **M** | **20** | | | **30** | | |
|---|---|---|---|---|---|---|
| **n** | **SRS** | **ERSS*** | **ERSS**** | **SRS** | **ERSS*** | **ERSS**** |
| 4 | 0.0047 | 0.0043 | 0.0045 | 0.0040 | 0.0038 | 0.0039 |
| 6 | 0.0042 | 0.0039 | 0.0040 | 0.0037 | 0.0036 | 0.0036 |
| 8 | 0.0040 | 0.0038 | 0.0038 | 0.0035 | 0.0035 | 0.0035 |
| 10 | 0.0039 | 0.0037 | 0.0037 | 0.0034 | 0.0035 | 0.0035 |
| **M** | **50** | | | **100** | | |
| **n** | **SRS** | **ERSS*** | **ERSS**** | **SRS** | **ERSS*** | **ERSS**** |
| 4 | 0.0034 | 0.0035 | 0.0035 | 0.0031 | 0.0032 | 0.0032 |
| 6 | 0.0032 | 0.0033 | 0.0034 | 0.0030 | 0.0031 | 0.0031 |
| 8 | 0.0032 | 0.0033 | 0.0033 | 0.0029 | 0.0031 | 0.0031 |
| 10 | 0.0031 | 0.0033 | 0.0032 | 0.0029 | 0.0031 | 0.0031 |

Table 2.10: Type I error comparison for Shewhart control charts constructed under SRS and ERSS for odd sized samples (ERSS* denotes control charts under ERSS and MMLE, ERSS** denotes charts under ERSS and BLUE)

| **M** | **20** | | | **30** | | |
|---|---|---|---|---|---|---|
| **n** | **SRS** | **ERSS*** | **ERSS**** | **SRS** | **ERSS*** | **ERSS**** |
| 3 | 0.0055 | 0.0031 | 0.0042 | 0.0045 | 0.0039 | 0.0055 |
| 5 | 0.0044 | 0.0020 | 0.0030 | 0.0038 | 0.0025 | 0.0035 |
| 7 | 0.0041 | 0.0022 | 0.0019 | 0.0035 | 0.0021 | 0.0022 |
| 9 | 0.0039 | 0.0023 | 0.0021 | 0.0035 | 0.0026 | 0.0029 |
| **M** | **50** | | | **100** | | |
| **n** | **SRS** | **ERSS*** | **ERSS**** | **SRS** | **ERSS*** | **ERSS**** |
| 3 | 0.0037 | 0.0025 | 0.0031 | 0.0032 | 0.0038 | 0.0032 |
| 5 | 0.0033 | 0.0027 | 0.0027 | 0.0030 | 0.0026 | 0.0031 |
| 7 | 0.0032 | 0.0020 | 0.0026 | 0.0029 | 0.0022 | 0.0022 |
| 9 | 0.0031 | 0.0034 | 0.0017 | 0.0029 | 0.0030 | 0.0026 |

# CHAPTER 3

# QUALITY CONTROL CHARTS UNDER LONG TAILED SYMMETRIC DISTRIBUTION

Normality assumption may not always satisfied in the quality control applications. The distribution of the population of interest may be skewed, platykurtic and leptokurtic or may have light tails or heavy tails. In the literature, there are various techniques which enable researchers to construct Shewhart control charts for nonnormal distributions. For instance, Chan and Cui [23] proposed a method to eliminate the skewness of a distribution, which leads to construction of Shewhart control chart. Another example is the kurtosis correction method proposed by Tadikamalla and Popescu [87], who obtained a new structure to construct Shewhart control charts. For other robust techniques in quality control charts; see Cetinyurek [21].

As stated by Shewhart [78], normality assumption is no more a must since CLT is a very powerful and applicable theorem in constructing Shewhart control charts. However, to eliminate judgment ranking errors in RSS, small samples should be taken. Sizes of 3 to 10 but mostly sample sizes of 4, 5 or 6 are preferable in this respect. Therefore, CLT may not be well applicable in this case, and normality assumption may be violated. Because of these reasons, the distributions of all estimators should be checked to obtain correct quantile probabilities.

In this chapter, the most commonly encountered form of distribution in the quality control applications is studied. The distribution namely known as long-tailed symmetric is symmetric and has more mass in the tails. The long-tailed symmetric distribution can be defined as

$$f(x,p) = \frac{\Gamma(p)}{\Gamma(p-0.5)}\frac{1}{\sqrt{k\pi}\sigma}\left\{1+\frac{(x-\mu)^2}{k\sigma^2}\right\}^{-p}; \quad -\infty < x < \infty \tag{3.1}$$

where $p$ is the shape parameter, $k = 2p-3$ and $p \geq 2$.

This is a special form of symmetric $p$ family distributions (Tiku and Suresh [97]).

Mean, variance and kurtosis of this distribution are given as

$$E(X) = \mu \,, \qquad Var(X) = \sigma^2 \quad \text{and} \quad \beta_2 = \frac{3(p-3/2)}{(p-5/2)} \,. \tag{3.2}$$

The kurtosis value of this distribution decreases as the shape parameter increases, and gets the value 3 (the kurtosis of normal) when $p = \infty$. For instance, for the values of $p = 2.5, 3, 5, 10, \infty$, the kurtosis of the distribution becomes $\infty, 9, 4.2, 3.4, 3$, respectively. For $1 \leq p < 2$, k is set to 1, and variance does not exist for the distribution. In this case, $\sigma$ becomes just a scale parameter. For $p = 1$, expected value of the distribution does not exist and $\mu$ becomes just a location parameter. Long tailed symmetric family is very appropriate especially in modeling samples containing outliers (Tiku and Akkaya [95]). For this study the transformation $t = \sqrt{\nu/k}((x-\mu)/\sigma)$ is considered. Following this, the random variable $t$ has a Student's t distribution with $\nu = 2p - 1$. In the simulations, this transformation will be used to generate random variables as well as calculating the pdf and cdf of long-tailed symmetric distribution.

## 3.1 Quality Control Charts with Simple Random Sampling

In order to compare the performances of quality control charts under SRS and RSS, parameter estimators of long-tailed symmetric distribution under SRS is obtained first. As stated by Tiku and Suresh [97], Vaughan [101]-[100], and Islam et al. [46], obtaining explicit forms of parameter estimators is not easy due to the form of the likelihood function. Iterative methods also fail when there exist outliers (Puthenpura and Sinha [69]); therefore, an easy and effective way to obtain parameter estimators for long-tailed symmetric distribution is the use of MMLE method due to Tiku [92], and Tiku and Suresh [97].

The log likelihood function of long-tailed symmetric distribution under SRS is given by

$$lnL = const * -nln\sigma - p\sum_{i=1}^{n} ln\left\{1 + (1/k)z_i^2\right\} \tag{3.3}$$

where $z_i = (x_i - \mu)/\sigma$ and $-\infty < x_i < \infty$. The partial derivatives with respect to $\mu$ and $\sigma$ are,

$$\frac{\partial lnL}{\partial \mu} = \frac{2p}{k\sigma}\sum_{i=1}^{n} \frac{z_i}{\left\{1 + (1/k)z_i^2\right\}} \,, \tag{3.4}$$

$$\frac{\partial lnL}{\partial \sigma} = -\frac{n}{\sigma} + \frac{2p}{k\sigma}\sum_{i=1}^{n} \frac{z_i^2}{\left\{1 + (1/k)z_i^2\right\}} \,. \tag{3.5}$$

The partial derivatives of the likelihood function have the nonlinear term

$$g(z) = \frac{z}{1 + \frac{1}{k}z^2} \tag{3.6}$$

which leads to multiple roots for $p < \infty$ (Vaughan [101]-[100]). Hence, it is not possible to find explicit solutions for estimators of $\mu$ and $\sigma$ by MLE method.

### 3.1.1 Modified Maximum Likelihood Estimation

Since MLEs are intractable for long-tailed symmetric distribution under SRS, the parameter estimators can be found by MMLE method proposed by Tiku [92], and Tiku and Suresh [97]. To obtain MMLEs of parameters, let nonlinear functions be expressed as a function of $z_{(i)}$, $i = 1, 2, ..., n$, where $z_{(i)} = \frac{x_{(i)} - \mu}{\sigma}$. Then, the partial derivatives can be written as

$$\frac{\partial lnL}{\partial \mu} \cong \frac{\partial lnL^*}{\partial \mu} = \frac{2p}{k\sigma} \sum_{i=1}^{n} g(z_{(i)}) \ , \tag{3.7}$$

$$\frac{\partial lnL}{\partial \sigma} \cong \frac{\partial lnL^*}{\partial \sigma} = -\frac{n}{\sigma} + \frac{2p}{k\sigma} \sum_{i=1}^{n} z_{(i)} g(z_{(i)}) \ . \tag{3.8}$$

Let also the function $g(z_{(i)})$ be linearized by using Taylor Series expansion around the expected value of $Z_{(i)}$ ($E(Z_{(i)}) = t_{(i)}$). Some of $t_{(i)}$ values are tabulated in the papers by Tiku and Kumra [96], and Vaughan [101]-[100].

$$g(z_{(i)}) \cong \alpha_i + \beta_i z_{(i)} \tag{3.9}$$

where

$$\alpha_i = \frac{\frac{2}{k} t_{(i)}^3}{\left\{1 + \frac{1}{k} t_{(i)}^2\right\}^2} \ , \qquad \beta_i = \frac{1 - \frac{1}{k} t_{(i)}^2}{\left\{1 + \frac{1}{k} t_{(i)}^2\right\}^2} \ . \tag{3.10}$$

Replacing the nonlinear term with the corresponding linearized form of $g(z_{(i)})$ in the partial derivatives of the likelihood function, $\hat{\mu}$ and $\hat{\sigma}$ are obtained as

$$\hat{\mu} = \frac{\sum_{i=1}^{n} \beta_i x_{(i)}}{m} \ , \qquad \hat{\sigma} = \frac{B + \sqrt{B^2 + 4nC}}{2\sqrt{n(n-1)}} \tag{3.11}$$

where

$$m = \sum_{i=1}^{n} \beta_i \ , \qquad B = \frac{2p}{k} \sum_{i=1}^{n} \alpha_i x_{(i)} \qquad \text{and} \qquad C = \frac{2p}{k} \sum_{i=1}^{n} \beta_i (x_{(i)} - \hat{\mu})^2 \ . \tag{3.12}$$

If the value of $C$ is negative, $\hat{\sigma}$ is not real and positive; therefore, $\alpha_i$ and $\beta_i$ can be replaced with $\alpha_i^*$ and $\beta_i^*$, respectively (3.13); see also Tiku et al. [99]. $C$ often gets negative values when $p \leq 3$. When $p > 3$, $C$ gets negative values very rarely.

$$\alpha_i^* = 0\,, \qquad \beta_i^* = \frac{1}{\left\{1 + \frac{1}{k}t_{(i)}^2\right\}}\,. \tag{3.13}$$

The MML estimators are, as stated in Chapter 2, asymptotically fully efficient and unbiased.

**Distribution of $\hat{\mu}$:** After obtaining $\hat{\mu}$ by using MMLE method, its distribution should be investigated for small samples. This is important for constructing Shewhart control limits since the desired type I error level may not be achieved due to non-normality. For this reason, one can practically calculate the skewness and kurtosis of $\hat{\mu}$ by a simulation study based on 100,000 runs; see Table 3.1.

Table 3.1: Skewness and kurtosis values of $\hat{\mu}$ under SRS

| p | n | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| 3 | $\sqrt{\beta_1}$ | -0.013 | 0.063 | 0.008 | 0.004 | 0.009 | 0.006 | -0.004 | 0.015 |
| | $\beta_2$ | 4.262 | 4.593 | 3.499 | 3.371 | 3.292 | 3.244 | 3.197 | 3.185 |
| 3.5 | $\sqrt{\beta_1}$ | -0.010 | 0.018 | -0.011 | -0.005 | -0.011 | 0.004 | 0.011 | -0.004 |
| | $\beta_2$ | 3.814 | 3.483 | 3.375 | 3.305 | 3.229 | 3.157 | 3.183 | 3.156 |
| 5 | $\sqrt{\beta_1}$ | 0.007 | 0.008 | -0.005 | 0.008 | 0.006 | 0.006 | 0.001 | -0.005 |
| | $\beta_2$ | 3.352 | 3.265 | 3.187 | 3.181 | 3.133 | 3.086 | 3.115 | 3.067 |
| 10 | $\sqrt{\beta_1}$ | 0.001 | -0.009 | 0.004 | -0.020 | -0.001 | -0.006 | 0.006 | 0.006 |
| | $\beta_2$ | 3.139 | 3.087 | 3.098 | 3.060 | 3.060 | 3.045 | 3.041 | 3.067 |

From the results in Table 3.1, it is obvious that the distribution of $\hat{\mu}$ is symmetric for all $p$ values included in the simulation study. On the other hand, the kurtosis of $\hat{\mu}$ is always greater than 3; therefore, normality assumption fails especially for smaller shape parameter values. Therefore, one needs to use a different approach to deal with longer tails in constructing control chart limits.

## 3.2 Quality Control Charts with Ranked Set Sampling

In this section, estimators for long-tailed symmetric distribution parameters $\mu$ and $\sigma$ under RSS are obtained. Since long-tailed symmetric distribution is also of location-scale type, the likelihood can be defined as

$$\frac{1}{\sigma}f\left(\frac{x-\mu}{\sigma}\right) \propto \frac{1}{\sigma}\left\{1 + \frac{(x-\mu)^2}{k\sigma^2}\right\}^{-p}; \qquad -\infty < x < \infty\,. \tag{3.14}$$

Further, the log-likelihood function can be written as

$$lnL \cong -nln\sigma + \sum_{i=1}^{n} lnf(z_{(i)}) + \sum_{i=1}^{n}(i-1)lnF(z_{(i)}) + \sum_{i=1}^{n}(n-i)ln(1-F(z_{(i)})) \quad (3.15)$$

where $z_{(i)} = (x_{(i)} - \mu)/\sigma$ and $F(z)$ is the cdf of the long-tailed symmetric distribution. The partial derivatives of the log likelihood are

$$\begin{aligned} \frac{\partial lnL}{\partial \mu} &= \left(\frac{1}{\sigma}\right)\left[\frac{2p}{k}\sum_{i=1}^{n}\frac{z_{(i)}}{1+\frac{1}{k}z_{(i)}^2} - \sum_{i=1}^{n}(i-1)\frac{f(z_{(i)})}{F(z_{(i)})}\right. \\ &\left. \quad +\sum_{i=1}^{n}(n-i)\frac{f(z_{(i)})}{1-F(z_{(i)})}\right], \qquad (3.16) \\ \frac{\partial lnL}{\partial \sigma} &= \left(\frac{1}{\sigma}\right)\left[\frac{2p}{k}\sum_{i=1}^{n}\frac{z_{(i)}}{1+\frac{1}{k}z_{(i)}^2}z_{(i)} - \sum_{i=1}^{n}(i-1)\frac{f(z_{(i)})}{F(z_{(i)})}z_{(i)}\right. \\ &\left. \quad +\sum_{i=1}^{n}(n-i)\frac{f(z_{(i)})}{1-F(z_{(i)})}z_{(i)} - n\right]. \qquad (3.17) \end{aligned}$$

The estimators of $\mu$ and $\sigma$ can not be obtained explicitly by MLE method due to the nonlinear terms in the following equations:

$$g_1(z_{(i)}) = \frac{z_{(i)}}{1+\frac{1}{k}z_{(i)}^2}, \quad g_2(z_{(i)}) = \frac{f(z_{(i)})}{F(z_{(i)})}, \quad g_3(z_{(i)}) = \frac{f(z_{(i)})}{1-F(z_{(i)})}; \quad 1 \le i \le n. \qquad (3.18)$$

### 3.2.1 Modified Maximum Likelihood Estimation

Because of the fact that the closed form solutions of the estimators can not be obtained by MLE method, MMLE can be applied in this section. To do that, the nonlinear terms in 3.18 can be linearized by Taylor Series expansion around the expected value of $z_{(i)}$. Write,

$$\begin{aligned} g_1(z_{(i)}) &\cong \alpha_{1i} + \beta_{1i}z_{(i)}, \qquad (3.19) \\ g_2(z_{(i)}) &\cong \alpha_{2i} - \beta_{2i}z_{(i)}, \qquad (3.20) \\ g_3(z_{(i)}) &\cong \alpha_{3i} + \beta_{3i}z_{(i)}; \quad 1 \le i \le n, \qquad (3.21) \end{aligned}$$

where

$$\beta_{1i} = \frac{1-\frac{1}{k}t_{(i)}^2}{\left\{1+\frac{1}{k}t_{(i)}^2\right\}^2}, \qquad \alpha_{1i} = \frac{\frac{2}{k}t_{(i)}^3}{\left\{1+\frac{1}{k}t_{(i)}^2\right\}^2}, \qquad (3.22)$$

$$\beta_{2i} = \frac{2pt_{(i)}}{k+t_{(i)}^2}\frac{f(t_{(i)})}{F(t_{(i)})} + \left(\frac{f(t_{(i)})}{F(t_{(i)})}\right)^2, \qquad \alpha_{2i} = \frac{f(t_{(i)})}{F(t_{(i)})} + t_{(i)}\beta_{2i}\ , \tag{3.23}$$

$$\beta_{3i} = \left(\frac{f(t_{(i)})}{1-F(t_{(i)})}\right)^2 - \frac{2pt_{(i)}}{k+t_{(i)}^2}\frac{f(t_{(i)})}{1-F(t_{(i)})} \quad \text{and} \quad \alpha_{3i} = \frac{f(t_{(i)})}{1-F(t_{(i)})} - t_{(i)}\beta_{3i}\ . \tag{3.24}$$

Thus, the MML estimators are obtained as follows:

$$\hat{\mu} = \sum_{i=1}^{n} a_i x_{(i)}\ ; \qquad a_i = \frac{u_i}{\sum_{i=1}^{n} u_i} \tag{3.25}$$

and

$$\hat{\sigma} = \frac{-B + \sqrt{B^2 + 4nC}}{2n}\ , \tag{3.26}$$

where

$$B = \sum_{i=1}^{n} x_{(i)} w_i\ , \qquad C = \sum_{i=1}^{n} \left(x_{(i)} - \hat{\mu}\right)^2 u_i\ , \tag{3.27}$$

$$u_i = \frac{2p}{k}\beta_{1i} + (i-1)\beta_{2i} + (n-i)\beta_{3i}\ , \tag{3.28}$$

$$w_i = (i-1)\alpha_{2i} - (n-i)\alpha_{3i} - \frac{2p}{k}\alpha_{1i}\ . \tag{3.29}$$

If $C$ is less than zero, then $\alpha_{1i}$ and $\beta_{1i}$ can be replaced practically with $\alpha_{1i}^*$ and $\beta_{1i}^*$ (Equation 3.13), respectively, so as to obtain a real and positive root of $\hat{\sigma}$ as mentioned in Section 3.1.1.

**Distribution of $\hat{\mu}$:** Before constructing Shewhart control charts, the distribution of the estimator $\hat{\mu}$ should be investigated as mentioned at the beginning of this chapter. Table 3.2 exhibits the skewness ($\sqrt{\beta_1}$) and kurtosis ($\beta_2$) values of the estimator $\hat{\mu}$ under different $p$ parameters obtained by Monte Carlo simulations.

Similar to $\hat{\mu}$ obtained under SRS, the distribution of the estimator $\hat{\mu}$ under RSS has a similar structure. It has a long-tailed symmetric distribution especially for smaller sample sizes and shape parameter values.

## 3.3 Simulation Results

To compare the estimators obtained by SRS and RSS under long-tailed symmetric distribution, Monte Carlo simulations with 100,000 repetitions were conducted. In

Table 3.2: Skewness and kurtosis values of $\hat{\mu}$ under RSS

| p | n | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| 3 | $\sqrt{\beta_1}$ | -0.025 | -0.003 | -0.002 | -0.012 | 0 | -0.010 | -0.024 | 0.018 |
| | $\beta_2$ | 4.532 | 3.690 | 3.419 | 3.313 | 3.200 | 3.126 | 3.098 | 3.138 |
| 3.5 | $\sqrt{\beta_1}$ | 0.008 | -0.014 | 0.010 | -0.014 | -0.007 | -0.010 | 0.012 | 0.007 |
| | $\beta_2$ | 3.904 | 3.511 | 3.343 | 3.223 | 3.141 | 3.090 | 3.096 | 3.045 |
| 5 | $\sqrt{\beta_1}$ | -0.003 | 0.023 | 0 | -0.001 | -0.002 | -0.003 | -0.003 | 0.007 |
| | $\beta_2$ | 3.404 | 3.233 | 3.146 | 3.149 | 3.040 | 3.031 | 3.032 | 3.056 |
| 10 | $\sqrt{\beta_1}$ | 0 | 0.001 | -0.001 | 0.001 | 0.014 | -0.003 | -0.012 | -0.002 |
| | $\beta_2$ | 3.162 | 3.097 | 3.051 | 3.036 | 3.070 | 3.021 | 3.033 | 3.006 |

Table 3.3: Comparison of RSS and SRS in the estimation of long-tailed symmetric distribution parameters with $p = 3$

| | $\hat{\mu}$ | | | | $\hat{\sigma}$ | | | |
|---|---|---|---|---|---|---|---|---|
| | **Bias** | | **MSE** | | **Bias** | | **MSE** | |
| **n** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** |
| 3 | -0.0009 | -0.0028 | 0.3166 | 0.1756 | 0.0529 | -0.0528 | 0.4482 | 0.2786 |
| 4 | 0.0017 | 0.0009 | 0.2266 | 0.0988 | 0.0647 | 0.0153 | 0.3296 | 0.2005 |
| 5 | -0.0006 | -0.0006 | 0.1765 | 0.0638 | 0.0645 | 0.0540 | 0.2480 | 0.1536 |
| 6 | 0.0004 | 0.0008 | 0.1439 | 0.0446 | 0.0627 | 0.0785 | 0.1994 | 0.1312 |
| 7 | -0.0004 | 0.0011 | 0.1230 | 0.0335 | 0.0591 | 0.0829 | 0.1677 | 0.1081 |
| 8 | 0.0013 | 0.0010 | 0.1063 | 0.0251 | 0.0580 | 0.0970 | 0.1432 | 0.0931 |
| 9 | 0.0003 | 0.0003 | 0.0937 | 0.0199 | 0.0518 | 0.1007 | 0.1230 | 0.0783 |
| 10 | -0.0013 | 0.0026 | 0.0838 | 0.0166 | 0.0488 | 0.1054 | 0.1077 | 0.0739 |

Table 3.4: Comparison of RSS and SRS in the estimation of long-tailed symmetric distribution parameters with $p = 3.5$

| | $\hat{\mu}$ | | | | $\hat{\sigma}$ | | | |
|---|---|---|---|---|---|---|---|---|
| | **Bias** | | **MSE** | | **Bias** | | **MSE** | |
| **n** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** |
| 3 | 0.0005 | 0.0000 | 0.3215 | 0.1697 | 0.0190 | -0.0779 | 0.3823 | 0.2380 |
| 4 | 0.0007 | -0.0025 | 0.2355 | 0.1009 | 0.0363 | -0.0032 | 0.2734 | 0.1800 |
| 5 | -0.0033 | -0.0009 | 0.1858 | 0.0673 | 0.0430 | 0.0386 | 0.2157 | 0.1416 |
| 6 | -0.0011 | 0.0004 | 0.1525 | 0.0473 | 0.0419 | 0.0523 | 0.1730 | 0.1107 |
| 7 | -0.0007 | 0.0020 | 0.1296 | 0.0351 | 0.0428 | 0.0690 | 0.1448 | 0.0909 |
| 8 | 0.0001 | 0.0011 | 0.1122 | 0.0267 | 0.0414 | 0.0782 | 0.1246 | 0.0797 |
| 9 | -0.0019 | -0.0021 | 0.0997 | 0.0212 | 0.0390 | 0.0869 | 0.1080 | 0.0681 |
| 10 | -0.0007 | 0.0025 | 0.0893 | 0.0177 | 0.0379 | 0.0855 | 0.0959 | 0.0569 |

the simulations, four different $p$ parameters were used, $p = 3, 3.5, 5, 10$ and the other parameters were set as $\mu = 0$ and $\sigma = 1$.

As can be seen from Tables 3.3, 3.4, 3.5, and 3.6, the biases of the estimators

Table 3.5: Comparison of RSS and SRS in the estimation of long-tailed symmetric distribution parameters with $p = 5$

| | $\hat{\mu}$ | | | | $\hat{\sigma}$ | | | |
|---|---|---|---|---|---|---|---|---|
| | **Bias** | | **MSE** | | **Bias** | | **MSE** | |
| n | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** |
| 3 | 0.0010 | 0.0043 | 0.3283 | 0.1769 | -0.0261 | -0.1159 | 0.3080 | 0.1968 |
| 4 | 0.0000 | -0.0006 | 0.2446 | 0.1032 | -0.0034 | -0.0449 | 0.2169 | 0.1375 |
| 5 | 0.0023 | 0.0008 | 0.1947 | 0.0684 | 0.0072 | 0.0053 | 0.1696 | 0.1050 |
| 6 | -0.0009 | 0.0006 | 0.1604 | 0.0490 | 0.0107 | 0.0236 | 0.1385 | 0.0832 |
| 7 | 0.0021 | 0.0005 | 0.1372 | 0.0357 | 0.0133 | 0.0326 | 0.1151 | 0.0658 |
| 8 | 0.0022 | -0.0004 | 0.1214 | 0.0278 | 0.0165 | 0.0441 | 0.1001 | 0.0561 |
| 9 | -0.0004 | -0.0033 | 0.1060 | 0.0221 | 0.0145 | 0.0496 | 0.0871 | 0.0478 |
| 10 | -0.0004 | -0.0003 | 0.0957 | 0.0178 | 0.0163 | 0.0566 | 0.0788 | 0.0433 |

Table 3.6: Comparison of RSS and SRS in the estimation of long-tailed symmetric distribution parameters with $p = 10$

| | $\hat{\mu}$ | | | | $\hat{\sigma}$ | | | |
|---|---|---|---|---|---|---|---|---|
| | **Bias** | | **MSE** | | **Bias** | | **MSE** | |
| n | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** |
| 3 | 0.0015 | 0.0000 | 0.3324 | 0.1732 | -0.0739 | -0.1564 | 0.2546 | 0.1739 |
| 4 | -0.0017 | 0.0015 | 0.2488 | 0.1016 | -0.0443 | -0.0799 | 0.1801 | 0.1138 |
| 5 | -0.0002 | -0.0039 | 0.1989 | 0.0703 | -0.0304 | -0.0372 | 0.1385 | 0.0826 |
| 6 | 0.0009 | -0.0037 | 0.1645 | 0.0495 | -0.0211 | -0.0176 | 0.1109 | 0.0624 |
| 7 | -0.0002 | 0.0045 | 0.1421 | 0.0376 | -0.0167 | -0.0015 | 0.0933 | 0.0507 |
| 8 | -0.0005 | -0.0021 | 0.1243 | 0.0288 | -0.0117 | 0.0123 | 0.0812 | 0.0421 |
| 9 | -0.0011 | -0.0009 | 0.1101 | 0.0225 | -0.0072 | 0.0208 | 0.0717 | 0.0351 |
| 10 | -0.0005 | 0.0002 | 0.0995 | 0.0187 | -0.0061 | 0.0292 | 0.0637 | 0.0309 |

of $\mu$ under both SRS and RSS are negligible, and the MSEs of the estimators under RSS are quite smaller than SRS counterparts. Likewise, the estimators of $\sigma$ under RSS yield always much smaller MSEs. To see the efficiencies of RSS estimators more clearly, relative efficiency (RE) table is compiled such that it is defined as $RE = MSE_{srs}/MSE_{rss}$ (see Table 3.7). RE greater than 1 means RSS is more efficient.

Table 3.7 shows that RSS is always more efficient than SRS. Moreover, as the sample size increases the efficiency of RSS increases too. However, since normality assumption is not satisfied, type I errors do not converge to the real value 0.0027 under $3\sigma$ limits. To solve this problem, a moment approach can be useful and the tails of the distribution of $\hat{\mu}$ can be modeled more efficiently.

Table 3.7: Relative efficiency of $\hat{\mu}$ and $\hat{\sigma}$ under SRS and RSS

| p | n | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| 3 | $\hat{\mu}$ | 1.80 | 2.29 | 2.77 | 3.23 | 3.67 | 4.24 | 4.70 | 5.04 |
| | $\hat{\sigma}$ | 1.61 | 1.64 | 1.61 | 1.52 | 1.55 | 1.54 | 1.57 | 1.46 |
| 3.5 | $\hat{\mu}$ | 1.89 | 2.33 | 2.76 | 3.23 | 3.69 | 4.20 | 4.71 | 5.04 |
| | $\hat{\sigma}$ | 1.61 | 1.52 | 1.52 | 1.56 | 1.59 | 1.56 | 1.59 | 1.68 |
| 5 | $\hat{\mu}$ | 1.86 | 2.37 | 2.85 | 3.28 | 3.84 | 4.36 | 4.79 | 5.37 |
| | $\hat{\sigma}$ | 1.57 | 1.58 | 1.61 | 1.66 | 1.75 | 1.78 | 1.82 | 1.82 |
| 10 | $\hat{\mu}$ | 1.92 | 2.45 | 2.83 | 3.32 | 3.78 | 4.32 | 4.90 | 5.32 |
| | $\hat{\sigma}$ | 1.46 | 1.58 | 1.68 | 1.78 | 1.84 | 1.93 | 2.04 | 2.06 |

Table 3.8: Type I error comparison for Shewhart control charts under SRS and RSS with normal approximation for long tailed symmetric distribution with $p = 3$

| **M** | **20** | | **30** | | **50** | | **100** | |
|---|---|---|---|---|---|---|---|---|
| **n** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** |
| 3 | 0.0108 | 0.0102 | 0.0095 | 0.0089 | 0.0086 | 0.0086 | 0.0080 | 0.0079 |
| 4 | 0.0089 | 0.0076 | 0.0080 | 0.0070 | 0.0073 | 0.0065 | 0.0067 | 0.0061 |
| 5 | 0.0075 | 0.0064 | 0.0068 | 0.0058 | 0.0062 | 0.0054 | 0.0057 | 0.0051 |
| 6 | 0.0066 | 0.0056 | 0.0060 | 0.0053 | 0.0055 | 0.0048 | 0.0051 | 0.0045 |
| 7 | 0.0064 | 0.0053 | 0.0058 | 0.0047 | 0.0053 | 0.0043 | 0.0049 | 0.0041 |
| 8 | 0.0059 | 0.0047 | 0.0054 | 0.0043 | 0.0050 | 0.0040 | 0.0046 | 0.0037 |
| 9 | 0.0053 | 0.0045 | 0.0048 | 0.0041 | 0.0045 | 0.0038 | 0.0042 | 0.0036 |
| 10 | 0.0052 | 0.0042 | 0.0047 | 0.0039 | 0.0043 | 0.0036 | 0.0040 | 0.0034 |

Table 3.9: Type I error comparison for Shewhart control charts under SRS and RSS with normal approximation for long tailed symmetric distribution with $p = 3.5$

| **M** | **20** | | **30** | | **50** | | **100** | |
|---|---|---|---|---|---|---|---|---|
| **n** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** |
| 3 | 0.0097 | 0.0089 | 0.0085 | 0.0077 | 0.0077 | 0.0072 | 0.0070 | 0.0068 |
| 4 | 0.0081 | 0.0071 | 0.0072 | 0.0062 | 0.0065 | 0.0058 | 0.0060 | 0.0055 |
| 5 | 0.0070 | 0.0058 | 0.0062 | 0.0054 | 0.0056 | 0.0049 | 0.0053 | 0.0046 |
| 6 | 0.0061 | 0.0053 | 0.0054 | 0.0048 | 0.0049 | 0.0044 | 0.0046 | 0.0041 |
| 7 | 0.0057 | 0.0049 | 0.0051 | 0.0044 | 0.0047 | 0.0041 | 0.0044 | 0.0038 |
| 8 | 0.0054 | 0.0046 | 0.0049 | 0.0042 | 0.0045 | 0.0039 | 0.0042 | 0.0036 |
| 9 | 0.0051 | 0.0042 | 0.0046 | 0.0039 | 0.0042 | 0.0036 | 0.0040 | 0.0034 |
| 10 | 0.0053 | 0.0041 | 0.0047 | 0.0038 | 0.0044 | 0.0035 | 0.0041 | 0.0033 |

## 3.4 Three-Moment t Approximation

Consider a random variable $Y$ from a distribution such that

$$|\sqrt{\beta_1}| < 0.1 \qquad \text{and} \qquad \beta_2 \geq 3\ ,$$

Table 3.10: Type I error comparison for Shewhart control charts under SRS and RSS with normal approximation for long tailed symmetric distribution with $p = 5$

| M | 20 | | 30 | | 50 | | 100 | |
|---|---|---|---|---|---|---|---|---|
| **n** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** |
| 3 | 0.0077 | 0.0073 | 0.0066 | 0.0064 | 0.0058 | 0.0058 | 0.0052 | 0.0054 |
| 4 | 0.0063 | 0.0059 | 0.0055 | 0.0052 | 0.0049 | 0.0048 | 0.0045 | 0.0045 |
| 5 | 0.0059 | 0.0052 | 0.0051 | 0.0046 | 0.0046 | 0.0042 | 0.0043 | 0.0040 |
| 6 | 0.0053 | 0.0046 | 0.0047 | 0.0043 | 0.0043 | 0.0040 | 0.0040 | 0.0037 |
| 7 | 0.0050 | 0.0045 | 0.0045 | 0.0041 | 0.0041 | 0.0037 | 0.0038 | 0.0035 |
| 8 | 0.0050 | 0.0041 | 0.0045 | 0.0037 | 0.0041 | 0.0035 | 0.0038 | 0.0033 |
| 9 | 0.0044 | 0.0040 | 0.0040 | 0.0037 | 0.0036 | 0.0034 | 0.0034 | 0.0032 |
| 10 | 0.0047 | 0.0038 | 0.0042 | 0.0036 | 0.0039 | 0.0033 | 0.0036 | 0.0031 |

Table 3.11: Type I error comparison for Shewhart control charts under SRS and RSS with normal approximation for long tailed symmetric distribution with $p = 10$

| M | 20 | | 30 | | 50 | | 100 | |
|---|---|---|---|---|---|---|---|---|
| **n** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** |
| 3 | 0.0066 | 0.0055 | 0.0054 | 0.0049 | 0.0046 | 0.0044 | 0.0041 | 0.0040 |
| 4 | 0.0054 | 0.0049 | 0.0047 | 0.0043 | 0.0041 | 0.0039 | 0.0037 | 0.0036 |
| 5 | 0.0050 | 0.0044 | 0.0044 | 0.0040 | 0.0038 | 0.0037 | 0.0035 | 0.0034 |
| 6 | 0.0048 | 0.0041 | 0.0042 | 0.0038 | 0.0038 | 0.0034 | 0.0035 | 0.0032 |
| 7 | 0.0046 | 0.0040 | 0.0040 | 0.0037 | 0.0037 | 0.0033 | 0.0034 | 0.0032 |
| 8 | 0.0044 | 0.0039 | 0.0038 | 0.0036 | 0.0035 | 0.0033 | 0.0032 | 0.0031 |
| 9 | 0.0042 | 0.0038 | 0.0037 | 0.0035 | 0.0034 | 0.0032 | 0.0032 | 0.0030 |
| 10 | 0.0042 | 0.0038 | 0.0038 | 0.0034 | 0.0034 | 0.0032 | 0.0032 | 0.0030 |

where $\sqrt{\beta_1}$ and $\beta_2$ are the skewness and kurtosis of Y, respectively. Then a shifted form of the random variable $Y$ has Student's t distribution with $\nu$ degrees of freedom.

$$\frac{Y + a}{b} \sim t_\nu \tag{3.30}$$

Constants $a$, $b$, and $\nu$ can be obtained by equating the first three moments on both sides of 3.30 (Tiku [91]).

Considering the distribution of this random variable, the absolute value of the skewness is less than 0.1 for all the $p$ parameters included in the simulation and for all sample sizes. Moreover, the kurtosis is always greater than 3, hence the $t$ approximation can be applied to construct Shewhart control chart for the process mean under both SRS and RSS.

**Quality Control Limits with $t$ Approximation under SRS:** Quality control limits can be obtained with the help of three-moment t-approximation.

$$\begin{aligned} UCL &= \bar{\hat{\mu}}_{srs} + t_{\nu,\alpha/2} b \ , \\ CL &= \bar{\hat{\mu}}_{srs} \ , \\ LCL &= \bar{\hat{\mu}}_{srs} - t_{\nu,\alpha/2} b \ , \end{aligned} \tag{3.31}$$

where

$$b = \sqrt{\frac{\nu - 2}{\nu} Var(\hat{\mu}_{srs})} \ , \qquad \nu = \frac{2(2\beta_2 - 3)}{\beta_2 - 3} \ , \tag{3.32}$$

and $\beta_2$ is the kurtosis of the distribution of $\hat{\mu}_{srs}$. Since the distribution is symmetric, $a$ is equal to zero. The exact variance of $\hat{\mu}_{srs}$ can not be easily obtained; therefore, one needs to use the asymptotic variance (see Tiku and Akkaya, 2004 [95])

$$Var(\hat{\mu}_{srs}) = \frac{\sigma^2}{M} \tag{3.33}$$

where

$$M = \frac{2pm}{k} \ , \qquad m = \sum_{i=1}^{n} \beta_i \tag{3.34}$$

and $\beta_i$ is defined in Equation 3.10. $\sigma$ can be estimated by $c_1 \hat{\sigma}_{srs}$. Here $c_1$ is a constant which can be obtained by simulation and makes $\hat{\sigma}_{srs}$ an unbiased estimator of $\sigma$.

**Quality Control Limits with $t$ Approximation under RSS:** Shewhart control chart limits with $t$ approximation can be found in the same form

$$\begin{aligned} UCL &= \bar{\hat{\mu}}_{rss} + t_{\nu,\alpha/2} b \ , \\ CL &= \bar{\hat{\mu}}_{rss} \ , \\ LCL &= \bar{\hat{\mu}}_{rss} - t_{\nu,\alpha/2} b \ , \end{aligned} \tag{3.35}$$

where

$$b = \sqrt{\frac{\nu - 2}{\nu} Var(\hat{\mu}_{rss})} \ , \qquad \nu = \frac{2(2\beta_2 - 3)}{\beta_2 - 3} \ . \tag{3.36}$$

$\beta_2$ is the kurtosis of the distribution of $\hat{\mu}_{rss}$, and variance of $\hat{\mu}_{rss}$ can be written as

$$Var(\hat{\mu}_{rss}) = \frac{\sigma^2 \sum_{i=1}^{n} u_i^2 Var(Z_{(i)})}{\sum_{i=1}^{n} u_i^{\ 2}} \tag{3.37}$$

where $u_i$ is defined in Equation 3.28 and $Var(Z_{(i)})$ is simply the variance of the standardized order statistic $z_{(i)}$

### 3.4.1 Simulations with Three-Moment t Approximation

At the beginning of this chapter, it has been shown that the estimators of the process mean under both SRS and RSS are not distributed as normal, hence the classical Shewhart control chart is not relevant for this case. In this part of simulations, Shewhart control charts are constructed with the three-moment t-approximation as explained above. Long tailed symmetric distribution with parameters $\mu = 0$, $\sigma = 1$, and $p = 3, 3.5, 5, 10$ are used in the simulations and the type I error is chosen as 0.0027.

Table 3.12: Type I error comparison for Shewhart control charts under SRS and RSS with three-moment t-approximation for long tailed symmetric distribution with $p = 3$

| M | 20 | | 30 | | 50 | | 100 | |
|---|---|---|---|---|---|---|---|---|
| n | SRS | RSS | SRS | RSS | SRS | RSS | SRS | RSS |
| 3 | 0.0038 | 0.0030 | 0.0033 | 0.0029 | 0.0030 | 0.0025 | 0.0027 | 0.0025 |
| 4 | 0.0036 | 0.0035 | 0.0034 | 0.0030 | 0.0029 | 0.0029 | 0.0028 | 0.0027 |
| 5 | 0.0036 | 0.0036 | 0.0035 | 0.0033 | 0.0030 | 0.0030 | 0.0028 | 0.0028 |
| 6 | 0.0038 | 0.0033 | 0.0035 | 0.0031 | 0.0031 | 0.0028 | 0.0028 | 0.0027 |
| 7 | 0.0042 | 0.0035 | 0.0036 | 0.0033 | 0.0033 | 0.0029 | 0.0030 | 0.0028 |
| 8 | 0.0039 | 0.0037 | 0.0035 | 0.0033 | 0.0033 | 0.0032 | 0.0030 | 0.0030 |
| 9 | 0.0038 | 0.0036 | 0.0035 | 0.0035 | 0.0032 | 0.0031 | 0.0029 | 0.0029 |
| 10 | 0.0037 | 0.0032 | 0.0033 | 0.0030 | 0.0030 | 0.0027 | 0.0028 | 0.0026 |

Table 3.13: Type I error comparison for Shewhart control charts under SRS and RSS with three-moment t-approximation for long tailed symmetric distribution with $p = 3.5$

| M | 20 | | 30 | | 50 | | 100 | |
|---|---|---|---|---|---|---|---|---|
| n | SRS | RSS | SRS | RSS | SRS | RSS | SRS | RSS |
| 3 | 0.0042 | 0.0036 | 0.0038 | 0.0031 | 0.0032 | 0.0028 | 0.0028 | 0.0026 |
| 4 | 0.0043 | 0.0034 | 0.0036 | 0.0031 | 0.0034 | 0.0029 | 0.0031 | 0.0027 |
| 5 | 0.0040 | 0.0035 | 0.0035 | 0.0030 | 0.0033 | 0.0028 | 0.0029 | 0.0026 |
| 6 | 0.0037 | 0.0035 | 0.0032 | 0.0032 | 0.0029 | 0.0029 | 0.0027 | 0.0028 |
| 7 | 0.0039 | 0.0036 | 0.0033 | 0.0033 | 0.0031 | 0.0031 | 0.0029 | 0.0028 |
| 8 | 0.0041 | 0.0038 | 0.0037 | 0.0035 | 0.0034 | 0.0032 | 0.0031 | 0.0030 |
| 9 | 0.0039 | 0.0035 | 0.0034 | 0.0032 | 0.0030 | 0.0029 | 0.0028 | 0.0028 |
| 10 | 0.0039 | 0.0039 | 0.0037 | 0.0034 | 0.0032 | 0.0032 | 0.0030 | 0.0031 |

Simulations of Shewhart control charts with three-moment t-approximation show that the type I errors approach to the target value 0.0027. Moreover, control charts under RSS give smaller type I errors.

Table 3.14: Type I error comparison for Shewhart control charts under SRS and RSS with three-moment t-approximation for long tailed symmetric distribution with $p = 5$

| **M** | **20** | | **30** | | **50** | | **100** | |
|---|---|---|---|---|---|---|---|---|
| **n** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** |
| 3 | 0.0046 | 0.0042 | 0.0040 | 0.0036 | 0.0034 | 0.0033 | 0.0029 | 0.0029 |
| 4 | 0.0043 | 0.0038 | 0.0036 | 0.0036 | 0.0032 | 0.0032 | 0.0028 | 0.0029 |
| 5 | 0.0044 | 0.0040 | 0.0038 | 0.0035 | 0.0033 | 0.0032 | 0.0030 | 0.0029 |
| 6 | 0.0039 | 0.0036 | 0.0033 | 0.0033 | 0.0030 | 0.0030 | 0.0028 | 0.0027 |
| 7 | 0.0041 | 0.0041 | 0.0035 | 0.0037 | 0.0031 | 0.0034 | 0.0029 | 0.0032 |
| 8 | 0.0043 | 0.0039 | 0.0039 | 0.0035 | 0.0034 | 0.0032 | 0.0032 | 0.0030 |
| 9 | 0.0035 | 0.0037 | 0.0031 | 0.0035 | 0.0028 | 0.0031 | 0.0027 | 0.0030 |
| 10 | 0.0041 | 0.0035 | 0.0037 | 0.0032 | 0.0032 | 0.0030 | 0.0031 | 0.0028 |

Table 3.15: Type I error comparison for Shewhart control charts under SRS and RSS with three-moment t-approximation for long tailed symmetric distribution with $p = 10$

| **M** | **20** | | **30** | | **50** | | **100** | |
|---|---|---|---|---|---|---|---|---|
| **n** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** |
| 3 | 0.0055 | 0.0046 | 0.0042 | 0.0036 | 0.0035 | 0.0033 | 0.0032 | 0.0029 |
| 4 | 0.0047 | 0.0041 | 0.0040 | 0.0036 | 0.0035 | 0.0032 | 0.0031 | 0.0029 |
| 5 | 0.0041 | 0.0041 | 0.0037 | 0.0037 | 0.0031 | 0.0032 | 0.0029 | 0.0030 |
| 6 | 0.0043 | 0.0040 | 0.0037 | 0.0034 | 0.0034 | 0.0032 | 0.0030 | 0.0030 |
| 7 | 0.0041 | 0.0035 | 0.0036 | 0.0032 | 0.0033 | 0.0028 | 0.0030 | 0.0027 |
| 8 | 0.0039 | 0.0037 | 0.0035 | 0.0033 | 0.0032 | 0.0031 | 0.0030 | 0.0030 |
| 9 | 0.0038 | 0.0036 | 0.0034 | 0.0031 | 0.0030 | 0.0030 | 0.0029 | 0.0028 |
| 10 | 0.0037 | 0.0036 | 0.0033 | 0.0033 | 0.0030 | 0.0031 | 0.0028 | 0.0029 |

# CHAPTER 4

# QUALITY CONTROL CHARTS UNDER SHORT TAILED SYMMETRIC DISTRIBUTION

As explained in Chapter 3, normality assumption is not very practical in quality control applications. Most of the data have more mass in tails; therefore long tailed symmetric distribution provides a good fit to that type of data. Sometimes, researchers may also encounter with some data having shorter tails, and the use of short tailed symmetric distribution could be a good solution for this situation. Although there are very few applications on short-tailed symmetric distributions in the literature, Akkaya and Tiku [2] claim that some real life data can fit to this distribution perfectly. Data sets available by Hand et al. [39], Kendall and Stuart [47], Montgomery and Peck [58], and Atkinson and Riani [12], are examples of a short-tailed symmetric distribution in practice. Advantage of the short tailed symmetric distribution is that it is appropriate for modeling samples containing inliers which have erroneous information in the middle of the sample.

Tiku and Vaughan [98] defined a family of short tailed symmetric distributions as

$$f(x) = \frac{K}{\sqrt{2\pi}\sigma}\left\{1+\frac{1}{2h}\left(\frac{x-\mu}{\sigma}\right)^2\right\}^2 \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}\ ; \quad -\infty < x < \infty \tag{4.1}$$

where $h = 2-d$ and $d < 2$ is a constant and $K = \frac{1}{\left\{\sum_{j=0}^{2}\binom{2}{j}\left(\frac{1}{2h}\right)^j \frac{(2j)!}{2^j j!}\right\}}$.

This distribution is a symmetric distribution and its kurtosis depends on the value of $d$ parameter. As $d$ increases the kurtosis of $X$ decreases. For instance, for the values of $d = 1.5, 1, 0.5, 0, -0.5, -\infty$, kurtosis takes values $1.71, 2.03, 2.26, 2.44, 2.56, 3$, respectively. The kurtosis of short-tailed symmetric distribution is always less than 3 and approaches to 3 only when d goes to $-\infty$. The distribution is always unimodal for $d \leq 0$, and it is generally multi-modal for $d > 0$; see Figure 4.1.

The expectation and variance of the distribution is given by

$$E(X) = \mu \; , \tag{4.2}$$

$$Var(X) = \mu_2 \sigma^2 = K \sum_{i=1}^{n} \binom{2}{j} \left( \frac{1}{2h} \right)^j \frac{\{2(j+1)!\}}{2^{j+1}(j+1)!} \sigma^2 \; . \tag{4.3}$$

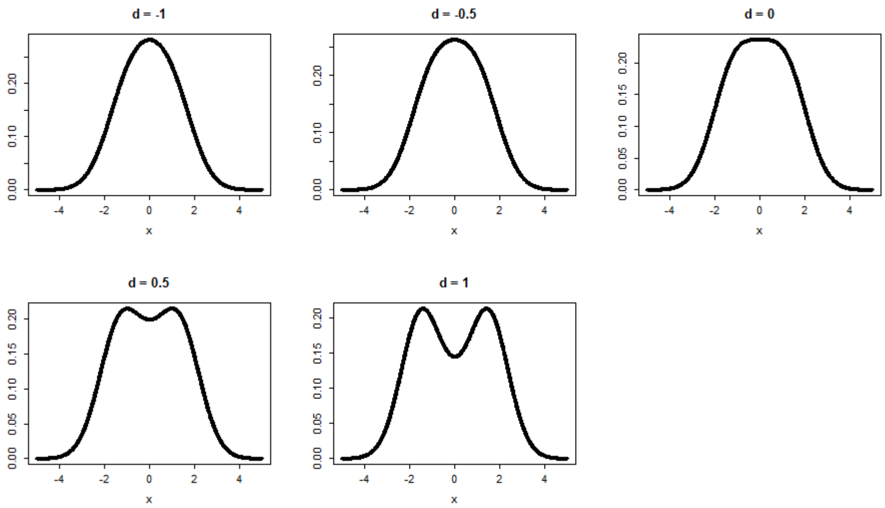


Figure 4.1: The density functions of short-tailed symmetric distribution with shape parameter $d = -1, -0.5, 0, 0.5, 1$, respectively.

## 4.1 Quality Control Charts with Simple Random Sampling

In the first part of the estimation process, parameter estimators are obtained under SRS. As in Section 3.1, maximum likelihood equations are intractable for short-tailed symmetric distribution too; therefore, MMLEs are obtained as given in Akkaya and Tiku [2]. MMLEs are as efficient as MLEs for small n and they are asymptotically fully efficient. The partial derivatives of log likelihood function are obtained as,

$$\frac{\partial lnL}{\partial \mu} = \frac{1}{\sigma} \sum_{i=1}^{n} z_i - \frac{2}{h\sigma} \sum_{i=1}^{n} g(z_i) \; , \tag{4.4}$$

$$\frac{\partial lnL}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma} \sum_{i=1}^{n} z_i^2 - \frac{2}{h\sigma} \sum_{i=1}^{n} z_i g(z_i) \; , \tag{4.5}$$

where

$$z_i = \frac{x_i - \mu}{\sigma} \; , \qquad g(z_i) = \frac{z_i}{1 + (1/2h) z_i^2} \; . \tag{4.6}$$

### 4.1.1 Modified Maximum Likelihood Estimation

Because of the non-linear function $g(z_i)$, closed form solutions for $\hat{\mu}$ and $\hat{\sigma}$ can not be obtained. To obtain MMLEs, $z_i$'s should be represented in the form of order statistics, $z_{(i)}$. After that, the non-linear function $g(z_{(i)})$ should be linearized by Taylor Series expansion around the expected value of $z_{(i)}$.

$$g(z_{(i)}) \cong \alpha_i + \gamma_i z_{(i)} \tag{4.7}$$

where

$$\alpha_i = \frac{\frac{1}{h}t_{(i)}^3}{\left\{1 + \frac{1}{2h}t_{(i)}^2\right\}^2} \quad \text{and} \quad \gamma_i = \frac{1 - \frac{1}{2h}t_{(i)}^2}{\left\{1 + \frac{1}{2h}t_{(i)}^2\right\}^2} . \tag{4.8}$$

Then, the estimators are obtained as follows:

$$\hat{\mu} = \frac{1}{m}\sum_{i=1}^{n} \beta_i x_{(i)} , \qquad \hat{\sigma} = \frac{-B + \sqrt{B^2 + 4nC}}{2\sqrt{n(n-1)}} \tag{4.9}$$

where

$$\beta_i = 1 - \frac{2}{h}\gamma_i , \qquad m = \sum_{i=1}^{n} \beta_i , \tag{4.10}$$

$$B = \frac{2}{h}\sum_{i=1}^{n} \alpha_i x_{(i)} \quad \text{and} \quad C = \sum_{i=1}^{n} \beta_i (x_{(i)} - \hat{\mu})^2 . \tag{4.11}$$

When $d \le 0$, $\beta_i$ $(1 \le i \le n)$ is always greater than zero and $C$ is positive which leads to real and positive $\hat{\sigma}$. On the other hand, some of the values of $\beta_i$ in the middle are less than zero for $d > 0$. In this case, $\hat{\sigma}$ is not real and positive. To obtain an estimator for $\sigma$ which is always real and positive, $\alpha_i^*$, $\gamma_i^*$, and $\beta_i^*$ can be used instead of $\alpha_i$, $\gamma_i$, and $\beta_i$. These modified coefficients are given by

$$\alpha_i^* = \frac{\frac{1}{h}t_{(i)}^3 + \left(1 - \frac{h}{2}\right) t_{(i)}}{\left\{1 + \frac{1}{2h}t_{(i)}^2\right\}^2} , \qquad \gamma_i^* = \frac{\frac{h}{2} - \frac{1}{2h}t_{(i)}^2}{\left\{1 + \frac{1}{2h}t_{(i)}^2\right\}^2} \quad \text{and} \quad \beta_i^* = 1 - \frac{2}{h}\gamma_i^* . \tag{4.12}$$

With this replacement, the asymptotic properties of the estimators do not change as

$$\alpha_i + \gamma_i z_{(i)} \cong \alpha_i^* + \gamma_i^* z_{(i)} \tag{4.13}$$

It should be noted that $\alpha_i = \alpha_i^*$ and $\gamma_i = \gamma_i^*$ when $d = 0$.

**Distribution of $\hat{\mu}$:** The distribution of $\hat{\mu}$ is important to determine how to construct confidence limits. If it is normally distributed, then classical Shewhart control chart method can be used. Otherwise, another approach is needed to construct chart limits. For this purpose, checking sample skewness and kurtosis values might be useful. Table 4.1 gives skewness ($\sqrt{\beta_1}$) and kurtosis ($\beta_2$) values of $\hat{\mu}$ obtained by Monte Carlo simulation based on 100,000 iterations.

Table 4.1: Skewness and kurtosis values of $\hat{\mu}$ under SRS

| d | n | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| -1 | $\sqrt{\beta_1}$ | -0.006 | 0.009 | -0.002 | 0.001 | -0.004 | 0.005 | -0.007 | 0.006 |
| | $\beta_2$ | 2.932 | 2.940 | 2.954 | 2.957 | 3.002 | 3.000 | 2.983 | 2.984 |
| -0.5 | $\sqrt{\beta_1}$ | 0.011 | 0.003 | -0.007 | 0.001 | -0.009 | 0 | -0.008 | 0.002 |
| | $\beta_2$ | 2.895 | 2.957 | 2.984 | 2.985 | 2.984 | 2.990 | 3.018 | 3.008 |
| 0 | $\sqrt{\beta_1}$ | -0.011 | -0.001 | 0.012 | 0.009 | -0.006 | -0.004 | 0.002 | 0.005 |
| | $\beta_2$ | 2.902 | 3.006 | 3.023 | 3.049 | 3.035 | 3.055 | 3.067 | 3.045 |
| 0.5 | $\sqrt{\beta_1}$ | -0.003 | 0.008 | -0.009 | 0.009 | 0.004 | -0.005 | -0.013 | 0 |
| | $\beta_2$ | 2.846 | 3.091 | 3.162 | 3.167 | 3.165 | 3.145 | 3.188 | 3.187 |
| 1 | $\sqrt{\beta_1}$ | -0.002 | 0.012 | -0.013 | 0.004 | 0.011 | -0.005 | 0.012 | 0.006 |
| | $\beta_2$ | 2.482 | 3.232 | 3.218 | 3.324 | 3.308 | 3.338 | 3.373 | 3.334 |

According to Table 4.1, the distribution of $\hat{\mu}$ seems to be quite close to normal for $d = -1, -0.5, 0$ since skewness values are very close to zero and kurtosis values are almost equal to 3. However, the distribution of $\hat{\mu}$ seems to be coming from a leptokurtic symmetric family for $d = 0.5, 1$. Therefore, a classical Shewhart control chart procedure can be applied when $d = -1, -0.5, 0$. For $d = 0.5, 1$, a three-moment t approximation provides more accurate results.

**Quality control limits for Shewhart control chart under SRS with normal approximation:** The MMLE of $\mu$ is unbiased since $E(\hat{\mu}) = \mu$. The exact variance of $\hat{\mu}$ is

$$Var(\hat{\mu}) = (\beta'\Omega\beta)\frac{\sigma^2}{m} \tag{4.14}$$

where $m$ is defined in Equation 4.10. Here $\beta$ is the column vector of $\beta_i$ coefficients in Equation 4.10 and $\Omega$ is the variance-covariance matrix of $Z_{(i)}$. But it is very hard to obtain $\Omega$; therefore, Akkaya and Tiku [2] obtained asymptotic variance of $\hat{\mu}$ as in Equation 4.15. It is given by

$$Var(\hat{\mu}) = \frac{\sigma^2}{nD} \tag{4.15}$$

where

$$D = 1 - \frac{2}{h}\left[\frac{1-\frac{1}{2h}}{1+2\frac{1}{2h}+3\frac{1}{2h}^2}\right]. \tag{4.16}$$

The Shewhart control limits for $\hat{\mu}$ satisfying normality assumption can be constructed as

$$
\begin{aligned}
UCL &= \hat{\mu} + 3\frac{c_1\bar{\hat{\sigma}}}{d_1\sqrt{nD}} \ , \\
CL &= \hat{\mu} \ , \\
LCL &= \hat{\mu} - 3\frac{c_1\bar{\hat{\sigma}}}{d_1\sqrt{nD}} \ ,
\end{aligned}
\tag{4.17}
$$

where $c_1$ is the constant to make $\hat{\sigma}$ an unbiased estimator of $\sigma$, and is obtained through simulations. The constant $d_1$ is also used to let the asymptotic variance of $\hat{\mu}$ be an unbiased estimator of exact variance and is also obtained through simulations.

## 4.2 Quality Control Charts with Ranked Set Sampling

In this section, parameters of short-tailed symmetric distribution are estimated under the RSS scheme. Each observation in RSS is coming from a different distribution and they are independent of each other since each order statistic is obtained from a different set of observations. Short-tailed symmetric distribution is also from a location-scale family; therefore, the likelihood can be obtained as in Equation 2.7 in Section 2.2. The only difference is,

$$
f(z) = \frac{K}{\sigma\sqrt{2\pi}}\left\{1 + \frac{1}{2h}z^2\right\}^2 \exp\left(-z^2/2\right) \ ; \quad -\infty < z < \infty \tag{4.18}
$$

where $z = (x-\mu)/\sigma$.

The partial derivatives of the log likelihood function are

$$
\begin{aligned}
\frac{\partial lnL}{\partial\mu} &= \left(\frac{1}{\sigma}\right)\left\{\sum_{i=1}^{n} z_{(i)} - \frac{2}{h}\sum_{i=1}^{n} g_1(z_{(i)}) - \sum_{i=1}^{n}(i-1)g_2(z_{(i)})\right. \\
&\quad \left. + \sum_{i=1}^{n}(n-i)g_3(z_{(i)})\right\}
\end{aligned}
\tag{4.19}
$$

$$
\begin{aligned}
\frac{\partial lnL}{\partial\sigma} &= \left(\frac{1}{\sigma}\right)\left\{\sum_{i=1}^{n} z_{(i)}^2 - \frac{2}{h}\sum_{i=1}^{n} z_{(i)}g_1(z_{(i)}) - \sum_{i=1}^{n}(i-1)z_{(i)}g_2(z_{(i)})\right. \\
&\quad \left. + \sum_{i=1}^{n}(n-i)z_{(i)}g_3(z_{(i)}) - n\right\}
\end{aligned}
\tag{4.20}
$$

where

$$g_1(z_{(i)}) = \frac{z_{(i)}}{1+(1/2h)z_{(i)}^2} \; , \tag{4.21}$$

$$g_2(z_{(i)}) = \frac{f(z_{(i)})}{F(z_{(i)})} \; , \tag{4.22}$$

$$g_3(z_{(i)}) = \frac{f(z_{(i)})}{1-F(z_{(i)})} \; ; \quad 1 \le i \le n \; . \tag{4.23}$$

The $g$ functions in the equations above are non-linear terms in the partial derivatives of the log likelihood function and prevent from obtaining explicit solutions of the maximum likelihood estimators.

### 4.2.1 Modified Maximum Likelihood Estimation

First of all, the non-linear $g$ functions should be linearized by Taylor Series expansion around the expected values of $Z_{(i)}$ as follows

$$g_1(z_{(i)}) \cong \alpha_{1i} + \beta_{1i} z_{(i)} \; , \tag{4.24}$$

$$g_2(z_{(i)}) \cong \alpha_{2i} - \beta_{2i} z_{(i)} \; , \tag{4.25}$$

$$g_3(z_{(i)}) \cong \alpha_{3i} + \beta_{3i} z_{(i)} \; ; \quad 1 \le i \le n \; . \tag{4.26}$$

where $\alpha_{1i}$ and $\beta_{1i}$ equal to $\alpha_i$ and $\gamma_i$ in Equation 4.8, respectively. Other coefficients are

$$\beta_{2i} = t_{(i)} \frac{f(t_{(i)})}{F(t_{(i)})} \left( 1 - \frac{4}{2h + t_{(i)}^2} \right) + \left( \frac{f(t_{(i)})}{F(t_{(i)})} \right)^2 \; , \tag{4.27}$$

$$\alpha_{2i} = \frac{f(t_{(i)})}{F(t_{(i)})} + t_{(i)} \beta_{2i} \; , \tag{4.28}$$

$$\beta_{3i} = \left( \frac{f(t_{(i)})}{1-F(t_{(i)})} \right)^2 - t_{(i)} \frac{f(t_{(i)})}{1-F(t_{(i)})} \left( 1 - \frac{4}{2h + t_{(i)}^2} \right) \; , \tag{4.29}$$

$$\alpha_{3i} = \frac{f(t_{(i)})}{1-F(t_{(i)})} - t_{(i)} \beta_{3i} \; ; \quad 1 \le i \le n \; . \tag{4.30}$$

When $d \le 0$, $\alpha_{1i}^*$ and $\beta_{1i}^*$ can be used instead of $\alpha_{1i}$ and $\beta_{1i}$ to obtain a real and positive estimator for $\sigma$. $\alpha_{1i}^*$ and $\beta_{1i}^*$ are equal to $\alpha_i^*$, $\gamma_i^*$ as in Equation 4.12, respectively. Then, MML estimators are obtained as

$$\hat{\mu} = \sum_{i=1}^{n} a_i x_{(i)} \; ; \qquad a_i = \frac{u_i}{\sum_{i=1}^{n} u_i} \tag{4.31}$$

and

$$\hat{\sigma} = \frac{-B + \sqrt{B^2 + 4nC}}{2n} \; , \tag{4.32}$$

where

$$B = \sum_{i=1}^{n} x_{(i)} w_i \ , \qquad C = \sum_{i=1}^{n} \left(x_{(i)} - \hat{\mu}\right)^2 u_i \ , \tag{4.33}$$

$$u_i = (i-1)\beta_{2i} + (n-i)\beta_{3i} + 1 - \frac{2}{h}\beta_{1i} \quad \text{and} \qquad w_i = \frac{2}{h}\alpha_{1i} + (i-1)\alpha_{2i} - (n-i)\alpha_{3i} \ . \tag{4.34}$$

It should also be noted that the MML estimators are asymptotically fully efficient and unbiased, and almost equal to ML estimators even for small sample sizes (Tiku and Akkaya [95]).

**Distribution of $\hat{\mu}$:** To simply analyze the distribution of $\hat{\mu}$, sample skewness and kurtosis values are examined based on a simulation of 100,000 runs. Table 4.2 gives the skewness ($\sqrt{\beta_1}$) and kurtosis ($\beta_2$) values of $\hat{\mu}$ under different $d$ parameters and different sample sizes $n = 3, 4, ..., 10$.

Table 4.2: Skewness and kurtosis values of $\hat{\mu}$ under RSS

| d | n | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| -1 | $\sqrt{\beta_1}$ | 0.032 | 0.011 | 0.051 | -0.011 | 0.004 | -0.049 | 0.005 | -0.020 |
| | $\beta_2$ | 2.994 | 2.944 | 3.004 | 2.941 | 3.019 | 2.992 | 2.992 | 3.060 |
| -0.5 | $\sqrt{\beta_1}$ | 0.014 | -0.006 | 0.014 | 0.018 | 0.037 | -0.001 | 0.009 | -0.012 |
| | $\beta_2$ | 2.993 | 2.926 | 3.048 | 3.061 | 3.021 | 2.939 | 2.967 | 3.007 |
| 0 | $\sqrt{\beta_1}$ | 0.015 | -0.008 | 0.011 | 0.045 | -0.014 | -0.018 | 0.023 | -0.020 |
| | $\beta_2$ | 2.926 | 2.995 | 3.032 | 3.075 | 3.052 | 3.023 | 3.015 | 3.016 |
| 0.5 | $\sqrt{\beta_1}$ | -0.009 | -0.002 | 0.003 | -0.024 | 0.006 | -0.007 | 0.055 | 0.002 |
| | $\beta_2$ | 3.013 | 3.113 | 3.070 | 3.088 | 3.017 | 3.023 | 3.094 | 3.006 |
| 1 | $\sqrt{\beta_1}$ | -0.076 | -0.006 | 0.002 | -0.011 | -0.008 | -0.015 | 0.041 | 0.044 |
| | $\beta_2$ | 2.992 | 3.294 | 3.151 | 3.215 | 3.219 | 3.043 | 3.183 | 3.080 |

The outcomes shown in Table 4.2 indicate that $\hat{\mu}$'s from the short-tailed symmetric distribution are approximately normally distributed when $d = -1, -0.5, 0$. On the other hand, $\hat{\mu}$'s from short-tailed symmetric distribution with $d = 0.5, 1$ do not satisfy normality since the kurtosis values are greater than 3. Therefore, classical Shewhart control chart can be constructed by using normal approximation for $d = -1, -0.5, 0$. As for $d = 0.5, 1$, a three-moment t approximation is again useful to construct control charts.

**Quality control limits for Shewhart control chart under RSS with normal approximation:** MML estimators are unbiased since $E(\hat{\mu}) = \mu$. The variance of the estimator $\hat{\mu}$ is

$$Var(\hat{\mu}_{rss}) = \frac{\sigma^2 \sum_{i=1}^{n} u_i^2 Var(z_{(i)})}{\sum_{i=1}^{n} u_i^{\ 2}} \tag{4.35}$$

where $u_i$ is defined in Equation 4.34.

For $\hat{\mu}$ satisfying normality assumption (when $d = -1, -0.5, 0$) the Shewhart control limits with type I error $\alpha = 0.0027$ are constructed as

$$\begin{aligned} UCL &= \bar{\hat{\mu}} + 3c_1\bar{\hat{\sigma}} \frac{\sqrt{\sum_{i=1}^{n} u_i^2 Var(z_{(i)})}}{\sum_{i=1}^{n} u_i}, \\ CL &= \bar{\hat{\mu}}, \\ LCL &= \bar{\hat{\mu}} - 3c_1\bar{\hat{\sigma}} \frac{\sqrt{\sum_{i=1}^{n} u_i^2 Var(z_{(i)})}}{\sum_{i=1}^{n} u_i}, \end{aligned} \tag{4.36}$$

where $c_1$ is a constant defined for unbiasedness purposes.

## 4.3 Simulation Results

In this section, efficiencies of the estimators under SRS and RSS are compared by Monte Carlo simulations based on 100,000 repetitions. In the simulations, short-tailed symmetric distribution with parameters $\mu = 0$, $\sigma = 1$, and 4 different $d$ parameters ($d = -1, -0.5, 0, 0.5, 1$) are used.

Table 4.3: Comparison of RSS and SRS in the estimation of short-tailed symmetric distribution parameters with $d = -1$

| | $\hat{\mu}$ | | | | $\hat{\sigma}$ | | | |
|---|---|---|---|---|---|---|---|---|
| | **Bias** | | **MSE** | | **Bias** | | **MSE** | |
| **n** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** |
| 3 | -0.0027 | -0.0014 | 0.5677 | 0.2936 | -0.1761 | -0.2443 | 0.2022 | 0.1573 |
| 4 | -0.0026 | 0.0098 | 0.4223 | 0.1735 | -0.1225 | -0.1617 | 0.1384 | 0.0957 |
| 5 | -0.0043 | -0.0008 | 0.3350 | 0.1162 | -0.0921 | -0.1104 | 0.1040 | 0.0647 |
| 6 | 0.0015 | 0.0024 | 0.2788 | 0.0867 | -0.0751 | -0.0804 | 0.0829 | 0.0474 |
| 7 | -0.0022 | 0.0031 | 0.2387 | 0.0640 | -0.0632 | -0.0591 | 0.0691 | 0.0364 |
| 8 | -0.0013 | 0.0027 | 0.2089 | 0.0500 | -0.0532 | -0.0445 | 0.0595 | 0.0286 |
| 9 | -0.0005 | 0.0013 | 0.1857 | 0.0401 | -0.0477 | -0.0362 | 0.0520 | 0.0237 |
| 10 | 0.0007 | 0.0002 | 0.1659 | 0.0328 | -0.0421 | -0.0299 | 0.0459 | 0.0193 |

Table 4.4: Comparison of RSS and SRS in the estimation of short-tailed symmetric distribution parameters with $d = -0.5$

| | $\hat{\mu}$ | | | | $\hat{\sigma}$ | | | |
|---|---|---|---|---|---|---|---|---|
| | **Bias** | | **MSE** | | **Bias** | | **MSE** | |
| **n** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** |
| 3 | -0.0027 | -0.0092 | 0.6065 | 0.3242 | -0.1900 | -0.2648 | 0.1960 | 0.1626 |
| 4 | 0.0024 | -0.0009 | 0.4545 | 0.1986 | -0.1308 | -0.1787 | 0.1350 | 0.0964 |
| 5 | -0.0006 | 0.0030 | 0.3615 | 0.1300 | -0.1003 | -0.1207 | 0.1011 | 0.0639 |
| 6 | 0.0011 | 0.0031 | 0.2948 | 0.0920 | -0.0798 | -0.0930 | 0.0800 | 0.0473 |
| 7 | 0.0009 | 0.0069 | 0.2536 | 0.0698 | -0.0659 | -0.0725 | 0.0668 | 0.0355 |
| 8 | 0.0007 | -0.0043 | 0.2225 | 0.0538 | -0.0564 | -0.0539 | 0.0564 | 0.0279 |
| 9 | 0.0024 | -0.0021 | 0.1959 | 0.0433 | -0.0500 | -0.0442 | 0.0493 | 0.0230 |
| 10 | -0.0021 | -0.0005 | 0.1785 | 0.0357 | -0.0446 | -0.0374 | 0.0439 | 0.0191 |

Table 4.5: Comparison of RSS and SRS in the estimation of short-tailed symmetric distribution parameters with $d = 0$

| | $\hat{\mu}$ | | | | $\hat{\sigma}$ | | | |
|---|---|---|---|---|---|---|---|---|
| | **Bias** | | **MSE** | | **Bias** | | **MSE** | |
| **n** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** |
| 3 | -0.0001 | 0.0008 | 0.6813 | 0.3799 | -0.2147 | -0.3005 | 0.1939 | 0.1710 |
| 4 | -0.0013 | -0.0009 | 0.5003 | 0.2252 | -0.1416 | -0.1983 | 0.1294 | 0.0991 |
| 5 | 0.0001 | -0.0014 | 0.3932 | 0.1477 | -0.1063 | -0.1450 | 0.0962 | 0.0658 |
| 6 | 0.0038 | -0.0006 | 0.3253 | 0.1044 | -0.0848 | -0.1078 | 0.0760 | 0.0461 |
| 7 | -0.0026 | 0.0009 | 0.2739 | 0.0790 | -0.0695 | -0.0856 | 0.0626 | 0.0353 |
| 8 | -0.0020 | 0.0013 | 0.2378 | 0.0603 | -0.0604 | -0.0659 | 0.0530 | 0.0263 |
| 9 | -0.0020 | 0.0029 | 0.2104 | 0.0495 | -0.0517 | -0.0555 | 0.0460 | 0.0221 |
| 10 | 0.0003 | 0.0019 | 0.1898 | 0.0402 | -0.0460 | -0.0469 | 0.0409 | 0.0183 |

Table 4.6: Comparison of RSS and SRS in the estimation of short-tailed symmetric distribution parameters with $d = 0.5$

| | $\hat{\mu}$ | | | | $\hat{\sigma}$ | | | |
|---|---|---|---|---|---|---|---|---|
| | **Bias** | | **MSE** | | **Bias** | | **MSE** | |
| **n** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** |
| 3 | -0.0029 | -0.0125 | 0.8672 | 0.5764 | -0.1584 | -0.2748 | 0.1872 | 0.1640 |
| 4 | -0.0036 | -0.0130 | 0.5783 | 0.2865 | -0.0986 | -0.1837 | 0.1223 | 0.0926 |
| 5 | -0.0013 | -0.0038 | 0.4448 | 0.1869 | -0.0703 | -0.1327 | 0.0885 | 0.0589 |
| 6 | 0.0008 | -0.0051 | 0.3615 | 0.1266 | -0.0550 | -0.1010 | 0.0696 | 0.0424 |
| 7 | -0.0019 | 0.0026 | 0.2998 | 0.0936 | -0.0436 | -0.0811 | 0.0564 | 0.0311 |
| 8 | 0.0010 | 0.0019 | 0.2587 | 0.0725 | -0.0359 | -0.0690 | 0.0474 | 0.0252 |
| 9 | -0.0010 | 0.0020 | 0.2260 | 0.0575 | -0.0301 | -0.0575 | 0.0410 | 0.0199 |
| 10 | -0.0040 | -0.0045 | 0.2017 | 0.0471 | -0.0262 | -0.0489 | 0.0360 | 0.0167 |

As demonstrated in Tables 4.3, 4.4, 4.5, 4.6, and 4.7, the biases of the estimator $\hat{\mu}$ under both SRS and RSS are all negligible. MSE values of $\hat{\mu}$ obtained for RSS

Table 4.7: Comparison of RSS and SRS in the estimation of short-tailed symmetric distribution parameters with $d = 1$

| | $\hat{\mu}$ | | | | $\hat{\sigma}$ | | | |
|---|---|---|---|---|---|---|---|---|
| | **Bias** | | **MSE** | | **Bias** | | **MSE** | |
| n | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** |
| 3 | -0.0028 | 0.0075 | 1.7300 | 2.0250 | 0.0685 | -0.0456 | 0.2765 | 0.1594 |
| 4 | -0.0061 | 0.0104 | 0.6866 | 0.4958 | -0.0137 | -0.1208 | 0.1276 | 0.0744 |
| 5 | 0.0012 | 0.0086 | 0.5442 | 0.3215 | 0.0044 | -0.0992 | 0.0877 | 0.0488 |
| 6 | -0.0061 | -0.0005 | 0.4086 | 0.1934 | 0.0060 | -0.0815 | 0.0656 | 0.0353 |
| 7 | -0.0018 | -0.0003 | 0.3352 | 0.1391 | 0.0082 | -0.0683 | 0.0522 | 0.0265 |
| 8 | -0.0021 | 0.0040 | 0.2808 | 0.0981 | 0.0078 | -0.0605 | 0.0433 | 0.0210 |
| 9 | -0.0024 | -0.0014 | 0.2425 | 0.0777 | 0.0073 | -0.0550 | 0.0362 | 0.0175 |
| 10 | -0.0001 | 0.0023 | 0.2119 | 0.0604 | 0.0083 | -0.0469 | 0.0317 | 0.0143 |

are always smaller than those for SRS. Biases of $\hat{\sigma}$ under RSS and SRS are almost equal to each other for some values of $n$. For other $n$'s biases under SRS are slightly smaller. However, MSE under RSS is always less than SRS. To compare the efficiencies of these estimators more clearly, relative efficiencies (RE) are obtained as in Table 4.8. Relative efficiency is constructed as in the form $MSE_{srs}/MSE_{rss}$; therefore, the values of RE less than one indicates that the estimator under SRS is more efficient than the RSS estimator.

Table 4.8: Relative efficiency of $\hat{\mu}$ and $\hat{\sigma}$ under SRS and RSS

| d | n | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| -1 | $\hat{\mu}$ | 1.93 | 2.43 | 2.88 | 3.22 | 3.73 | 4.18 | 4.63 | 5.05 |
| | $\hat{\sigma}$ | 1.29 | 1.45 | 1.61 | 1.75 | 1.90 | 2.08 | 2.19 | 2.38 |
| -0.5 | $\hat{\mu}$ | 1.87 | 2.29 | 2.78 | 3.20 | 3.63 | 4.14 | 4.52 | 5.00 |
| | $\hat{\sigma}$ | 1.21 | 1.40 | 1.58 | 1.69 | 1.88 | 2.03 | 2.14 | 2.30 |
| 0 | $\hat{\mu}$ | 1.79 | 2.22 | 2.66 | 3.11 | 3.47 | 3.94 | 4.25 | 4.72 |
| | $\hat{\sigma}$ | 1.13 | 1.31 | 1.46 | 1.65 | 1.77 | 2.02 | 2.08 | 2.23 |
| 0.5 | $\hat{\mu}$ | 1.50 | 2.02 | 2.38 | 2.86 | 3.20 | 3.57 | 3.93 | 4.28 |
| | $\hat{\sigma}$ | 1.14 | 1.32 | 1.50 | 1.64 | 1.81 | 1.88 | 2.06 | 2.15 |
| 1 | $\hat{\mu}$ | 0.85 | 1.38 | 1.69 | 2.11 | 2.41 | 2.86 | 3.12 | 3.51 |
| | $\hat{\sigma}$ | 1.74 | 1.71 | 1.80 | 1.86 | 1.97 | 2.06 | 2.07 | 2.21 |

According to the relative efficiency table (Table 4.8), it can be said that the estimators under RSS are highly efficient compared to those under SRS since RE values are all greater than 1. Moreover, as the sample size increases, the efficiency also increases.

In the second part of the simulations, type I errors of Shewhart control charts are obtained under short-tailed symmetric distribution with $d = -1, -0.5, 0, 0.5, 1$ and for both SRS and RSS. In the simulations, the true parameter values of $\mu$ and $\sigma$ are chosen to be 0 and 1, respectively, and the control limits are set for $\alpha = 0.0027$. As mentioned in the previous sections, the process mean does not satisfy normality for some shape parameters (which are $d = 0.5, 1$) of short-tailed symmetric distribution. Although they do not satisfy normality, we construct Tables 4.12 and 4.13 to see the effect of violation for the assumption on the type I error.

Table 4.9: Type I error comparison for Shewhart control charts under SRS and RSS for short-tailed symmetric distribution with $d = -1$

| | $M = 20$ | | $M = 30$ | | $M = 50$ | | $M = 100$ | |
|---|---|---|---|---|---|---|---|---|
| **n** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** |
| 3 | 0.0045 | 0.0044 | 0.0038 | 0.0036 | 0.0028 | 0.0031 | 0.0027 | 0.0027 |
| 4 | 0.0044 | 0.0037 | 0.0033 | 0.0032 | 0.0030 | 0.0028 | 0.0027 | 0.0025 |
| 5 | 0.0039 | 0.0038 | 0.0033 | 0.0033 | 0.0033 | 0.0030 | 0.0026 | 0.0027 |
| 6 | 0.0037 | 0.0037 | 0.0036 | 0.0033 | 0.0032 | 0.0030 | 0.0031 | 0.0028 |
| 7 | 0.0040 | 0.0033 | 0.0037 | 0.0029 | 0.0028 | 0.0027 | 0.0027 | 0.0025 |
| 8 | 0.0041 | 0.0034 | 0.0033 | 0.0031 | 0.0031 | 0.0029 | 0.0027 | 0.0027 |
| 9 | 0.0039 | 0.0037 | 0.0034 | 0.0034 | 0.0028 | 0.0031 | 0.0029 | 0.0030 |
| 10 | 0.0039 | 0.0034 | 0.0032 | 0.0031 | 0.0032 | 0.0028 | 0.0026 | 0.0026 |

Table 4.10: Type I error comparison for Shewhart control charts under SRS and RSS for short-tailed symmetric distribution with $d = -0.5$

| | $M = 20$ | | $M = 30$ | | $M = 50$ | | $M = 100$ | |
|---|---|---|---|---|---|---|---|---|
| **n** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** |
| 3 | 0.0048 | 0.0041 | 0.0036 | 0.0034 | 0.0028 | 0.0029 | 0.0024 | 0.0025 |
| 4 | 0.0041 | 0.0042 | 0.0035 | 0.0036 | 0.0032 | 0.0032 | 0.0024 | 0.0029 |
| 5 | 0.0040 | 0.0038 | 0.0033 | 0.0033 | 0.0028 | 0.0030 | 0.0029 | 0.0028 |
| 6 | 0.0040 | 0.0036 | 0.0035 | 0.0032 | 0.0030 | 0.0029 | 0.0026 | 0.0027 |
| 7 | 0.0039 | 0.0033 | 0.0035 | 0.0030 | 0.0032 | 0.0028 | 0.0032 | 0.0025 |
| 8 | 0.0038 | 0.0037 | 0.0034 | 0.0034 | 0.0031 | 0.0031 | 0.0028 | 0.0029 |
| 9 | 0.0038 | 0.0036 | 0.0032 | 0.0033 | 0.0033 | 0.0031 | 0.0027 | 0.0028 |
| 10 | 0.0038 | 0.0034 | 0.0036 | 0.0031 | 0.0030 | 0.0028 | 0.0029 | 0.0027 |

Tables 4.9, 4.10, and 4.11 show the performances of SRS and RSS estimators for short-tailed symmetric distribution with shape parameter $d = -1, -0.5, 0$, respectively. The process means obtained from these distributions satisfy normality assumption. Therefore, the type I errors are very close to 0.0027. Type I errors obtained under RSS are smaller for small sample sizes and sets with fewer samples.

Due to the fact that normality is not satisfied for the process means from short-tailed symmetric distributions with $d = 0.5, 1$ type I errors of Shewhart control charts are not as small as 0.0027, which necessitates using a moment approach to model the quantiles

Table 4.11: Type I error comparison for Shewhart control charts under SRS and RSS for short-tailed symmetric distribution with $d = 0$

| | $M = 20$ | | $M = 30$ | | $M = 50$ | | $M = 100$ | |
|---|---|---|---|---|---|---|---|---|
| **n** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** |
| 3 | 0.0044 | 0.0042 | 0.0032 | 0.0035 | 0.0025 | 0.0029 | 0.0023 | 0.0025 |
| 4 | 0.0049 | 0.0039 | 0.0035 | 0.0034 | 0.0033 | 0.0030 | 0.0027 | 0.0027 |
| 5 | 0.0040 | 0.0042 | 0.0035 | 0.0037 | 0.0033 | 0.0034 | 0.0032 | 0.0031 |
| 6 | 0.0042 | 0.0037 | 0.0038 | 0.0033 | 0.0038 | 0.0030 | 0.0031 | 0.0028 |
| 7 | 0.0045 | 0.0036 | 0.0037 | 0.0032 | 0.0035 | 0.0029 | 0.0032 | 0.0028 |
| 8 | 0.0041 | 0.0037 | 0.0038 | 0.0033 | 0.0033 | 0.0031 | 0.0033 | 0.0029 |
| 9 | 0.0039 | 0.0038 | 0.0033 | 0.0035 | 0.0034 | 0.0033 | 0.0034 | 0.0032 |
| 10 | 0.0042 | 0.0034 | 0.0037 | 0.0032 | 0.0034 | 0.0030 | 0.0031 | 0.0029 |

Table 4.12: Type I error comparison for Shewhart control charts under SRS and RSS for short-tailed symmetric distribution with $d = 0.5$

| | $M = 20$ | | $M = 30$ | | $M = 50$ | | $M = 100$ | |
|---|---|---|---|---|---|---|---|---|
| **n** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** |
| 3 | 0.0041 | 0.0043 | 0.0032 | 0.0036 | 0.0025 | 0.0030 | 0.0016 | 0.0026 |
| 4 | 0.0047 | 0.0046 | 0.0038 | 0.0041 | 0.0034 | 0.0037 | 0.0028 | 0.0034 |
| 5 | 0.0051 | 0.0044 | 0.0044 | 0.0040 | 0.0039 | 0.0037 | 0.0038 | 0.0035 |
| 6 | 0.0050 | 0.0039 | 0.0045 | 0.0036 | 0.0042 | 0.0033 | 0.0042 | 0.0031 |
| 7 | 0.0050 | 0.0038 | 0.0043 | 0.0035 | 0.0041 | 0.0032 | 0.0042 | 0.0030 |
| 8 | 0.0045 | 0.0039 | 0.0045 | 0.0036 | 0.0044 | 0.0034 | 0.0033 | 0.0032 |
| 9 | 0.0047 | 0.0034 | 0.0041 | 0.0031 | 0.0039 | 0.0029 | 0.0034 | 0.0027 |
| 10 | 0.0044 | 0.0037 | 0.0042 | 0.0034 | 0.0040 | 0.0032 | 0.0044 | 0.0031 |

Table 4.13: Type I error comparison for Shewhart control charts under SRS and RSS for short-tailed symmetric distribution with $d = 1$

| | $M = 20$ | | $M = 30$ | | $M = 50$ | | $M = 100$ | |
|---|---|---|---|---|---|---|---|---|
| **n** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** |
| 3 | 0.0023 | 0.0042 | 0.0015 | 0.0034 | 0.0011 | 0.0028 | 0.0008 | 0.0024 |
| 4 | 0.0053 | 0.0055 | 0.0048 | 0.0049 | 0.0041 | 0.0045 | 0.0040 | 0.0042 |
| 5 | 0.0050 | 0.0049 | 0.0047 | 0.0044 | 0.0042 | 0.0041 | 0.0039 | 0.0039 |
| 6 | 0.0059 | 0.0052 | 0.0059 | 0.0048 | 0.0056 | 0.0046 | 0.0052 | 0.0044 |
| 7 | 0.0062 | 0.0046 | 0.0066 | 0.0043 | 0.0054 | 0.0040 | 0.0053 | 0.0038 |
| 8 | 0.0056 | 0.0044 | 0.0059 | 0.0041 | 0.0052 | 0.0039 | 0.0052 | 0.0037 |
| 9 | 0.0059 | 0.0043 | 0.0055 | 0.0040 | 0.0049 | 0.0038 | 0.0048 | 0.0036 |
| 10 | 0.0058 | 0.0044 | 0.0054 | 0.0041 | 0.0052 | 0.0039 | 0.0050 | 0.0037 |

of the distribution of $\hat{\mu}$ under both distributions.

## 4.4 Three-Moment t Approximation

As mentioned in Section 3.4, three-moment t-approximation can be applicable to approximate the distributions of random variables having absolute skewness less than 0.1 and kurtosis greater than 3. The distribution of the parameter estimator $\hat{\mu}$ from short-tailed symmetric distribution with $d = 0.5, 1$ is quite relevant for this approximation. Since the distribution of $\hat{\mu}$ is not close to normal, classical Shewhart control chart method can not be applied. Instead, control charts for SRS and RSS are constructed by t-approximation. As before, write

$$\frac{\hat{\mu} + a}{b} \sim t_\nu \tag{4.37}$$

where the constants $a$, $b$, and $\nu$ are obtained by equating the first three moments on both sides of 4.37 (Tiku [91]).

**Quality Control Limits with $t$ Approximation under SRS:** Shewhart control chart limits under SRS by using three-moment t approximation are

$$\begin{aligned} UCL &= \bar{\hat{\mu}}_{srs} + t_{\nu,\alpha/2} b \ , \\ CL &= \bar{\hat{\mu}}_{srs} \ , \\ LCL &= \bar{\hat{\mu}}_{srs} - t_{\nu,\alpha/2} b \ , \end{aligned} \tag{4.38}$$

where

$$b = \sqrt{\frac{\nu - 2}{\nu} Var(\hat{\mu}_{srs})} \quad \text{and} \quad \nu = \frac{2(2\beta_2 - 3)}{\beta_2 - 3} \ . \tag{4.39}$$

$\beta_2$ is the simulated kurtosis of the distribution of $\hat{\mu}_{srs}$, and can be obtained through Monte Carlo simulations. Variance of $\hat{\mu}_{srs}$ is defined in the same form as in Equation 4.15.

**Quality Control Limits with $t$ Approximation under RSS:** Shewhart control chart limits with three-moment t-approximation can be obtained as follows under RSS:

$$\begin{aligned} UCL &= \bar{\hat{\mu}}_{rss} + t_{\nu,\alpha/2} b \ , \\ CL &= \bar{\hat{\mu}}_{rss} \ , \\ LCL &= \bar{\hat{\mu}}_{rss} - t_{\nu,\alpha/2} b \ , \end{aligned} \tag{4.40}$$

where

$$b = \sqrt{\frac{\nu - 2}{\nu} Var(\hat{\mu}_{rss})} \ , \qquad \nu = \frac{4(\beta_2 - 1.5)}{\beta_2 - 3} \ ; \tag{4.41}$$

$\beta_2$ being the simulated kurtosis of the distribution of $\hat{\mu}_{rss}$. Variance of $\hat{\mu}_{rss}$ is similarly obtained as in Equation 4.35.

### 4.4.1 Simulations with Three-Moment t Approximation

In this part, Monte Carlo simulations for Shewhart control charts are conducted by three-moment t approximation for the process means. The parameters of the distribution are set as $\mu = 0$ and $\sigma = 1$ and the $\alpha$ value is chosen as 0.0027 to construct control limits.

Table 4.14: Type I error comparison for Shewhart control charts under SRS and RSS for short-tailed symmetric distribution with $d = 0.5$

| | $M = 20$ | | $M = 30$ | | $M = 50$ | | $M = 100$ | |
|---|---|---|---|---|---|---|---|---|
| **n** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** |
| 4 | 0.0039 | 0.0033 | 0.0033 | 0.0030 | 0.0028 | 0.0027 | 0.0025 | 0.0025 |
| 5 | 0.0037 | 0.0040 | 0.0033 | 0.0036 | 0.0029 | 0.0034 | 0.0026 | 0.0032 |
| 6 | 0.0038 | 0.0032 | 0.0033 | 0.0029 | 0.0029 | 0.0027 | 0.0028 | 0.0025 |
| 7 | 0.0036 | 0.0036 | 0.0033 | 0.0034 | 0.0030 | 0.0030 | 0.0029 | 0.0029 |
| 8 | 0.0037 | 0.0038 | 0.0034 | 0.0035 | 0.0031 | 0.0032 | 0.0030 | 0.0031 |
| 9 | 0.0032 | 0.0027 | 0.0029 | 0.0025 | 0.0028 | 0.0023 | 0.0025 | 0.0021 |
| 10 | 0.0034 | 0.0036 | 0.0031 | 0.0033 | 0.0028 | 0.0031 | 0.0027 | 0.0028 |

Table 4.15: Type I error comparison for Shewhart control charts under SRS and RSS for short-tailed symmetric distribution with $d = 1$

| | $M = 20$ | | $M = 30$ | | $M = 50$ | | $M = 100$ | |
|---|---|---|---|---|---|---|---|---|
| **n** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** | **SRS** | **RSS** |
| 4 | 0.0036 | 0.0030 | 0.0029 | 0.0026 | 0.0025 | 0.0023 | 0.0023 | 0.0021 |
| 5 | 0.0036 | 0.0035 | 0.0031 | 0.0031 | 0.0029 | 0.0029 | 0.0026 | 0.0026 |
| 6 | 0.0037 | 0.0034 | 0.0033 | 0.0033 | 0.0031 | 0.0031 | 0.0029 | 0.0030 |
| 7 | 0.0037 | 0.0030 | 0.0036 | 0.0027 | 0.0033 | 0.0024 | 0.0032 | 0.0023 |
| 8 | 0.0035 | 0.0040 | 0.0033 | 0.0038 | 0.0032 | 0.0035 | 0.0030 | 0.0034 |
| 9 | 0.0032 | 0.0030 | 0.0030 | 0.0027 | 0.0028 | 0.0025 | 0.0027 | 0.0024 |
| 10 | 0.0033 | 0.0038 | 0.0032 | 0.0034 | 0.0030 | 0.0033 | 0.0029 | 0.0028 |

Tables 4.14 and 4.15 show that by using the three-moment t approximation, the type I errors converge to the target value 0.0027. Moreover, type I errors obtained under RSS are mostly smaller than SRS. There seems to be a weird sequence between the type I error values for the same number of samples. Sometimes values get very small and do not monotonically decrease as the sample size increases. The reason for this pattern arises from the shape of the distributions which are bimodal.

# CHAPTER 5

# APPLICATIONS

## 5.1 A REAL LIFE APPLICATION OF SHEWHART CONTROL CHART: LONG TAILED SYMMETRIC DISTRIBUTION

In this section, the use of Shewhart control charts under the proposed sampling method RSS is shown by a real life data. The data is from the article by Tadikamalla and Popescu [87]. They obtained the data set from a company which produces electronic components for various different sectors, such as military, medical companies. The data that is going to be used in the application is about optical lenses, and the characteristic that is going to be assessed is the center thickness of these lenses. The data was collected when the process was under control; therefore, it is expected that the sample statistics obtained from this data should vary within the control limits.

The data was scaled by some constants as $X = a + bY$ to preserve the confidentiality of the company. This transformation did not effect the distribution of the data, it only changed the mean and variance. The writers obtained the skewness and the kurtosis of the data as $-0.24$ and 5.1, respectively. Figure 5.1 displays the histogram of the data.

According to sample skewness and kurtosis values, it can be said that this data may be a good fit for long-tailed symmetric distribution. In the estimation process of control chart, the shape parameter of long-tailed symmetric distribution is treated as known. Therefore, the shape parameter should be estimated beforehand. The shape parameter can be estimated along the same lines as explained by Surucu and Sazak [86], and Tiku and Akkaya [95]. By following these studies, $\hat{\mu}$ and $\hat{\sigma}$ were obtained under various values of the shape parameter $p$, and the set of parameters maximizing the log-likelihood function was chosen. Figure 5.2 shows the log-likelihood values for the corresponding $p$ values.

The vertical line in Figure 5.2 shows the position of the $p$ parameter which maximizes the log-likelihood and the estimate is $\hat{p} = 2.9$. Other parameter estimates are obtained as $\hat{\mu} = 60.16$ and $\hat{\sigma} = 2.29$. To visually check the goodness-of-fit of this data set under the long-tailed symmetric distribution with shape parameter $p = 2.9$, a Q-Q plot was

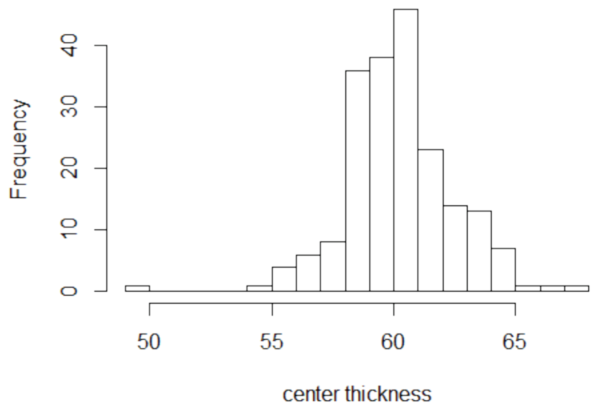


Figure 5.1: The histogram of the data, center thickness

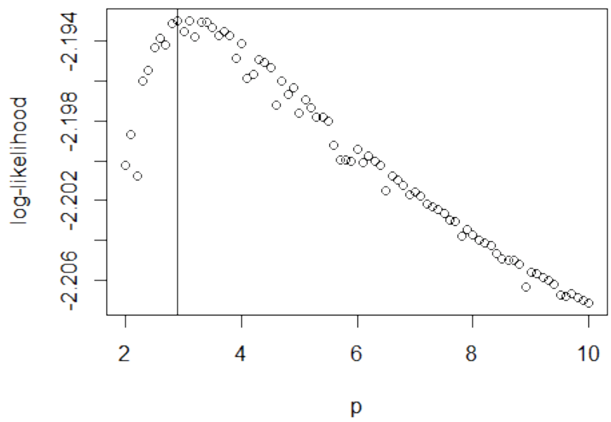


Figure 5.2: The plot of the estimation process

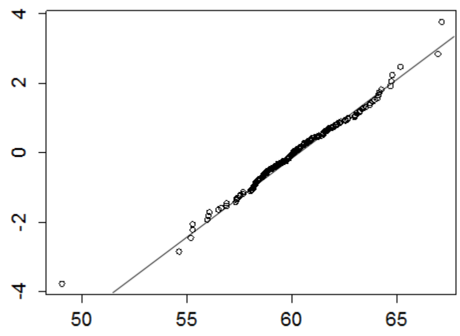


Figure 5.3: Q-Q plot of the center thickness data

constructed as in Figure 5.3.

The Q-Q plot in Figure 5.3 shows that the data exhibits a straight line structure. Therefore, it can be said that this data set is a good fit for the long-tailed symmetric distribution with shape parameter $p = 2.9$.

After determining the distribution from which the data comes from, Shewhart control charts were constructed under both the SRS and RSS methods. As explained in Chapter 3, the normality assumption for the sample means of long-tailed symmetric distribution is not acceptable. Therefore, quality control charts should be constructed with t-approximation. In this application, to show the appropriateness of the three-moment t approximation method, Shewhart control charts were constructed under both SRS and RSS, using three-moment t and normal approximations. To obtain equal sized SRS and RSS, sample sizes were chosen as $n = 3$ and 20 sub-groups were obtained. It can be noted that the total sample size was 200 in the original study.

The straight lines in Figures 5.4 and 5.5 show the upper and lower limits of Shewhart control chart obtained by normal approximation, whereas the dashed lines were obtained by t approximation. As mentioned above, the sample was collected when the process was under control, thus it is expected that all of the sample statistics should vary within the control limits. However the 3rd point in Figure 5.4 and the 17th point in Figure 5.5 are out-of Shewhart control limits obtained by normal approximation. The limits by three-moment approach cover all of the sample statistics in the plots which proves the accuracy of this method when compared to classical Shewhart control limits obtained with normal approximation.

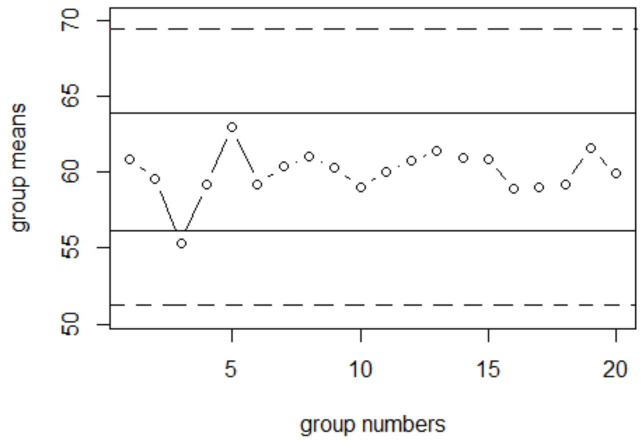


Figure 5.4: The Shewhart control chart under SRS with both normal and t approximations

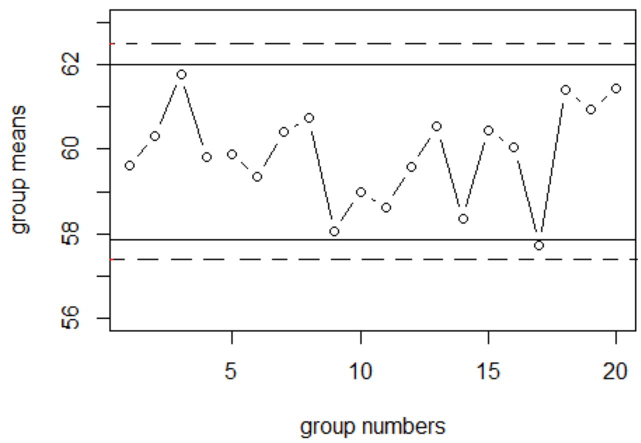


Figure 5.5: The Shewhart control chart under RSS with both normal and t approximations

## 5.2 AN EXAMPLE OF SHEWHART CONTROL CHART WITH SIMULATED DATA: SHORT TAILED SYMMETRIC DISTRIBUTION

In order to illustrate the concepts explained in this study, an example with a simulated data set is provided in this section. The example is prepared FOR short-tailed symmetric distribution with parameters $d = -1$, $\mu = 0$, and $\sigma = 1$. The simulated data sets and parameter estimators $\hat{\mu}$, $\hat{\sigma}$ are given in Tables B.1- B.3 in appendix. Sample size and sub-group size are chosen as $n = 5$ and $M = 50$, respectively.

Shewhart control charts constructed under both SRS and RSS are given in Figures 5.6-5.7. These samples can be treated as the samples from in-control process since all of them are generated from one distribution with fixed parameters. Therefore, it is expected that all sample statistics should vary within the control limits.

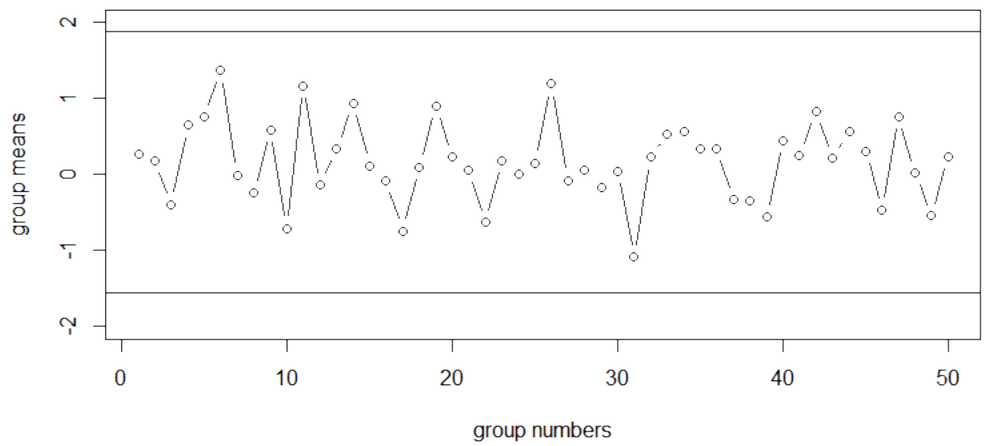


Figure 5.6: The Shewhart control chart under SRS for in-control situation

As expected, the sample points in Figures 5.6 and 5.7 are all within the control limits for both SRS and RSS. To compare the ability of detecting an out-of control process of charts under SRS and RSS, one observation was shifted according to Tiku's outlier model (Tiku [93]-[94]). In Tiku's outlier model, $3\sigma$ is added to the maximum value in the simulated data in order to generate an outlier.

Figures 5.8-5.9 show that Shewhart control chart under SRS could not detect the outlier whereas the outlier point generated is out-of control limits in the control chart under RSS. This example is a good proof of what is claimed in this study.

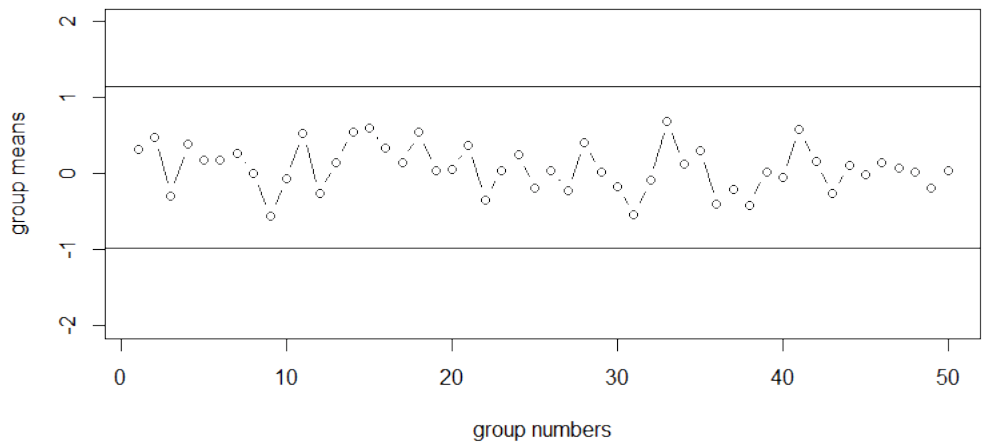


Figure 5.7: The Shewhart control chart under RSS for in-control situation

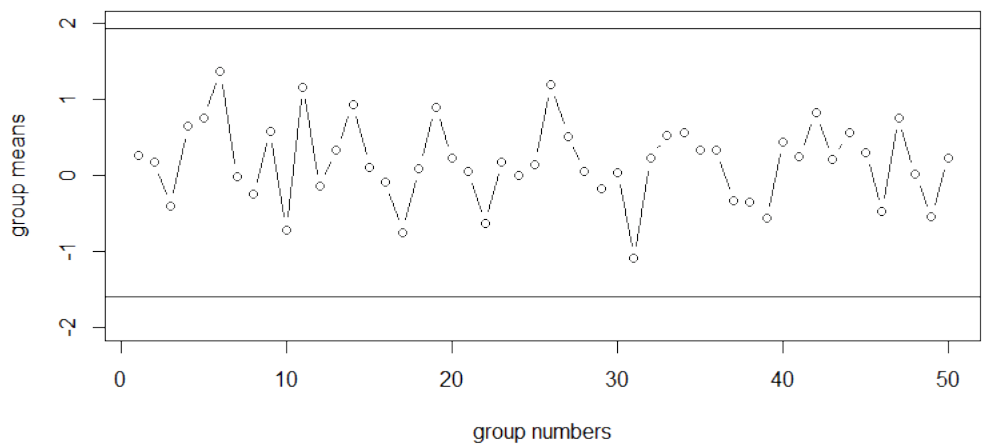


Figure 5.8: The Shewhart control chart under SRS for out-of control situation

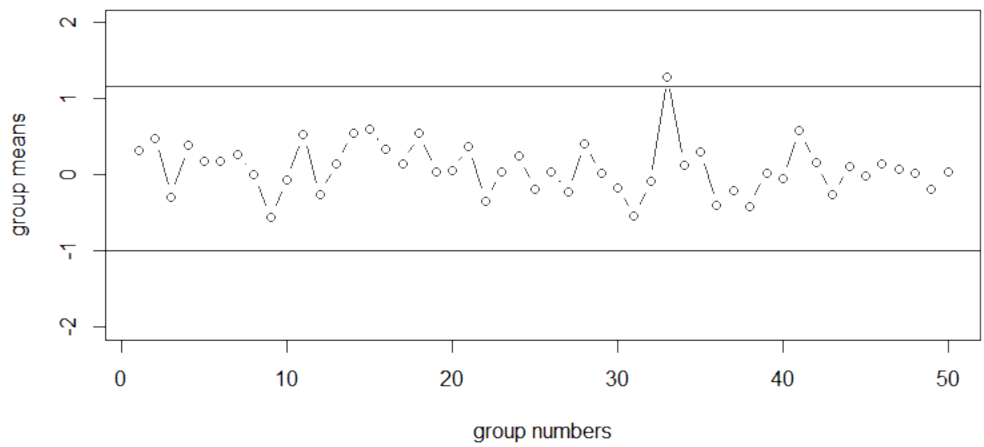


Figure 5.9: The Shewhart control chart under RSS for out-of control situation

# CHAPTER 6

# CONCLUSION AND FUTURE STUDIES

In this thesis, a new sampling method RSS has been adapted into the Shewhart control chart. The sampling method RSS holds more information than the SRS counterpart; therefore, especially for samples of small sizes, RSS provides samples which are better representatives of population of interest. Shewhart control chart method is one of the best application areas for this sampling method. In the literature, there are some papers applying RSS in Shewhart control charts (see Section 1.1.3); however, closed form solutions for parameter estimators are not given in these studies. In fact, due to the form of the likelihood function, the explicit solutions of MLEs can not be obtained under RSS. Therefore, in this study, an estimation method proposed by Tiku [92], and Tiku and Suresh [97] has been applied to obtain closed form solutions of parameter estimators under RSS and its modification ERSS for Shewhart control charts. The method is namely known as MMLE. It is asymptotically unbiased and fully efficient.

Shewhart control chart under RSS is compared with its SRS counterpart. In these comparisons, three different distributions are used. The first distribution is normal which is the most frequently used distribution in Shewhart control chart studies. The parameter estimators of normal under RSS are obtained by MMLE method and the Shewhart control charts obtained with this method yield smaller type I errors. Moreover, a modification of RSS, known as extreme ranked set sampling, is used to obtain quality control charts. The sampling design of ERSS differs for even and odd sized samples, hence the estimators have different forms for both even and odd sized samples. Type I errors of Shewhart control charts give better results under ERSS when compared to SRS.

The second distribution used in the comparison of control charts under different sampling designs is long tailed symmetric distribution. Since the variable of interest in this study is to control the process mean for constructing control charts, it can be said that the distribution of the process need not be normal, which follows from the CLT. However, this statement is true only for large samples. The efficiency of the sampling method RSS strictly depends on the correctness of the judgment ranks; therefore, sample sizes should be kept at minimum, such as 4, 5 or 6 to minimize the judgment ranking errors. However, this is not enough to ensure normality of statistics coming from long-tailed symmetric distribution. This necessitates using a different approach so as to construct control chart boundaries. Three-moment t approximation is very useful in this respect and provide the desired results.

Short-tailed symmetric distribution is the last distribution considered in this study. Unlike long-tailed symmetric distribution, some forms of this distribution is appropriate for normal approximation. On the other hand, some forms violate normality assumption, and three-moment t approximation is applicable. The simulation results show that Shewhart control charts under RSS result in smaller type I errors.

Simulations in this thesis are conducted by using R statistical software, version 2.15.1. Some sample codes are given in Appendix C.

In the estimation of quality control limits $\bar{\hat{\mu}}$ is used; however, depending on the sample size some robust methods, such as trimmed mean, median can be utilized as $\hat{\hat{\mu}}$. The performances of these control charts can be compared as a future study. Moreover, many different modifications of RSS can be compared under each distribution for estimating Shewhart control limits. The efficiency of the sampling method may vary according to the shape of the distribution; therefore, Shewhart control charts by SRS, RSS, and modifications of RSS can also be compared under skewed distributions. In this study, it is assumed that judgment rankings are perfect, however in practice, there may be judgment ranking errors. Therefore, RSS and its modifications can be compared under judgment ranking errors for constructing Shewhart control charts. Also, different sampling methods can be applied to the new types of control charts; such as CUSUM and EWMA charts as a future study.

# REFERENCES


[1] M. Abujiya and H. Muttlak. Quality control chart for the mean using double ranked set sampling. *Journal of Applied Statistics*, 31(10):1185–1201, 2004.

[2] A.D. Akkaya and M.L. Tiku. Short-tailed distributions and inliers. *Test*, 17:282–296, 2008.

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

[4] A.D. Al-Nasser, A.I. Al-Omari, and M. Al-Rawwash. Monitoring the process mean based on quality control charts using on folded ranked set sampling. *Pakistan Journal of Statistics and Operation Research*, 9(1):79–91, 2013.

[5] A.D. Al-Nasser and M. Al-Rawwash. A control chart based on ranked data. *Journal of Applied Sciences*, 7(14):1936–1941, 2007.

[6] A.D. Al-Nasser and A.B. Mustafa. Robust extreme ranked set sampling. *Journal of Statistical Computation and Simulation*, 79(7):859–867, 2009.

[7] A.I. Al-Omari and A.D. Al-Nasser. Statistical quality control limits for the sample mean chart using robust extreme ranked set sampling. *Economis Quality Control*, 27:73–89, 2011.

[8] A.I. Al-Omari and K.H. Jaber. Percentile double ranked set sampling. *Journal of Mathematics and Statistics*, 4(1):60–64, 2008.

[9] M.F. Al-Saleh and S.A. Al-Hadhrami. Estimation of the mean of the exponential distribution using moving extremes ranked set sampling. *Statistical Papers*, 44:367–382, 2003.

[10] M.F. Al-Saleh and S.A. Al-Hadrami. Parametric estimation for the location parameter for symmetric distributions using moving extremes ranked set sampling with application to trees data. *Environmetrics*, 14:651–664, 2013.

[11] M.F. Al-Saleh and M.A. Al-Kadiri. Double-ranked set sampling. *Statistics & Probability Letters*, 48:205–212, 2000.

[12] A. Atkinson and M. Riani. *Robust Diagnostic Regression Analysis*. Springer, New York, 2000.

[13] D.S. Bai and I.S. Choi. X-bar and r control charts for skewed populations. *Journal of Quality Technology*, 27(2):120–131, 1995.

[14] Z.D. Bai and Z. Chen. On the theory of ranked set sampling and its ramifications. *Journal of Statistical Planning and Inference*, 109:81–99, 2003.

[15] N. Balakrishnan and T. Li. Blues of parameters of generalized geometric distribution using ordered ranked set sampling. *Communications in Statistics-Simulation and Computation*, 34:253–266, 2005.

[16] N. Balakrishnan and T. Li. Confidence intervals for quantiles and tolerance intervals based on ordered ranked set samples. *Annals of the Institute of Statistical Mathematics*, 58(4):757–777, 2006.

[17] N. Balakrishnan and T. Li. Fisher information and tukey's linear sensitivity measure based on orderedranked set samples. In *Advances in Distribution Theory, Order Statistics, and Inference*, pages 187–206. Birkhauser Boston, 2006.

[18] N. Balakrishnan and T. Li. Ordered ranked set samples and applications to inference. *Journal of Statistical Planning and Inference*, 138:3512–3524, 2008.

[19] C.N. Bouza. Sampling using ranked sets: concepts, results and perspectives. *Revista Investigacion Operacional*, 26(3):1–19, 2005.

[20] R. Caulcutt. The rights and wrongs of control charts. *Applied Statistics*, 44(3):279–288, 1995.

[21] A. Cetinyurek. Robust control charts. Master's thesis, Middle East Technical University, 2006.

[22] S.P. Chakraborti, P. Van der Laan, and S.T. Bakir. Nonparametric control charts: an overview and some results. *Journal of Quality Technology*, 33:304–315, 2001.

[23] L.K. Chan and H.J. Cui. Skewness correction x-bar and r charts for skewed distributions. *Nav Res Logist*, pages 555–573, 2003.

[24] Z. Chen. On ranked-set sample quantiles and their applications. *Journal of Statistical Planning and Inference*, 83:125–135, 2000.

[25] Z. Chen. Ranked set sampling: its essence and some new applications. *Environmental and Ecological Statistics*, 14:355–363, 2007.

[26] Z. Chen and Z. Bai. The optimal ranked set sampling scheme for parametric families. *The Indian Journal of Statistics*, 62(2):178–192, 2000.

[27] Z. Chen, Z. Bai, and B.K. Sinha. *Ranked Set Sampling Theory and Applications*. Springer, 2003.

[28] W.K. Chiu. The economic design of cusum charts for controlling normal means. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 23(3):420–433, 1974.

[29] F. Choobineh and J.L. Ballard. Control-limits of qc charts for skewed distributions using weighted-variance. *IEEE Transactions on Reliability*, R-36(4):473–477, 1987.

[30] Y.M. Chou, A.M. Polansky, and R.L. Mason. Transforming non-normal data to normality in statistical process control. *Journal of Quality Technology*, 30(2):133–141, 1998.

[31] H. A. David and D.N. Levine. Ranked set sampling in the presence of judgment error. *Biometrics*, 28:553–555, 1972.

[32] T.R. Dell and J.L. Clutter. Ranked set sampling with order statistics backgroun. *Biometrics*, 28:545–553, 1972.

[33] A.J. Duncan. *Quality Control and Industrial Statistics*. Irwin Inc., 1965.

[34] E.A. Elsayed and A. Chen. An economic design of x-bar control chart using quadratic loss function. *International Journal of Production Research*, 32(4):873–887, 1994.

[35] W.D. Ewan. When and how to use cu-sum charts. *Technometrics*, 5(1):1–22, 1963.

[36] E.B. Ferrel. Control charts for lognormal universe. *Industrial Quality Control*, 15:4–6, 1958.

[37] M.A. Girshick and H. Rubin. A bayes approach to a quality control model. *The Annals of Mathematical Statistics*, 23:114–125, 1952.

[38] L.S. Halls and T.R. Dell. Trial of ranked set sampling for forage yields. *Forest Science*, 12(1):22–26, 1966.

[39] D.J. Hand, F. Daly, A.D. Lunn, K.J. McConway, and E. Ostrowski. *Small Data Sets*. Chapman & Hall, New York, 1994.

[40] H.L. Harter. Expected values of normal orderstatistics. *Biometrika*, 48:151–165, 1961.

[41] D.M. Hawkins. Multivariate quality control based on regression-adjusted variable. *Technometrics*, 33:61–75, 1991.

[42] D.M. Hawkins and D.H. Olwell. *Cumulative Sum Control Charts and Charting for Quality Improvement*. New York: Springer-Verlag, 1999.

[43] C. Ho and K. Case. Economic design of control charts: a literature review for 1981-1991. *Journal of Quality Technology*, 26:39–53, 1994.

[44] H. Hotelling. Multivariate quality control. In C. Eisenhart, M.W. Hastay, and W.A. Wallis, editors, *Techniques of Statistical Analysis*. New York: McGraw-Hill, 1947.

[45] International Conference on Networking and Information Technology. *Effect of non normality on statistical control charts*, 2010.

[46] M.Q. Islam, M.L. Tiku, and F. Yildirim. Non-normal regression: part i: skew distributions. *Communications in Statistics - Theory and Methods*, 30:993–1020, 2001.

[47] M.D. Kendal and A. Stuart. *The advanced theory of statistics*, volume 3. Charles Griffin, London, 1968.

[48] P.H. Kvam and F.J. Samaniego. On maximum likelihood estimation based on ranked set sampling with applications to reliability. In A. Basu, editor, *Advances in Reliability*, pages 215–229. 1993.

[49] R.Y. Liu. Control charts for multivariate processes. *Journal of the American Statistical Association*, 90:1380–1387, 1995.

[50] C.A. Lowry, W.H. Woodall, C.W. Champ, and S.E. Rigdon. A multivariate exponentially weighted moving average control chart. *Technometrics*, 34:46–53, 1992.

[51] C.W. Lu and Jr Reynolds, M.R. Ewma control charts for monitoring the mean of autocorrelated processes. *Journal of Quality Technology*, 31:166–188, 1999.

[52] J.M. Lucas and M.S. Saccucci. Exponentially weighted moving average control schemes: properties and enhancements. *Technometrics*, 32:1–12, 1990.

[53] S.N. MacEachern, O. Ozturk, D.A. Wolfe, and G.V. Stark. A new ranked set sample estimator of variance. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 64(2):177–188, 2002.

[54] R.L. Mason, C.W. Champ, N.D. Tracy, S.J. Wierda, and J.C. Young. Assessment of multivariate process control techniques. *Journal of Quality Technology*, 29:140–143, 1997.

[55] G.A. McIntyre. A method for unbiased selective sampling, using ranked sets. *Australian Journal of Agricultural Research*, 3(4):385 – 390, 1952.

[56] D.C. Montgomery. The economic design of control charts: A review and literature survey. *Journal of Quality Technology*, 12(2), 1980.

[57] D.C. Montgomery. *Introduction to Statistical Quality Control*. John Wiley & Sons, Inc., 6th edition, 2009.

[58] D.C. Montgomery and E.A. Peck. *Introduction to Linear Regression Analysis*. Wiley, New York, 1991.

[59] H. Muttlak and W. Al-Sabah. Statistical quality control based on ranked set sampling. *Journal of Applied Statistics*, 30(9):1055–1078, 2003.

[60] R.W. Nahhas, D.A. Wolfe, and H. Chen. Ranked set sampling: cost and optimal set size. *Biometrics*, 58(4):964–971, 2002.

[61] P.R. Nelson. Control charts for weibull processes with standards given. *IEEE Transactions on Reliability*, 28:283–387, 1979.

[62] K. Nishina and W. Peng-Hsiung. Performance of cusum charts from the viewpoint of change-point estimation in the presence of autocorrelation. *Quality and Reliability Engineering International*, 12:3–8, 1996.

[63] O. Ozturk. Ranked set sample inference under a symmetry restriction. *Journal of Statistical Planning and Inference*, 102:317–336, 2002.

[64] O. Ozturk. Robust joint estimation of location and scale parameters in ranked set samples. *Journal of Statistical Planning and Inference*, 127:295–308, 2005.

[65] O. Ozturk. Parametric estimation of location and scale parameters in ranked set sampling. *Journal of Statistical Planning and Inference*, 141:1616–1622, 2011.

[66] E.S. Page. Control chatrs for the mean of a normal population. *Journal of the Royal Statistical Society. Series B (Methodological)*, 16(1):131–135, 1954.

[67] G.P. Patil, D. Egemen, and B. Surucu. Ranked set sampling. In *Encyclopedia of Environmetrics*. John Wiley & Sons, 2nd edition, 2013.

[68] G.P. Patil, G.D. Johnson, D. Egemen, and B. Surucu. Ranked set sampling methodology for environmental data. In N. Balakrishnan, editor, *Methods and Applications of Statistics in the Atmospheric and Earth Sciences*. John Wiley & Sons, Inc., 2013.

[69] S. Puthenpura and N.K. Sinha. Modified maximum likelihood method for the robust estimation of system parametrs from very noisy data. *Automatica*, 22:231–235, 1986.

[70] C.P. Quesenberry. The effect of sample size on estimated limits for x-bar and x control charts. *Journal of Quality Technology*, 25(4):237–247, 1993.

[71] S.W. Roberts. Control charts based on geometric moving average. *Technometrics*, 1:239–250, 1959.

[72] T.P. Ryan. *Statistical Methods for Quality Improvement*. New York: Wiley, 1989.

[73] R.D. Salazar and A.K. Sinha. Control chart x bar based on ranked set sampling. Technical report, Department of Probability and Statistics, Centro de Investigation en Matematicas (CIMAT), 1997.

[74] H.M. Samawi, M.S. Ahmed, and W. Abu-Dayyeh. Estimating the population mean using extreme ranked set sampling. *Biometrical Journal*, 38(5):577–586, 1996.

[75] M.R. Shadid, M.Z. Raqab, and A.I. Al-Omari. Modified blues and blies of the location and scale parameters and the population mean using ranked set sampling. *Journal of Statistical Computation and Simulation*, 81(3):261–274, 2011.

[76] A.B. Shaibu and H.A. Muttlak. A comparison of the maximum likelihood estimators under ranked set sapling some of its modifications. *Applied Mathematics and Computation*, 129:441–453, 2002.

[77] A.B. Shaibu and H.A. Muttlak. Estimating the parameters of the normal, exponential and gamma distributions using median and extreme ranked set samples. *Statistica*, 64(1):75–98, 2004.

[78] W.A. Shewhart. *Economic Control of Quality of Manufactured Product*. D. Van Nostrand Company Inc. Princeton, New Jersey, 1931.

[79] A.N. Shiryaev. On optimum methods in quickest detection problem. *Theory of Probability and Applications*, 8:22–46, 1963.

[80] A.K. Sinha. On some recent developments in ranked set sampling. *Bulletin of Informatics and Cybernetics*, 37:137–160, 2005.

[81] T.A. Spedding and P.L. Rawlings. Non-normality in statistical process control measurements. *International Journal of Quality & Reliability Management*, 11(6):27–37, 1994.

[82] S.L. Stokes. Estimation of variance using judgment ordered ranked set samples. *Biometrics*, 36:35–42, 1980.

[83] S.L. Stokes. Inferences on the correlation coefficient in bivariate normal populations from ranked set samples. *Journal of the American Statistical Association*, 75:989–995, 1980.

[84] S.L. Stokes and T.W. Sager. Characterization of a ranked set sample with application to estimating distribution functions. *Journal of the American Statistical Association*, 83:374–381, 1988.

[85] Z.G. Stoumbos, JR. M.R. Reynolds, T.R. Ryan, and W.H. Woodall. Thestate of statistical process control as we proceed into the 21st century. *Journal of the American Statistical Association*, 95(451):992–998, 2000.

[86] B. Surucu and H.S. Sazak. Monitoring reliability for a three-parameter weibull distribution. *Reliability Engineering and System Safety*, 94:503–508, 2009.

[87] P.R. Tadikamalla and D.G. Popescu. Kurtosis correction method for x-bar and r control charts for long-tailed symmetrical distributions. *Naval Research Logistics*, 54(4), 2007.

[88] G. Tagaras. A survey of recent developments in the design of adaptive control charts. *Journal of Quality Technology*, 30:212–231, 1998.

[89] K. Takahasi and K. Wakimoto. 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, 1968.

[90] H.M. Taylor. The economic design of cumulative sum control charts. *Technometrics*, 10:479–488, 1968.

[91] M.L. Tiku. Usefulness of three-moment chi-square and t approximations. *Journal of Indian Society and Agricultural Statistics*, 18:4–16, 1966.

[92] M.L. Tiku. Estimating the mean and standard deviation from a censored normal sample. *Biometrika*, 54:155–165, 1967.

[93] M.L. Tiku. A new statistic for testing suspected outliers. *Communications in Statistics*, 4(8):737–752, 1975.

[94] M.L. Tiku. Rejoinder: Comment on "a new statistic for testing suspected outliers". *Communications in Statistics - Theory and Methods*, A6:1417–1422, 1977.

[95] M.L. Tiku and A.D. Akkaya. *Robust Estimation and Hypothesis Testing.* New Age International (P) Limited, Publishers, 2004.

[96] M.L. Tiku and S. Kumra. Expected values and variances and covariances of order statistics for a family of symmetric distributions (students t). *Mathematical Statistics, American Mathematical society: Providence, RI*, 8:141–270, 1981.

[97] M.L. Tiku and R.P. Suresh. A new method of estimation for location and scale parameters. *Journal of Statistical Planning and Inference*, 30:281–292, 1992.

[98] M.L. Tiku and D.C. Vaughan. A family of short-tailed symmetric distributions. Technical report, McMaster University, Canada, 1999.

[99] M.L. Tiku, W.K. Wong, D.C. Vaughan, and G. Bian. Time series models in non-normal situations: symmetric innovations. *Journal of Time Series Analysis*, 21:571–596, 2000.

[100] D.C. Vaughan. Expected values, variances and covariances of order statistics for students t distribution with two degrees of freedom. *Communications in Statistics - Simulation and Computation*, 21:391–404, 1992.

[101] D.C. Vaughan. On the tiku-suresh method of estimation. *Communications in Statistics - Theory and Methods*, 21:451–469, 1992.

[102] Y. Wang, Z. Chen, and J. Liu. General ranked set sampling with cost considerations. *Biometrics*, 60(2):556–561, 2004.

[103] D.A. Wolfe. Ranked set sampling: an approach to more efficient data collection. *Statistical Science*, 19(4):636–643, 2004.

[104] D.A. Wolfe. Ranked set sampling. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(4):460–466, 2010.

[105] Z. Wu. Asymmetric control limits of the x-bar chart for skewed process distributions. *International Journal of Quality & Reliability Management*, 13(9):49–60, 1996.

[106] S.A. Yourstone and W.J. Zimmer. Non-normality and thedesign of control charts for averages. *Decision Sciences*, 23(5):1099–1113, 1992.

# APPENDIX A

# BIAS OF THE SCALE PARAMETER ESTIMATOR FOR LONG TAILED SYMMETRIC DISTRIBUTION

Table A.1: Bias of the scale parameter estimator ($\hat{\sigma}_{mml}$) under RSS

| n \ $p$ | 3 | 3.5 | 5 | 10 |
|---|---|---|---|---|
| 3 | -0.0528 | -0.0779 | -0.1159 | -0.1564 |
| 4 | 0.0153 | -0.0032 | -0.0449 | -0.0799 |
| 5 | 0.0540 | 0.0386 | 0.0053 | -0.0372 |
| 6 | 0.0785 | 0.0523 | 0.0236 | -0.0176 |
| 7 | 0.0829 | 0.0690 | 0.0326 | -0.0015 |
| 8 | 0.0970 | 0.0782 | 0.0441 | 0.0123 |
| 9 | 0.1007 | 0.0869 | 0.0496 | 0.0208 |
| 10 | 0.1054 | 0.0855 | 0.0566 | 0.0292 |
| 11 | 0.1047 | 0.0837 | 0.0556 | 0.0256 |
| 12 | 0.1031 | 0.0900 | 0.0553 | 0.0316 |
| 13 | 0.0998 | 0.0833 | 0.0568 | 0.0311 |
| 14 | 0.0987 | 0.0809 | 0.0578 | 0.0326 |
| 15 | 0.0978 | 0.0792 | 0.0593 | 0.0332 |
| 16 | 0.0955 | 0.0792 | 0.0604 | 0.0338 |
| 17 | 0.0942 | 0.0799 | 0.0561 | 0.0342 |
| 18 | 0.0940 | 0.0776 | 0.0551 | 0.0345 |
| 19 | 0.0912 | 0.0769 | 0.0537 | 0.0348 |
| 20 | 0.0860 | 0.0726 | 0.0534 | 0.0347 |
| 30 | 0.0679 | 0.0591 | 0.0446 | 0.0321 |
| 50 | 0.0473 | 0.0420 | 0.0318 | 0.0230 |
| 100 | 0.0266 | 0.0236 | 0.0184 | 0.0130 |

# APPENDIX B

# SIMULATED DATA SETS FROM SHORT TAILED SYMMETRIC DISTRIBUTION UNDER SRS AND RSS

Table B.1: Simulated SRS sample with $n = 5$ and $M = 50$ from short-tailed symmetric distribution with parameters $d = -1$, $\mu = 0$, and $\sigma = 1$

| group | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $\hat{\mu}_{SRS}$ | $\hat{\sigma}_{SRS}$ |
|---|---|---|---|---|---|---|---|
| 1 | -0.619 | 0.387 | -0.754 | 0.176 | 2.170 | 0.375 | 0.871 |
| 2 | 1.036 | 1.862 | -1.865 | 1.710 | -1.836 | 0.100 | 1.297 |
| 3 | -1.490 | 1.952 | 0.331 | -1.023 | -1.728 | -0.248 | 1.118 |
| 4 | -0.045 | 0.352 | 1.728 | -0.441 | 1.661 | 0.663 | 0.701 |
| 5 | 1.892 | -0.757 | 1.812 | 1.882 | -1.043 | 0.633 | 1.037 |
| 6 | 1.686 | 2.720 | 1.424 | -0.022 | 1.045 | 1.363 | 0.754 |
| 7 | -0.381 | -1.361 | 0.152 | 1.112 | 0.448 | -0.040 | 0.696 |
| 8 | -0.208 | -0.803 | 0.292 | -0.952 | 0.438 | -0.251 | 0.444 |
| 9 | 0.505 | 1.392 | 1.212 | -1.172 | 1.028 | 0.464 | 0.775 |
| 10 | -2.148 | -2.257 | 1.277 | -0.722 | 0.296 | -0.660 | 1.102 |
| 11 | 0.612 | 1.597 | 0.714 | 1.480 | 1.429 | 1.140 | 0.324 |
| 12 | 0.074 | -0.020 | -0.361 | -0.423 | 0.058 | -0.149 | 0.167 |
| 13 | 0.646 | 1.370 | -0.851 | 0.952 | -0.398 | 0.311 | 0.674 |
| 14 | 2.863 | 1.248 | -0.706 | -0.062 | 1.381 | 0.961 | 1.027 |
| 15 | -0.887 | 1.162 | -0.227 | 1.182 | -0.632 | 0.142 | 0.691 |
| 16 | -0.057 | -2.625 | 0.120 | 1.040 | 1.122 | -0.241 | 1.134 |
| 17 | -0.728 | -2.495 | -0.651 | 0.652 | -0.469 | -0.784 | 0.863 |
| 18 | -1.186 | 1.838 | -0.218 | 1.189 | -1.125 | 0.166 | 0.970 |
| 19 | 0.116 | 2.250 | 1.607 | 0.510 | 0.040 | 0.977 | 0.699 |
| 20 | 2.470 | 0.066 | -0.411 | -0.787 | -0.192 | 0.387 | 0.972 |
| 21 | -1.104 | 0.792 | -0.403 | -0.074 | 1.086 | 0.050 | 0.650 |
| 22 | 0.057 | -1.183 | 0.311 | -0.281 | -2.062 | -0.703 | 0.713 |
| 23 | -1.645 | 0.543 | 0.975 | 1.240 | -0.183 | 0.082 | 0.855 |
| 24 | 0.393 | -1.287 | 0.048 | -0.193 | 1.057 | -0.025 | 0.651 |
| 25 | -0.040 | 1.448 | -1.841 | 0.051 | 1.154 | 0.080 | 0.963 |

Table B.2: Simulated SRS sample cont.

| group | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $\hat{\mu}_{SRS}$ | $\hat{\sigma}_{SRS}$ |
|---|---|---|---|---|---|---|---|
| 26 | 0.304 | 2.531 | -0.069 | 2.210 | 1.011 | 1.214 | 0.816 |
| 27 | 0.371 | -0.633 | -1.747 | -1.728 | 3.351 | 0.147 | 1.557 |
| 28 | 0.161 | 0.775 | 0.767 | -0.089 | -1.305 | -0.017 | 0.631 |
| 29 | -2.203 | -0.041 | 0.928 | -0.094 | 0.556 | -0.283 | 0.915 |
| 30 | 0.397 | 0.493 | 1.549 | -1.737 | -0.495 | -0.006 | 0.921 |
| 31 | -2.770 | -1.968 | -1.284 | 0.241 | 0.378 | -1.097 | 0.984 |
| 32 | 2.272 | -0.967 | 0.350 | -0.602 | 0.107 | 0.333 | 0.943 |
| 33 | 0.528 | 1.070 | 1.768 | -0.732 | 0.017 | 0.527 | 0.712 |
| 34 | 2.269 | -1.664 | 1.153 | 0.021 | 1.066 | 0.485 | 1.108 |
| 35 | -1.457 | 0.117 | 0.301 | 0.078 | 2.705 | 0.422 | 1.143 |
| 36 | -1.299 | 2.759 | 0.594 | -0.069 | -0.314 | 0.442 | 1.143 |
| 37 | 0.467 | 1.833 | -0.249 | -1.859 | -1.790 | -0.254 | 1.133 |
| 38 | -1.101 | 0.429 | -0.708 | -1.072 | 0.718 | -0.295 | 0.603 |
| 39 | -3.815 | 1.142 | -2.123 | 1.229 | 0.795 | -0.784 | 1.624 |
| 40 | -0.306 | 2.483 | 0.420 | -1.937 | 1.567 | 0.408 | 1.267 |
| 41 | 1.282 | 0.351 | -0.590 | -1.102 | 1.282 | 0.205 | 0.768 |
| 42 | -1.753 | -0.366 | 3.083 | 2.235 | 0.946 | 0.786 | 1.425 |
| 43 | 1.055 | -0.142 | 1.572 | 0.188 | -1.598 | 0.165 | 0.910 |
| 44 | 0.895 | 1.061 | 2.064 | 1.470 | -2.685 | 0.341 | 1.411 |
| 45 | 0.941 | 0.811 | -0.101 | -0.078 | -0.022 | 0.350 | 0.358 |
| 46 | 0.328 | -1.011 | -1.819 | -1.163 | 1.351 | -0.385 | 0.934 |
| 47 | -1.147 | 0.396 | 1.775 | 0.764 | 2.047 | 0.695 | 0.940 |
| 48 | 0.885 | -1.380 | 0.017 | 0.629 | -0.038 | -0.038 | 0.658 |
| 49 | 0.537 | -1.997 | 0.231 | -0.507 | -0.988 | -0.589 | 0.744 |
| 50 | -1.057 | 2.698 | -0.198 | -1.132 | 0.867 | 0.380 | 1.168 |

Table B.3: Simulated RSS sample with $n = 5$ and $M = 50$ from short-tailed symmetric distribution with parameters $d = -1$, $\mu = 0$, and $\sigma = 1$

| group | $x_{(1)}$ | $x_{(2)}$ | $x_{(3)}$ | $x_{(4)}$ | $x_{(5)}$ | $\hat{\mu}_{RSS}$ | $\hat{\sigma}_{RSS}$ |
|---|---|---|---|---|---|---|---|
| 1 | -1.798 | -0.963 | 1.002 | 1.318 | 2.091 | 0.330 | 1.006 |
| 2 | -0.134 | 0.747 | 0.126 | -0.045 | 1.686 | 0.471 | 0.579 |
| 3 | -2.624 | -0.042 | -0.381 | 0.223 | 1.406 | -0.275 | 0.933 |
| 4 | -1.170 | -0.574 | -0.532 | 1.255 | 2.946 | 0.378 | 1.056 |
| 5 | -0.775 | -1.134 | -0.057 | 1.665 | 1.237 | 0.188 | 0.808 |
| 6 | -0.502 | -1.681 | -0.930 | 0.914 | 3.153 | 0.170 | 1.278 |
| 7 | 0.057 | -1.449 | -0.040 | 0.304 | 2.508 | 0.253 | 1.004 |
| 8 | -1.918 | -0.311 | -0.224 | 0.528 | 1.967 | 0.010 | 0.861 |
| 9 | -2.350 | -1.615 | 0.022 | -0.417 | 1.626 | -0.555 | 0.954 |
| 10 | -1.753 | 0.263 | -0.322 | 1.282 | 0.176 | -0.051 | 0.801 |
| 11 | 0.669 | 0.203 | -0.608 | 0.992 | 1.389 | 0.524 | 0.646 |
| 12 | -2.090 | -1.057 | 0.387 | 1.137 | 0.352 | -0.243 | 0.853 |
| 13 | -1.910 | -0.803 | -0.020 | 2.244 | 1.248 | 0.165 | 1.060 |
| 14 | -0.522 | 1.090 | -0.287 | 0.141 | 2.360 | 0.553 | 0.873 |
| 15 | -0.423 | -0.522 | -0.044 | 1.182 | 2.820 | 0.592 | 0.893 |
| 16 | 0.530 | -0.041 | 0.493 | -0.533 | 1.244 | 0.324 | 0.654 |
| 17 | -1.284 | -0.787 | 0.290 | 1.354 | 1.198 | 0.158 | 0.731 |
| 18 | -1.132 | -0.193 | -0.142 | 2.668 | 1.530 | 0.560 | 1.019 |
| 19 | -2.331 | 0.017 | 1.615 | -0.094 | 1.030 | 0.054 | 1.081 |
| 20 | -1.380 | -1.054 | -0.732 | 0.521 | 2.968 | 0.051 | 1.119 |
| 21 | -1.865 | -0.766 | 0.191 | 2.516 | 1.825 | 0.392 | 1.143 |
| 22 | -1.859 | -1.330 | -0.004 | 0.292 | 1.212 | -0.340 | 0.758 |
| 23 | -2.523 | 0.351 | -0.133 | 1.470 | 1.056 | 0.065 | 1.033 |
| 24 | -1.701 | -1.099 | 0.764 | 0.629 | 2.659 | 0.241 | 1.058 |
| 25 | -2.377 | -0.488 | -0.411 | -0.403 | 2.768 | -0.190 | 1.178 |

Table B.4: Simulated RSS sample cont.

| group | $x_{(1)}$ | $x_{(2)}$ | $x_{(3)}$ | $x_{(4)}$ | $x_{(5)}$ | $\hat{\mu}_{RSS}$ | $\hat{\sigma}_{RSS}$ |
|---|---|---|---|---|---|---|---|
| 26 | -0.933 | 0.048 | 0.296 | -0.318 | 1.159 | 0.047 | 0.525 |
| 27 | -1.613 | -0.469 | -1.069 | 0.851 | 1.199 | -0.215 | 0.773 |
| 28 | -0.795 | -1.591 | 0.957 | 0.944 | 2.568 | 0.401 | 1.056 |
| 29 | -1.031 | -1.685 | -0.590 | 1.138 | 2.260 | 0.008 | 1.044 |
| 30 | -1.105 | -1.819 | 0.378 | 0.650 | 1.076 | -0.171 | 0.815 |
| 31 | -2.234 | -2.037 | -0.314 | 0.718 | 1.174 | -0.540 | 0.961 |
| 32 | -1.426 | -1.419 | -0.210 | 0.574 | 2.053 | -0.095 | 0.903 |
| 33 | -1.007 | -0.510 | 0.867 | 0.796 | 3.317 | 0.680 | 1.043 |
| 34 | -0.817 | -1.208 | 0.191 | 1.126 | 1.398 | 0.134 | 0.732 |
| 35 | -0.380 | -1.132 | 0.991 | 1.793 | 0.272 | 0.313 | 0.904 |
| 36 | -1.226 | -1.235 | -0.988 | -0.108 | 1.573 | -0.407 | 0.764 |
| 37 | -3.031 | 1.212 | -0.897 | -0.105 | 1.776 | -0.194 | 1.363 |
| 38 | -2.014 | -2.123 | -1.356 | 1.005 | 2.421 | -0.422 | 1.273 |
| 39 | -0.643 | -1.767 | 0.378 | 1.615 | 0.558 | 0.028 | 0.942 |
| 40 | -2.012 | -0.810 | 0.259 | 1.153 | 1.185 | -0.038 | 0.844 |
| 41 | -0.633 | -1.072 | 0.844 | 1.102 | 2.652 | 0.566 | 0.938 |
| 42 | -1.121 | 0.325 | 0.091 | 0.298 | 1.212 | 0.164 | 0.535 |
| 43 | -2.559 | -0.260 | 0.375 | -0.287 | 1.411 | -0.261 | 0.941 |
| 44 | -1.579 | -1.436 | 1.049 | 1.054 | 1.482 | 0.112 | 0.954 |
| 45 | -1.440 | -0.910 | -0.154 | 1.845 | 0.644 | 0.008 | 0.873 |
| 46 | -2.179 | -0.564 | -0.342 | 1.990 | 1.865 | 0.165 | 1.101 |
| 47 | -1.866 | -0.270 | 0.599 | -0.370 | 2.317 | 0.075 | 0.995 |
| 48 | -0.789 | -0.769 | -0.104 | 0.629 | 1.173 | 0.025 | 0.533 |
| 49 | -0.766 | -0.887 | -0.516 | -0.227 | 1.506 | -0.190 | 0.640 |
| 50 | -1.875 | -0.035 | -0.189 | 0.902 | 1.424 | 0.054 | 0.783 |

# APPENDIX C

# R CODES FOR SHEWHART CONTROL CHARTS UNDER NORMAL, LONG-TAILED SYMMETRIC AND SHORT-TAILED SYMMETRIC DISTRIBUTIONS

```
# QUALITY CONTROL CHART BY RANKED SET SAMPLING UNDER NORMAL DISTRIBUTION

cchart_rss_normal <- function(M,n,vec_n,vec2_n)
{
N = 10000
N_test = 10000

for(h in 1:N){
#######  GENERATING M SAMPLE OF SIZE n
      for(group in 1:M){
         for(i in 1:n)
             rss[i,] = rnorm(n, mean = 0, sd = 1)

         rss = t(apply(rss,1,sort))
         rss_samp= diag(rss)

         sum_urxr = 0
         sum_urxr_sq =0
    sum_wrxr = 0
         for(i in 1:n){
w1=dnorm(vec_n[1,i], mean = 0, sd = 1)/pnorm(vec_n[1,i],
mean = 0, sd = 1)
w2=dnorm(vec_n[1,i], mean = 0, sd = 1)/(1-pnorm(vec_n[1,i],
mean = 0, sd = 1))

b1 = vec_n[1,i]*w1 + w1^2
b2 = w2^2 - vec_n[1,i]*w2
a1 = w1 + vec_n[1,i]*b1
a2 = w2 - vec_n[1,i]*b2
```

```
sum_urxr = sum_urxr + (((i-1)*b1 + (n-i)*b2 + 1)*rss_samp[1,i])
sum_urxr_sq = sum_urxr_sq + (((i-1)*b1 + (n-i)*b2 + 1)
*(rss_samp[1,i])^2)
        sum_wrxr = sum_wrxr + (((i-1)*a1 - (n-i)*a2) * rss_samp[1,i])
         }

         estimates[group,1] = sum_urxr/sum_ur
    B = sum_wrxr
    C = sum_urxr_sq - (2*estimates[group,1]*sum_urxr)
+ (estimates[group,1]^2 * sum_ur)
         estimates[group,2] = (-B + sqrt(B^2 + (4*n*C)))/(2*n)
         }   ####### END OF GENERATING M SAMPLE OF SIZE n
est_hatbar[h,1] = mean(estimates[,1])
est_hatbar[h,2] = mean(estimates[,2])

limits[h,1] = est_hatbar[h,1] - (3 * A1 * est_hatbar[h,2])
limits[h,2] = est_hatbar[h,1] + (3 * A1 * est_hatbar[h,2])
###########  START TESTING THE LIMITS
        count = 0  # for counting the values out of limits
for(test in 1:N_test){

 for(i in 1:n)
             rss[i,] = rnorm(n, mean = 0, sd = 1)

         rss = t(apply(rss,1,sort))
         rss_samp = diag(rss)

         sum_urxr = 0
         for(i in 1:n){

w1=dnorm(vec_n[1,i], mean = 0, sd = 1)/pnorm(vec_n[1,i],
mean = 0, sd = 1)
w2=dnorm(vec_n[1,i], mean = 0, sd = 1)/(1-pnorm(vec_n[1,i],
mean = 0, sd = 1))

b1 = vec_n[1,i]*w1 + w1^2
b2 = w2^2 - vec_n[1,i]*w2
sum_urxr = sum_urxr + (((i-1)*b1 + (n-i)*b2 + 1)*rss_samp[1,i])
         }

         mu_hat = sum_urxr/sum_ur

 if ((mu_hat<limits[h,1]) || (mu_hat>limits[h,2])) count=count+1
```

```
}  # end of test loop N_test
###########  END OF TESTING THE LIMITS
alpha[h,1] = count / N_test

}  # end of big for N
}   # end of function

# QUALITY CONTROL CHART BY RANKED SET SAMPLING
# UNDER SHORT TAILED SYMMETRIC DISTRIBUTION

cchart_rss_sts <- function(M, n, d){
N = 1000
N_test = 10000

h = 2-d
K = 1 / (1 + (1/h) + (0.75/h^2))

for (ii in 1:N) {
for (group in 1:M) {
rss_samp = sts_random_samples[id[group],]
rss_samp = as.matrix(rss_samp)
dim(rss_samp) = c(1,n)

sum_ur_star = 0
sum_urxr = 0
sum_urxr_star = 0
sum_wrxr = 0
sum_urxrsq = 0
for (i in 1:n) {
 w1 = dtikuv(vec_n[1,i],d) / ptikuv(vec_n[1,i],d)
 w2 = dtikuv(vec_n[1,i],d) / (1 - ptikuv(vec_n[1,i],d))

 b1 = (1 - (vec_n[1,i]^2/(2*h))) / (1 + (vec_n[1,i]^2/(2*h)))^2
 b2 = w1^2 + (w1 * vec_n[1,i] * (1 - (4/(2*h + vec_n[1,i]^2))))
 b3 = w2^2 - (w2 * vec_n[1,i] * (1 - (4/(2*h + vec_n[1,i]^2))))

 if (d<=0) { a1 = (vec_n[1,i]^3/h) / (1+(vec_n[1,i]^2/(2*h)))^2
b1_star = b1
} else{a1=((vec_n[1,i]^3/h)+(vec_n[1,i]*(1-(h/2))))
/(1+(vec_n[1,i]^2/(2*h)))^2
b1_star=((h/2)-(vec_n[1,i]^2/(2*h)))
/(1+(vec_n[1,i]^2/(2*h)))^2
}
```

```
a2 = w1 + (vec_n[1,i] * b2)
a3 = w2 - (vec_n[1,i] * b3)

sum_ur_star = sum_ur_star + ((i-1)*b2 + (n-i)*b3 + 1
- (2/h)*b1_star)
sum_urxr = sum_urxr + (rss_samp[1,i] * ((i-1)*b2 + (n-i)*b3 + 1
- (2/h)*b1))
sum_urxr_star=sum_urxr_star+(rss_samp[1,i]*((i-1)*b2+(n-i)*b3
  + 1 - (2/h)*b1_star))
sum_wrxr = sum_wrxr + (rss_samp[1,i] * ((2/h)*a1 + (i-1)*a2
- (n-i)*a3))
sum_urxrsq=sum_urxrsq+(rss_samp[1,i]^2 * ((i-1)*b2 + (n-i)*b3
+ 1 - (2/h)*b1_star))
}

est_group[group,1] = sum_urxr / sum_ur

B = sum_wrxr
C = sum_urxrsq-(2*est_group[group,1]*sum_urxr_star)
+(est_group[group,1]^2 * sum_ur_star)
est_group[group,2] = (-B + sqrt(B^2 + (4*n*C))) / (2 * n)
}

est_bars[ii,] = apply(est_group,2,mean)
limits[ii,1] = est_bars[ii,1] - (3 * A1 * est_bars[ii,2])
limits[ii,2] = est_bars[ii,1] + (3 * A1 * est_bars[ii,2])


count = 0
test_id = sample(dim_samples[1], N_test, replace=F)
#################### START TESTING THE LIMITS
for(test in 1:N_test) {
rss_test = sts_random_samples[test_id[test],]
rss_test = as.matrix(rss_test)
dim(rss_test) = c(1,n)

sum_urxr = 0
for(j in 1:n) {
w1 = dtikuv(vec_n[1,j],d) / ptikuv(vec_n[1,j],d)
w2 = dtikuv(vec_n[1,j],d) / (1 - ptikuv(vec_n[1,j],d))
b1 = (1 - (vec_n[1,j]^2/(2*h))) / (1 + (vec_n[1,j]^2/(2*h)))^2
b2 = w1^2 + (w1 * vec_n[1,j] * (1 - (4/(2*h + vec_n[1,j]^2))))
```

```
b3 = w2^2 - (w2 * vec_n[1,j] * (1 - (4/(2*h + vec_n[1,j]^2))))

sum_urxr = sum_urxr + (rss_test[1,j] * ((j-1)*b2 + (n-j)*b3
  + 1 - (2/h)*b1))
}

mu_hat = sum_urxr / sum_ur

if (mu_hat < limits[ii,1] || mu_hat > limits[ii,2]) count=count+1
}  ############################  END OF TESTING THE LIMITS

alpha[ii,1] = count / N_test

}  }

# QUALITY CONTROL CHART BY RANKED SET SAMPLING
# UNDER LONG TAILED SYMMETRIC DISTRIBUTION
# THREE-MOMENT T APPROXIMATION

cchart_rss_lts_tmoment <- function(M,n,p)
{
  N = 1000
  N_test = 10000

  v=2*p-1
  k=2*p-3

  dof_v = (2*(2*kurt - 3)) / (kurt - 3)
  for(h in 1:N){
    for(group in 1:M){

      rss = rt(n*n,v)*sqrt(k/v)
      dim(rss) = c(n,n)
      rss = t(apply(rss,1,sort))
      rss_samp = diag(rss)

      sum_ur = 0
      sum_urxr = 0
      sum_urxr_sq =0
      sum_wrxr = 0
      for(i in 1:n){

        w1 = (dt((vec_n[1,i]*sqrt(v/k)), df=v) * sqrt(v/k))
```

```
/ (pt((vec_n[1,i]*sqrt(v/k)), df=v))
        w2 = (dt((vec_n[1,i]*sqrt(v/k)), df=v) * sqrt(v/k))
/ (1 - pt((vec_n[1,i]*sqrt(v/k)), df=v))

        b1 = (1 - (vec_n[1,i]^2 / k))/(1 + (vec_n[1,i]^2 / k))^2
        b2 = w1*((2*p*vec_n[1,i])/(k+vec_n[1,i]^2)) + w1^2
        b3 = w2^2 - (w2*((2*p*vec_n[1,i])/(k+vec_n[1,i]^2)))
        a1 = ((2/k)*vec_n[1,i]^3)/(1 + (vec_n[1,i]^2/k))^2
        a2 = w1 + (b2*vec_n[1,i])
        a3 = w2 - (b3*vec_n[1,i])

        sum_ur = sum_ur + (((2*p/k)*b1) + ((i-1)*b2) + ((n-i)*b3))
        sum_urxr = sum_urxr + (rss_samp[i]*(((2*p/k)*b1) + ((i-1)*b2)
+ ((n-i)*b3)))
        sum_urxr_sq = sum_urxr_sq + (rss_samp[i]^2 * (((2*p/k)*b1)
+ ((i-1)*b2) + ((n-i)*b3)))
        sum_wrxr = sum_wrxr + (rss_samp[i]*(((i-1)*a2) - ((n-i)*a3)
- ((2*p/k)*a1)))
      }

      estimates[group,1] = sum_urxr/sum_ur
      B = sum_wrxr
      C = sum_urxr_sq - ((sum_urxr)^2/sum_ur)

      if (C<0) {  # checking whether sigma is real and positive
        sum_ur = 0
        sum_urxr = 0
        sum_urxr_sq =0
        sum_wrxr = 0
        for(i in 1:n){

          w1 = (dt((vec_n[1,i]*sqrt(v/k)), df=v) * sqrt(v/k))
/ (pt((vec_n[1,i]*sqrt(v/k)), df=v))
          w2 = (dt((vec_n[1,i]*sqrt(v/k)), df=v) * sqrt(v/k))
/(1 - pt((vec_n[1,i]*sqrt(v/k)), df=v))

          b1 = 1/(1 + (vec_n[1,i]^2 / k))
          b2 = w1*((2*p*vec_n[1,i])/(k+vec_n[1,i]^2)) + w1^2
          b3 = w2^2 - (w2*((2*p*vec_n[1,i])/(k+vec_n[1,i]^2)))
          a2 = w1 + (b2*vec_n[1,i])
          a3 = w2 - (b3*vec_n[1,i])

          sum_ur = sum_ur + (((2*p/k)*b1) + ((i-1)*b2) + ((n-i)*b3))
```

```
          sum_urxr = sum_urxr + (rss_samp[i]*(((2*p/k)*b1) + ((i-1)*b2)
+ ((n-i)*b3)))
          sum_urxr_sq = sum_urxr_sq + (rss_samp[i]^2 * (((2*p/k)*b1)
+ ((i-1)*b2) + ((n-i)*b3)))
          sum_wrxr = sum_wrxr + (rss_samp[i]*(((i-1)*a2) - ((n-i)*a3)))
        }

        B = sum_wrxr
        C = sum_urxr_sq - ((sum_urxr)^2/sum_ur)

      }  # end of checking whether sigma is real and positive

      estimates[group,2] = (-B + sqrt(B^2 + (4*n*C)))/(2*n)
      }
    est_hatbar[h,] = apply(estimates,2,mean)
    bb = (sqrt((dof_v - 2)/dof_v) * A1 * est_hatbar[h,2])

    limits[h,1] = bb * qt((0.0027/2), dof_v) + est_hatbar[h,1]
    limits[h,2] = bb * qt((1-(0.0027/2)), dof_v) + est_hatbar[h,1]
    ###########  START TESTING THE LIMITS
    count = 0  # for counting the values out of limits
    for(test in 1:N_test){
      rss = rt(n*n,v)*sqrt(k/v)
      dim(rss) = c(n,n)
      rss = t(apply(rss,1,sort))
      rss_samp = diag(rss)

      sum_ur = 0
      sum_urxr = 0
      for(i in 1:n){

        w1 = (dt((vec_n[1,i]*sqrt(v/k)), df=v) * sqrt(v/k))
/ (pt((vec_n[1,i]*sqrt(v/k)), df=v))
        w2 = (dt((vec_n[1,i]*sqrt(v/k)), df=v) * sqrt(v/k))
/ (1 - pt((vec_n[1,i]*sqrt(v/k)), df=v))

        b1 = (1 - (vec_n[1,i]^2 / k))/(1 + (vec_n[1,i]^2 / k))^2
        b2 = w1*((2*p*vec_n[1,i])/(k+vec_n[1,i]^2)) + w1^2
        b3 = w2^2 - (w2*((2*p*vec_n[1,i])/(k+vec_n[1,i]^2)))

        sum_ur = sum_ur + (((2*p/k)*b1) + ((i-1)*b2) + ((n-i)*b3))
        sum_urxr = sum_urxr + (rss_samp[i]*(((2*p/k)*b1) + ((i-1)*b2)
+ ((n-i)*b3)))
```

```
    }

    mu_hat = sum_urxr/sum_ur

  if ((mu_hat<limits[h,1]) || (mu_hat>limits[h,2])) count=count+1
  }
  ###########  END OF TESTING THE LIMITS
  alpha[h,1] = count / N_test
}  }   # end of function
```