Multi-Stage Estimation Methods with Termination Defined Via Gini's Mean Difference or Mean Absolute Deviation

Jun Hu
University of Connecticut - Storrs, jun.hu@uconn.edu

Follow this and additional works at: https://opencommons.uconn.edu/dissertations

Recommended Citation
https://opencommons.uconn.edu/dissertations/1872
Multi-Stage Estimation Methods with Termination Defined Via Gini’s Mean Difference or Mean Absolute Deviation

Jun Hu, Ph.D.
University of Connecticut, 2018

ABSTRACT

This thesis consists of multi-stage methodologies handling two fundamental estimation problems. These are (i) fixed-width confidence interval estimation (FCIE), and (ii) minimum risk point estimation (MRPE) problems for the unknown mean \( \mu \) of a normal distribution whose variance \( \sigma^2 \) is also assumed unknown.

We first develop purely sequential estimation methodologies for both FCIE and MRPE problems. New stopping rules are constructed by replacing the sample variance with appropriate multiples of Gini’s mean difference (GMD) and mean absolute deviation (MAD) in defining the conditions for boundary crossing. A number of asymptotic first-order consistency, efficiency, and risk efficiency properties associated with these new estimation strategies has been investigated. These are followed by summaries obtained from extensive sets of simulations by drawing samples from (i) normal universes or (ii) mixture-normal universes where samples may be reasonably treated as observations from a normal universe in a large majority of simulations. We also include illustrations using sales data and horticulture data.
By revisiting Stein (1945,1949) as well as Mukhopadhyay and Duggan (1997), we then move on to propose new two-stage estimation methodologies under both MRPE and FCIE configurations for a normal mean $\mu$ when a lower bound of variance $\sigma^2_L (0 < \sigma_L < \sigma)$ is known to us. Again, unbiased estimators based on sample standard deviation, GMD and MAD are constructed to define the final sample sizes. The new methodologies prove to enjoy both asymptotic first-order and second-order properties, followed by simulated performances. Real data illustrations of horticulture data are also included.

Next, we further explore the asymptotic second-order approximations for the regret function associated with the purely sequential MRPE methodologies, providing a general structure.

Overall, we empirically feel confident that our newly developed GMD-based or MAD-based multi-stage estimation methodologies are more robust for practical purposes when we compare them with the sample variance-based methodologies respectively, especially when suspect outliers may be expected. We conclude with some interesting directions of future research work that we can follow to make our multi-stage estimation methodologies more widely applicable for a lot of inference problems.
Multi-Stage Estimation Methods with Termination Defined
Via Gini’s Mean Difference or Mean Absolute Deviation

Jun Hu

B.Sc., Wuhan University, China, 2014

A Dissertation
Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy
at the
University of Connecticut
2018
Doctor of Philosophy Dissertation

Multi-Stage Estimation Methods with Termination Defined
Via Gini’s Mean Difference or Mean Absolute Deviation

Presented by

Jun Hu, B.Sc.

Major Advisor

Nitis Mukhopadhyay

Associate Advisor

Joseph Glaz

Associate Advisor

Vladimir Pozdnyakov

University of Connecticut

2018
ACKNOWLEDGEMENTS

In the blink of an eye, it has been four years since I joined the Department of Statistics at the University of Connecticut in Fall 2014. Standing here in retrospect, I would like to express my sincere and foremost gratitude to my mother Xuefang Zhong and my father Lifeng Hu for their infinite moral and emotional support in every stage of my life. I am also grateful to a list of my other family members and friends throughout the four years of my Ph.D. program, without whose support I could have never ever gone this far, with a special mention to Aiqing Zhong, Yuqing Zhong and Xiangyi Li.

I would also like to thank my advisor Professor Nitis Mukhopadhyay from the bottom of my heart. It was so fortunate for me to have taken Prof. Mukhopadhyay’s math stat courses when I came to UConn, where I was very impressed by his passion for Statistics. His comprehensive knowledge in Statistics as well as wide-ranged collection of anecdotes always kept me spellbound. I believe it was the best decision, which was not difficult to make at all, to select Prof. Mukhopadhyay as my academic advisor. Over these past four years, his abundant experience has kept me from getting detoured along the way; and his warm encouragements have eliminated most frustrations in my research. I also highly appreciated Prof. Mukhopadhyay’s dedication in his teaching and research as well as his commitment to students. This has firmed my desire to plunge into academia further upon my graduation. I cherished the opportunity to work with Prof. Mukhopadhyay under his guidance, and hope to continue working with him in the future.

My sincere gratitude also goes to Prof. Ming-Hui Chen, who played the role as my program advisor during my initial year at UConn. I benefited a lot from his advise about
how to get acclimated to the program. I would like to extend my gratitude to Prof. Joseph
Glaz and Prof. Vladimir Pozdnyakov, my associate advisors for their continuous support in
research. They were always ready to provide help whenever I had any questions. I would
also like to especially thank Prof. Richard Vitale and Prof. Zhiyi Chi for their time to take
charge in my general exam. I am also grateful to all the other professors who have taught
me during the four years of my Ph.D. program.

Many thanks are owed to the following department staff: Megan Petsa, Tracy Burke
and Anthony Luis for their dedicated service over these past few years, which have made
our lives much easier and let us feel at home.

And finally, last but not least, I would like to thank a lot of my fellow graduate students,
without whom my time at UConn would have never been full of joy and memories. I will
cherish our friendship forever.
Table of Contents

Chapter 1: Introduction 1

1.1. Literature Review 1

1.2. Thesis Outline 4

Chapter 2: Formulations for Two Fundamental Estimation Problems:

Fixed-Width Confidence Intervals and Minimum Risk Point Estimation 6

2.1. Fixed-Width Confidence Interval Estimation (FCIE) 6

2.2. Minimum Risk Point Estimation (MRPE) 8

2.3. Introducing Gini’s Mean Difference (GMD) and Mean Absolute Deviation (MAD) 8

Chapter 3: Purely Sequential Methodologies Involving GMD or MAD 11

3.1. Fixed-Width Confidence Intervals 12

3.1.1. Some Useful Lemmas 13

3.1.2. Asymptotic First-Order Properties of the Purely Sequential Methodology (3.1.1)-(3.1.2) 17

3.1.3. Limited Robustness Properties of the Purely Sequential Methodology (3.1.1)-(3.1.2) 19

3.1.4. Simulations and Data Analysis 21

3.1.4.1. Samples from a Normal Population 22

3.1.4.2. Samples from an $\epsilon$-Contaminated Normal Population (3.1.7) 24
3.1.4.3. Some Overall Sentiments 26

3.2. Minimum Risk Point Estimation 28

3.2.1. Asymptotic First-Order Properties of the Purely Sequential Methodology (3.2.1)-(3.2.2) 29

3.2.2. Simulations and Data Analysis 35

3.2.2.1. Samples from a Normal Population 36

3.2.2.2. Samples from an $\epsilon$-Contaminated Normal Population (3.1.7) 37

3.2.2.3. Some Overall Sentiments 37

3.3. Illustrations with Real Datasets 38

3.3.1. Illustrations Using Sales Data 39

3.3.2. Illustrations Using Horticulture Data 40

3.4. Overall Concluding Thoughts 42

Chapter 4: Two-Stage Methodologies Involving GMD or MAD 54

4.1. Minimum Risk Point Estimation 55

4.1.1. Existing Two-Stage MRPE Methodologies 55

4.1.2. Substituting Sample Standard Deviation with Alternative Estimators 57

4.1.3. Properties of the Two-Stage Methodology (4.1.4)-(4.1.5) with $m$ from (4.1.2) 58

4.1.4. Simulations and Data Analysis 66

4.2. Fixed-Width Confidence Intervals 67

4.2.1. A Formulation 68

4.2.2. Properties of the Two-Stage Methodology (4.2.2)-(4.2.3) 69
4.2.3. Simulations and Data Analysis 71
4.3. Illustrations with Real Datasets 72
  4.3.1. MRPE Methodologies 73
  4.3.2. FCIE Methodologies 74
4.4. Overall Concluding Thoughts 74

Chapter 5: Second-Order Asymptotics in a Class of Purely Sequential MRPE Methodologies 80
  5.1. A Formulation 81
    5.1.1. Sufficient Conditions on Robust Estimators 81
    5.1.2. Purely Sequential MRPE Methodologies Involving $V_n$ 82
  5.2. Main Results 83
  5.3. An Application: Purely Sequential Methodology (3.2.1)-(3.2.2) 90
    5.3.1. Second-Order Approximations of $\omega_i(c), i = 0, 1, 2$ 90
    5.3.2. Simulations and Data Analysis 91
  5.4. Overall Concluding Thoughts 91

Chapter 6: A Brief Summary 94

Chapter 7: Future Work 95
  7.1. Other Multi-Stage Methodologies 95
  7.2. Distribution-Free Methodologies 96
  7.3. Multivariate Multi-Stage Estimation 97

List of Tables 98

Bibliography 99
Chapter 1

Introduction

In this thesis, we revisit both of the fixed-width confidence interval estimation (FCIE) and minimum risk point estimation (MRPE) problems under multi-stage stopping rules for an unknown normal mean $\mu$ when the population variance $\sigma^2$ also remains unknown. We propose to replace the sample standard deviation with appropriate multiples of either Gini’s mean difference (GMD) or mean absolute deviation (MAD) in defining the stopping boundaries. We do in order to accommodate possibilities of encountering outlying observations even though we consider a normal population distribution under investigation.

1.1. Literature Review

The sampling designs of multi-stage methodologies have been growing rapidly over the past decades. Stein (1945,1949) had put forward his ingenious two-stage strategy for gathering data. His fundamental contribution solved the fixed-width confidence interval estimation (FCIE) problem for a normal mean $\mu$ with preassigned coverage probability when the variance $\sigma^2$ is unknown. Later, Anscombe (1952,1953) developed breakthrough large-sample theories of sequential estimation. In his 1952 paper, Anscombe gave the fundamental formulation of the random central limit theorem. His 1953 paper proposed a purely sequential sampling strategy for the same problem. Ray (1957) broadened Anscombe’s (1952,1953) ideas by modifying the stopping boundary.

Chow and Robbins (1965) developed pioneering asymptotic theories of FCIE problems
in the distribution-free case and laid the foundation for ground-breaking results. These eventually led to a broad class of crucial nonlinear renewal theoretic tools for arriving at asymptotic second-order approximations developed by Woodroofe (1977,1982) and Lai and Siegmund (1977,1979).

In order to reduce the oversampling problem associated with Stein’s (1945,1949) two-stage methodology, Mukhopadhyay (1980) designed a modification giving rise to a surprising two-stage estimation strategy with the asymptotic efficiency property. This direction led Ghosh and Mukhopadhyay (1981) to give the foundation of the notion of asymptotic second-order efficiency property. Mukhopadhyay and Duggan (1997,1999) revisited this problem when a positive lower bound of the population variance \( \sigma^2 \) is known and proved that their two-stage methodology enjoyed both asymptotic first-order and second-order efficiency properties.

In a parallel path, Robbins (1959) introduced the pioneering formulation for minimum risk point estimation (MRPE) problems for a normal mean \( \mu \). He put forward a purely sequential stopping rule to estimate \( \mu \). Starr (1966) and Starr and Woodroofe (1969) subsequently broadened Robbins’ (1959) methodology in more details and concluded a number of interesting asymptotic properties enjoyed by that procedure. Second-order approximations were summarized by Woodroofe (1977,1982), Siegmund (1985), Ghosh et al. (1997) among other sources. Mukhopadhyay and Duggan (1997,1999) also incorporated their new Stein-type (1945,1949) methodology in this area.

In traditional normal problems, the boundary crossing criteria often utilized the sample variance unbiasedly estimating the population variance \( \sigma^2 \) at every step of the way. In a different direction, in lieu of using the sample variance, some researchers had looked into other estimators such as those based on GMD and MAD, both explicitly defined in (2.3.1).
GMD was originally proposed by Gini (1914,1921).

Later, Nair (1936) developed the standard error of GMD so that one could specify an unbiased estimator for the population variance based on GMD and its standard error. This was way before Hoeffding (1948,1961) came up with his pathbreaking theory of $U$-statistics. The literature on sequential $U$-statistics is very broad. One may review from Lee (1990), Mukhopadhyay and Vik (1985), Sen (1981,1985), Jurečková and Sen (1996) and other sources.

Revisiting Downton (1966) and Barnett et al. (1967), David (1968) used a linear function of order statistics as an estimator of the standard deviation of a normal distribution which turned out to be essentially GMD. See also Sen (1986). Barnett et al. (1967) established an unbiased estimator of the standard deviation based on GMD ($G$) and investigated the role of a sample mean standardized with $G$ as a competitor of the usual Student’s $t$ test when the sample size is fixed. Mukhopadhyay and Chattopadhyay (2013) recently revisited this problem charting interesting new directions.

Yitzhaki (2003) studied properties of GMD as a superior measure of variability than the usual sample variance in some non-normal cases. Meanwhile, MAD had been adopted as an estimator of the standard deviation, too. Herrey (1965) derived the expression of the variance of MAD using results from Helmert (1876). Herrey (1965) focused on MAD and constructed confidence intervals based on MAD instead of the sample standard deviation. Babu and Rao (1992) worked on expansions for statistics involving MAD. We may note that MAD is widely used in portfolio optimization models. See Konno and Yamazaki (1991) where MAD was preferred. One may also refer to Markowitz (1959).

In Chattopadhyay and Mukhopadhyay (2013), both GMD and MAD were used to construct two-stage FCIE methodologies for a normal mean $\mu$ in the presence of suspect outliers.
and these were compared with the customary Stein-type (1945,1949) two-stage methodology based on the sample variance. We have found interesting recent trends in sequential estimation problems for Gini’s index itself from Chattopadhyay and De (2016) and De and Chattopadhyay (2017). Mukhopadhyay and Hu (2017,2018) and Hu and Mukhopadhyay (2018) have extended the application of GMD and MAD into more multi-stage sampling strategies for FCIE and MRPE problems.


1.2. Thesis Outline

In this thesis we introduce multi-stage methodologies associated with two fundamental estimation problems, namely

(i) the fixed-width confidence interval estimation (FCIE), and
(ii) the minimum risk point estimation (MRPE).

We mainly focus on the purely sequential and two-stage sampling designs with the purpose to estimate an unknown normal mean.

Chapter 2 begins with basic formulations of FCIE and MRPE for an unknown mean $\mu$ in a $N(\mu, \sigma^2)$ population where $\sigma^2$ is also assumed unknown. Alternative robust estimators of $\sigma^2$ other than the sample variance are introduced, including appropriate functions of GMD and MAD.
In Chapter 3, we develop purely sequential FCIE and MRPE methodologies with estimators of $\sigma^2$ based on GMD and MAD as well as the customary sample variance in defining the stopping boundary conditions. We discuss a series of interesting asymptotic first-order properties as well as the limited robustness, followed by extensive sets of simulations as a reasonable validation. Real data analyses are also provided to address the possible application of our newly proposed methodologies in the fields of economics and agriculture with illustrations using (i) sales data and (ii) horticulture data. The chapter largely comes from the publication, Mukhopadhyay and Hu (2017).

In the contexts of both FCIE and MRPE problems, Chapter 4 introduces newly proposed GMD-based and MAD-based two-stage estimation methodologies, assuming that a positive lower bound $\sigma_L^2$ for the variance $\sigma^2$ is known to us in the spirit of Mukhopadhyay and Duggan (1997, 1999). We lay out a number of desirable asymptotic first-order and second-order properties. And simulated performances as well as illustrations with real datasets are also presented. Chapter 4 is based on the publication, Mukhopadhyay and Hu (2018).

In Chapter 5, we again focus on the purely sequential MRPE methodologies based on robust estimators of $\sigma$. However, we provide a general structure with appropriate sufficient conditions on such estimators, which will allow us in general to claim that the associated methodologies would enjoy both asymptotic first-order and second-order asymptotic properties. The chapter is based on Hu and Mukhopadhyay (2018), which has been submitted for publication.

Chapter 6 gives a brief summary of our work. In Chapter 7, we provide a lot of possible directions that we can follow for future research to make our methodologies more applicable for practical purposes.
Chapter 2

Formulations for Two Fundamental Estimation Problems: Fixed-Width Confidence Intervals and Minimum Risk Point Estimation

More than forty years ago, Ghosh and Mukhopadhyay (1976) gave a broad review of both sequential and purely sequential methodologies for constructing fixed-width confidence interval and minimum risk point estimators for unknown \( \mu \) in a \( N(\mu, \sigma^2) \) population having \( \sigma^2 \) also unknown. In the context of our present investigation, that paper remains relevant even today.

We begin with a sequence of independent observations \( X_1, X_2, \ldots \) from a \( N(\mu, \sigma^2) \) population with \(-\infty < \mu < \infty, 0 < \sigma^2 < \infty\), both parameters unknown. Having recorded \( X_1, \ldots, X_n, n \geq 2 \), we denote the customarily used unbiased estimators for \( \mu \) and \( \sigma^2 \) as follows:

- Sample mean: \( \bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i \)
- Sample variance: \( S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2 \).

In Sections 2.1-2.2, we respectively lay out the fixed-width confidence interval estimation methodology and the minimum risk point estimation methodology.

2.1. Fixed-Width Confidence Interval Estimation (FCIE)

In order to estimate the mean \( \mu \), first we aim to construct a confidence interval with preassigned (i) fixed-width \( 2d(> 0) \) and (ii) the confidence coefficient at least \( 1 - \alpha \), or
approximately at least $1 - \alpha$, $0 < \alpha < 1$. The customary fixed-width confidence interval

$$J_n = [\bar{X}_n \pm d] \quad (2.1.1)$$

is considered, which additionally must satisfy the requirement that $P_{\mu,\sigma}\{\mu \in J_n\}$ is at least $1 - \alpha$, or approximately $1 - \alpha$, with $0 < \alpha < 1$ preassigned. Dantzig (1940) proved that there did not exist any fixed-sample-size procedure to solve the problem. Pathbreaking papers that provided a solution to this problem were due to Stein (1945, 1949) by invoking the ingenious two-stage sampling methodology.

From the problem specifications, one has:

$$P_{\mu,\sigma}\{\mu \in J_n\} = P_{\mu,\sigma}\{|\bar{X}_n - \mu| \leq d\} = 2\Phi(\sqrt{nd}/\sigma) - 1,$$

so that

$$P_{\mu,\sigma}\{\mu \in J_n\} \geq 1 - \alpha \quad (2.1.2)$$

provided that $n$ is the smallest integer satisfying:

$$\sqrt{nd}/\sigma \geq a \Leftrightarrow n \geq a^2\sigma^2/d^2 = C, \text{ say.}$$

Here, we write $a \equiv z_{\alpha/2}$, the upper $100(\alpha/2)$% point of a standard normal distribution. Thus, we define the optimal fixed sample size, had $\sigma^2$ been known, as follows:

$$C \equiv C(d) = a^2\sigma^2/d^2, \quad (2.1.3)$$

by tacitly disregarding the fact that $C$ may not be an integer.
2.2. Minimum Risk Point Estimation (MRPE)

On the other hand, Robbins (1959) proposed his original and fundamental formulation of a MRPE problem for the normal mean $\mu$. We will consider the loss function given by the squared error loss (SEL) due to estimation of $\mu$ with $\bar{X}_n$ plus linear cost of sampling:

$$L_n = L_n(\mu, \bar{X}_n) = A(\bar{X}_n - \mu)^2 + cn,$$

where $A$ and $c$ are both known. (2.2.1)

Here, $A(>0)$ is an appropriate weight function and $c(>0)$ is the cost of each observation and “$n$” obviously indicates a sample size.

The risk function associated with (2.2.1) is then given by:

$$R_n(c) = E_{\mu, \sigma} [L_n(\mu, \bar{X}_n)] = A\sigma^2 n^{-1} + cn.$$ (2.2.2)

We can thus obtain the optimal fixed sample size, $n^* = n^*(c)$, had $\sigma^2$ been known, which minimizes the risk function $R_n(c)$ from (2.2.2). We have:

$$n^* = n^*(c) = \sigma \sqrt{A/c},$$ (2.2.3)

again by tacitly disregarding the fact that $n^*$ may not be an integer.

2.3. Introducing Gini’s Mean Difference (GMD) and Mean Absolute Deviation (MAD)

One should note that in (2.1.3) and (2.2.3), $\sigma^2$ is actually assumed unknown. Therefore, it is essential for us to estimate $\sigma^2$ by updating its estimator at various stages of the methodologies as needed. A customarily used unbiased estimator of $\sigma^2$ is the sample
variance, which, however, in certain scenarios, has its own drawback in estimating the population variance. For example, when skewness appears pronounced due to the existence of possible outliers even though the parent population is assumed normal, sample variance may not estimate $\sigma^2$ well. In nonsequential portfolio theory, Markowitz (1959) suggested using semivariance as a substitute of the sample variance.

Back to multi-stage estimation problems, Mukhopadhyay (1982) opened the possibilities of using estimators of $\sigma^2$ other than the sample variance. Recall that Chattopadhyay and Mukhopadhyay (2013), estimators based on GMD and MAD were used for constructing two-stage fixed-width confidence interval methodologies for a normal mean $\mu$ in the presence of suspect outliers and these were compared with the customary Stein-type (1945,1949) two-stage methodology based on the sample variance.

In this thesis, we therefore propose to push similar ideas lot further and utilize alternative more robust estimators for $\sigma^2$ based on (i) GMD and (ii) MAD, as needed, formally defined in (2.3.1).

Having recorded $X_1, ..., X_n$ from a $N(\mu, \sigma^2)$ population, we define GMD and MAD as follows:

(i) GMD: $G_n = \frac{(n)}{2}^{-1} \Sigma_{1 \leq i < j \leq n} |X_i - X_j|$

(ii) MAD: $M_n = n^{-1} \Sigma_{i=1}^{n} |X_i - \bar{X}_n|$

Unbiased estimators for the population variance $\sigma^2$ based on GMD or MAD were proposed in Nair (1936) and Herrey (1965) respectively. In the spirit of notation used in Chattopadhyay and Mukhopadhyay (2013), we denote two unbiased estimators as well as the sample variance by $U_{n;i}^2, i = 0, 1, 2$, for $\sigma^2$ as follows:
where we denote
\[
c_{n;1} = \left\{ \frac{4}{\pi n(n-1)} \left( \frac{\pi}{3} (n+1) + 2(n-2)\sqrt{3} + n^2 - 5n + 6 \right) \right\}^{1/2},
\]
\[
c_{n;2} = \left\{ \left( \frac{\pi}{2} + \arcsin \left( \frac{1}{n-1} \right) - n + \sqrt{n(n-2)} \right) \frac{2(n-1)}{\pi n^2} + \frac{2(n-1)}{\pi n} \right\}^{1/2}.
\]

We also construct unbiased estimators for \( \sigma \) based on GMD or MAD as well as the sample standard deviation as follows:

\[
(0) \quad \hat{\sigma} \equiv V_{n;0} = \frac{\Gamma \left( \frac{n-1}{2} \right)}{\sqrt{2\pi} \Gamma \left( \frac{n}{2} \right)} S_n;
\]
\[
(1) \quad \hat{\sigma} \equiv V_{n;1} = \sqrt{\frac{\pi}{2}} G_n;
\]
\[
(2) \quad \hat{\sigma} \equiv V_{n;2} = \sqrt{\frac{\pi n}{2(n-1)}} M_n.
\]

One will find many more specific details in Chapters 3 and 4, respectively, dealing with the two fundamental multi-stage estimation problems, namely, (i) FCIE and (ii) MRPE for the normal mean. Throughout the text, we write \( \phi(z) = (\sqrt{2\pi})^{-1} \exp(-z^2/2) \) and \( \Phi(z) = \int_{-\infty}^{z} \phi(y)dy, \ z \in R. \)
Chapter 3

Purely Sequential Methodologies
Involving GMD or MAD

In this chapter, we propose the GMD-based and MAD-based purely sequential methodologies for both fixed-width confidence interval estimation (FCIE) and minimum risk point estimation (MRPE) problems. In comparison, the customary sample variance-based purely sequential methodology is also included. The chapter is mainly based on Mukhopadhyay and Hu (2017).

Section 3.1 develops purely sequential estimation strategies with stopping boundaries based on the sample variance, GMD or MAD (defined in 2.3.1-2.3.3) for the FCIE problems. Three main results have been summarized in Theorems 3.1.1-3.1.3, which show that our newly proposed purely sequential FCIE strategies enjoy asymptotic first-order efficiency and asymptotic consistency properties.

Section 3.2 develops purely sequential estimation strategies with stopping boundaries based on the sample standard deviation, GMD or MAD (defined in 2.3.1 and 2.3.4) for the MRPE problems. Three main results are summarized in Theorems 3.2.1-3.2.3, which show that our newly proposed purely sequential MRPE strategies enjoy asymptotic first-order efficiency and first-order risk efficiency.

In both Sections 3.1 and 3.2, we have presented summaries of data analysis obtained via extensive sets of simulations with random samples drawn from a normal as well as a mixture normal distribution (3.1.7). The mixture normal scenarios were so constructed that
outlying observations could be expected, but the gathered data could overwhelmingly be
treated as random samples from a normal universe for practical purposes. We are reassured
that the new GMD-based and MAD-based purely sequential methodologies are more robust
(Theorem 3.1.4) than the existing sample variance-based purely sequential methodologies.
This happens in both problems when we may expect to encounter some outliers as we
sample from a normal universe.

Section 3.3 includes illustrations of both FCIE and MRPE methodologies for the mean
($\mu$) using two real datasets: (a) net profit or loss percentages (sales data) from department
stores (McNair 1930; [http://www.stat.ufl.edu/~winner/], and (b) number of days that seeds
of marigold varieties 2 and 3 needed to flower (horticulture data) from Mukhopadhyay et
al. (2004b).

The chapter ends with some brief overall thoughts in Section 3.4.

3.1. Fixed-Width Confidence Intervals

Under the fixed-width confidence interval formulation (2.1.1)-(2.1.3), we propose the
following sampling procedure to determine the final sample size. Beginning with pilot data
$X_1, \ldots, X_m$ of size $m$, $m \geq 2$, we record one additional observation at-a-time successively as
needed until we stop according to the following stopping rule:

$$N_i = N_i(d) = \inf \{n \geq m : n \geq a^2 U_{ni}^2 / d^2 \}, i = 0, 1, 2.$$  \hspace{1cm} (3.1.1)

Observe that the index $i = 0, 1, 2$ respectively corresponds to the customary sample variance-
based, the GMD-based or the MAD-based methodology with $U_{ni}^2, i = 0, 1, 2$, defined via
(2.3.2)-(2.3.3).

That is, for $i = 0, 1, 2$, if $m \geq a^2 U_{ni}^2 / d^2$ is satisfied, we do not take any additional
observation and the sample size is $N_i = m$. Otherwise, we record one more observation and obtain updated $U_{m+1;i}^2$ in order to check with the stopping rule (3.1.1). We terminate sampling at the first time $N_i = n(\geq m)$ such that $n \geq a^2 U_{n;i}^2/d^2$ occurs. Finally, having observed the full dataset \{$N_i, X_1, \ldots, X_m, X_{m+1}, \ldots, X_{N_i}$\}, we construct the fixed-width confidence interval for $\mu$ as follows:

$$J_{N_i} = [\bar{X}_{N_i} \pm d]. \quad (3.1.2)$$

We should add that the customary sample variance-based methodology was first developed by Anscombe (1952, 1953), Ray (1957), and Chow and Robbins (1965). In Anscombe (1953), he used "$a"", the upper 100$(\alpha/2)$\% point of a standard normal distribution and his stopping rule was similar to (3.1.1) with $i = 0$. Ray (1957) used the upper 100$(\alpha/2)$\% point of the Student’s $t_{n-1}$ distribution instead of "$a" in defining the stopping boundary $(= t_{n-1, \alpha/2}^2 S_n^2/d^2)$ in (3.1.1). Obviously, \lim_{n \to \infty} t_{n-1, \alpha/2} = a$. One may refer to Mukhopadhyay (2010), Gut and Mukhopadhyay (2010), and also look at the references included therein. Chow and Robbins (1965) used a general sequence of positive constants \{${a}_n; n \geq 2$\} instead of "$a" in defining their stopping boundary $(= a_n^2 S_n^2/d^2)$ in (3.1.1) where \lim_{n \to \infty} a_n = a$.

For the purely sequential methodologies \{${N_i, J_{N_i}}$\} from (3.1.1)-(3.1.2), we can clearly claim that $P_{\mu, \sigma}\{N_i < \infty\} = 1$ and $N_i \uparrow \infty$ w.p.1 as $d \downarrow 0, i = 0, 1, 2$.

### 3.1.1. Some Useful Lemmas

We begin with a number of technical lemmas (Lemmas 3.1.1-3.1.4) which will be helpful in justifying some of the more substantial properties.
Lemma 3.1.1. The statistics \( \{U_{k;i}^2, k \leq n\}, i = 0, 1, 2, \) are distributed independently of \( \{X_n, n \geq m\} \).

Proof: We fix \( \sigma = \sigma_0(>0) \) and \( n(\geq m) \), but otherwise they remain arbitrary. Now, one can show easily that \( \{U_{k;i}^2, k \leq n\} \) consists of ancillary statistics for \( \mu \) and \( X_n \) is a complete and sufficient statistic for \( \mu \) in the model \( N(\mu, \sigma_0^2) \). Thus, by Basu's (1955) theorem, we immediately claim that \( \{U_{k;i}^2, k \leq n\} \) and \( X_n \) are independent under the model \( N(\mu, \sigma_0^2) \) for every fixed \( \sigma_0^2 \). But, since \( \sigma_0 \) is fixed but arbitrary, it means that \( \{U_{k;i}^2, k \leq n\} \) and \( X_n \) are independently distributed. ■

Lemma 3.1.2. \( \sum_{i=1}^{2} \text{P}_n \sigma^2 \) as \( n \to \infty \) for \( i = 0, 1, 2 \).

Proof: With \( i = 0 \), it is obvious that \( S_n^2 \xrightarrow{\text{P}_n} \sigma^2 \) as \( n \to \infty \).

With \( i = 1 \), we see that \( G_n \) from (2.3.1) is a \( U \)-statistic so that \( G_n \xrightarrow{\text{P}_n} E_{\mu, \sigma} [\mid X_1 - X_2 \mid] \) as \( n \to \infty \) in view of Hoeffding (1948,1961). See also Lee (1990). Since \( X_1 - X_2 \sim N(0, 2\sigma^2) \), we have \( E_{\mu, \sigma} [\mid X_1 - X_2 \mid] = 2\sigma/\sqrt{\pi} \) and thus \( G_n^2 \xrightarrow{\text{P}_n} 4\sigma^2/\pi \) as \( n \to \infty \). Also, \( c_n^{-2} \to \pi/4 \) as \( n \to \infty \). Thus, \( U_{n;1}^2 \xrightarrow{\text{P}_n} \sigma^2 \) as \( n \to \infty \).

With \( i = 2 \), we work with \( M_n \) from (2.3.1) and observe that with probability 1 (w.p.1):

\[
n^{-1}\sum_{i=1}^{n} |X_i - \mu| - |X_n - \mu| \leq M_n \leq n^{-1}\sum_{i=1}^{n} |X_i - \mu| + |X_n - \mu|.
\]

Obviously,

\[
|X_n - \mu| \xrightarrow{\text{P}_n} 0, \quad n^{-1}\sum_{i=1}^{n} |X_i - \mu| \xrightarrow{\text{P}_n} E_{\mu, \sigma} [\mid X_1 - \mu \mid] = \sigma\sqrt{2/\pi},
\]

as \( n \to \infty \). Therefore, \( M_n^2 \xrightarrow{\text{P}_n} 2\sigma^2/\pi \) as \( n \to \infty \). Combining this with the fact that \( c_n^{-2} \to \pi/2 \) as \( n \to \infty \), one immediately concludes that \( U_{n;2}^2 \xrightarrow{\text{P}_n} \sigma^2 \) as \( n \to \infty \). ■
Lemma 3.1.3. For any arbitrary $p > 1$, we have $E_{\mu, \sigma} \left[ \sup_n U_{n;i}^p \right] < \infty$, $i = 0, 1, 2$.

Proof. With $i = 1$, we have $U_{n;1}^p = c_{n;1}^p \mathcal{G}_n^p$. Clearly, $c_{n;1}^p \to (\pi/4)^{p/2} < 1$ as $n \to \infty$, so that there exists large enough $n_1$ and $c_{n;1}^p < 1$ for all $n > n_1$. Let us denote

$$k = \max \left\{ c_{1;1}^p, \ldots, c_{n;1}^p, 1 \right\}$$

and we can express:

$$\sup_n U_{n;1}^p \leq k \sup_n \left\{ \left( \sum_{1 \leq i < j \leq n} |X_i - X_j| \right)^p \right\} \leq 4^p k \left\{ \sup_n n^{-1} \sum_{1 \leq i \leq n} |X_i - \mu| \right\}^p .$$

We know that $|X_i - \mu|, i = 1, 2, \ldots$ are independent and identically distributed (i.i.d.) and belong to $L^p, p > 1$, and so does $\sup_n \left\{ n^{-1} \sum_{1 \leq i \leq n} |X_i - \mu| \right\}$ by Wiener's (1939) ergodic theorem. Hence, the result follows from (3.1.3).

With $i = 2$, we have $U_{n;2}^p = c_{n;2}^p \mathcal{M}_n^p$. Again, clearly, $c_{n;2}^p \to (\pi/2)^{p/2} < 1$ as $n \to \infty$, so that there exists large enough $n_2$ and $c_{n;2}^p < 2^p$ for all $n > n_2$. Let us denote

$$l = \max \left\{ c_{1;2}^p, \ldots, c_{n;2}^p, 2^p \right\}$$

and we can express:

$$\sup_n U_{n;2}^p \leq l \sup_n \left\{ n^{-1} \sum_{1 \leq i \leq n} |X_i - \mathcal{X}_n| \right\}^p \leq l \left\{ \sup_n \left( n^{-1} \sum_{1 \leq i \leq n} |X_i - \mu| + |\mathcal{X}_n - \mu| \right) \right\}^p ,$$

15
in the spirit of (3.1.3). Again, \( \sup_n \left\{ n^{-1} \sum_{i \leq j \leq n} |X_i - \mu| + |\overline{X}_n - \mu| \right\} \) belongs to \( L^p \) by invoking Wiener’s (1939) ergodic theorem so that we have \( E_{\mu,\sigma} \left[ \sup_n U_{n/2}^p \right] < \infty \).

With \( i = 0 \), one may refer to Chow and Robbins (1965) or look at another source.

**Lemma 3.1.4.** Under the purely sequential estimation strategy \((N_i, \overline{X}_{N_i})\) from (3.1.1)-(3.1.2), for all fixed \( \mu, \sigma, d, \) and \( \alpha \), we conclude the following properties:

(i) \( E_{\mu,\sigma}[N_i] < \infty, V_{\mu,\sigma}[N_i] < \infty; \)

(ii) \( E_{\mu,\sigma}[\overline{X}_{N_i}] = \mu, V_{\mu,\sigma}[\overline{X}_{N_i}] = \sigma^2 E_{\mu,\sigma}[N_i^{-1}]; \)

for \( i = 0, 1, 2 \).

**Proof:**

Part (i): We surely have \( P_{\mu,\sigma}\{N_i < \infty\} = 1 \). The stopping rule (3.1.1) provides the inequality (w.p.1):

\[
N_i - 1 \leq m - 1 + a^2 U_{N_i-1;i}^2 / d^2,
\]

which gives (w.p.1):

\[
N_i^2 \leq m^2 + 2m \frac{a^2}{d^2} \sup_n U_{n;i}^2 + \frac{a^4}{d^4} \sup_n U_{n;i}^4,
\]

(3.1.4)

\( i = 0, 1, 2 \). Thus, Lemma 3.1.3 and (3.1.4) show that \( E_{\mu,\sigma}[N_i^2] < \infty \) with \( p = 4 \). Part (i) follows.

Part (ii): The event \( \{N_i = n\} \) depends only on the vector \( U_i^* = \left( U_{m+1,i}^2, U_{m+2,i}^2, ..., U_{n;i}^2 \right) \) whereas \( U_i^* \) is independent of \( \overline{X}_n \) in view of Lemma 3.1.1. As a result, the event \( \{N_i = n\} \) and \( \overline{X}_n \) are independent for every fixed \( n \geq m \).
Thus, we have:

\[
E_{\mu,\sigma}[\bar{X}_{N_i}]
= \sum_{n=m}^{\infty} E_{\mu,\sigma}[\bar{X}_{N_i}|N_i = n] P_{\mu,\sigma}\{N_i = n\}
= \sum_{n=m}^{\infty} E_{\mu,\sigma}[\bar{X}_n] P_{\mu,\sigma}\{N_i = n\}
= \mu.
\]

Similarly, we can express:

\[
E_{\mu,\sigma}[\bar{X}^2_{N_i}] = \mu^2 + \sigma^2 E_{\mu,\sigma}[N_i^{-1}],
\]

\[i = 0, 1, 2.\] Part (ii) follows. 

3.1.2. Asymptotic First-Order Properties of the Purely Sequential Methodology (3.1.1)-(3.1.2)

In this section, we lay down a number of interesting first-order asymptotic properties in the form of Theorems 3.1.1-3.1.3 associated with the estimation strategies \((N_i, J_{N_i})\) proposed in (3.1.1)-(3.1.2), \(i = 0, 1, 2.\)

**Theorem 3.1.1.** For the stopping time \(N_i\) defined by the purely sequential estimation strategy (3.1.1)-(3.1.2), for all fixed \(\mu, \sigma, d,\) and \(\alpha,\) we have \(Q_i \equiv \sqrt{N_i}(\bar{X}_{N_i} - \mu)/\sigma \sim N(0, 1), i = 0, 1, 2.\)

**Proof:** For all \(z \in R,\) we have:

\[
P_{\mu,\sigma}\{Q_i \leq z\}
= \sum_{n=m}^{\infty} P_{\mu,\sigma}\{Q_i \leq z | N_i = n\} P_{\mu,\sigma}\{N_i = n\}
\]
which simplifies to \( \Phi(z), i = 0, 1, 2. \) ■

**Theorem 3.1.2.** For the stopping time \( N_i \) defined by the purely sequential estimation strategy \((3.1.1)-(3.1.2)\), for all fixed \( \mu, \sigma, \) and \( \alpha \), we have:

\[
\lim_{d \to 0} E_{\mu, \sigma}[N_i/C] = 1, \ i = 0, 1, 2 \ [\text{Asymptotic First-Order Efficiency}], \tag{3.1.5}
\]

with \( C \) defined by \((2.1.3)\).

**Proof.** We prove the theorem along the lines of Chow and Robbins (1965). Let us denote \( y_{n,i} = U_{n,i}/\sigma^2 \), and we know from Lemma 3.1.2 that \( \lim_{n \to \infty} y_{n,i} = 1 \) w.p.1. Denote \( f(n) = n \) so that \( \lim_{n \to \infty} f(n) = \infty \) and \( \lim_{n \to \infty} f(n)/f(n-1) = 1 \). Then, the stopping rule \((3.1.1)\) would match the stopping rule of Chow and Robbins (1965). Using Lemma 3.1.3, we claim:

\[
E_{\mu, \sigma}[^{\sup_n y_{n,i}}] < \infty \text{ which immediately shows (3.1.5).} \ ■
\]

**Theorem 3.1.3.** For the stopping time \( N_i \) and the proposed fixed-width confidence interval \( J_{N_i} \) defined by the purely sequential estimation strategy \((3.1.1)-(3.1.2)\), for all fixed \( \mu, \sigma, \) and \( \alpha \), we have:

\[
\lim_{d \to 0} P_{\mu, \sigma}\{\mu \in J_{N_i}\} = 1 - \alpha, \ i = 0, 1, 2 \ [\text{Asymptotic Consistency}], \tag{3.1.6}
\]

**Proof.** First, it is obvious that \( \lim_{d \to 0} N_i/C = 1 \) with \( C \) defined by \((2.1.3)\). Next, we express:

\[
P_{\mu, \sigma}\{\mu \in J_{N_i}\} = E_{\mu, \sigma}\left[2\Phi\left(N_i^{1/2}d/\sigma\right) - 1\right] \to 1 - \alpha,
\]
as \(d \to 0\), which follows from the dominated convergence theorem.

Theorems 3.1.2-3.1.3 give us reasonable assurances that the terminal sample sizes on an average and the confidence coefficients associated with the constructed confidence intervals will hover around \(C\) and \(1 - \alpha\) respectively when \(C\) (or \(d\)) is large (small). Small and moderate sample size performances of the purely sequential estimation strategies \((N_i, J_{N_i})\) from (3.1.1)-(3.1.2) will be summarized shortly via computer simulations.

3.1.3. Limited Robustness Properties of the Purely Sequential Methodology (3.1.1)-(3.1.2)

Here, we briefly address some of the issues surrounding potentially limited robustness properties associated with the GMD-based or MAD-based purely sequential estimation strategies (3.1.1)-(3.1.2) with \(i = 1, 2\) when compared with the existing and customary strategy (3.1.1)-(3.1.2) with \(i = 0\).

We deal with a situation when the normality of the parent population is satisfied, but one may encounter some outliers during the sequential process of gathering observations. Since an outlier is expected to be either extremely large or extremely small, it is reasonable to presume that the true population variance, say, \(\sigma^2\) would be pulled upward compared with \(\sigma^2\), the assumed variance of the population without outliers.

As a consequence, if we fix the significance level \(1 - \alpha\) and the confidence interval’s width \(2d\), the true behind-the-scene optimal fixed sample size \(C\) will go up. That is, if \(C'\) (or \(C\)) corresponds to the optimal fixed sample size when the population produces some (or no) outliers, then it may be reasonable to expect \(C' > C\).

For convenience, we simulate data from a population that will be expected to produce
some outliers by generating observations from a mixture normal distribution specified by:

\[(1 - \epsilon)N(\mu, \sigma_1^2) + \epsilon N(\mu, \sigma_2^2)\] where \(\sigma_2 > \sigma_1\) and \(0 < \epsilon < 1\) is small, \(\text{(3.1.7)}\)

that is, we view the normal population \(N(\mu, \sigma_1^2)\) as mildly contaminated allowing it to possibly generate some outliers. Next, we summarize some interesting conclusions.

**Theorem 3.1.4.** Under the model \((3.1.7)\), with \(U_{n;i}^2\) defined in \((2.3.2)-(2.3.3)\), \(i = 0, 1, 2\), and for all fixed \((\mu, \sigma_1, \sigma_2) \in R \times R^+ \times R^+\), we have:

\[\sigma_1^2 < E_{\mu,\sigma_1,\sigma_2}[U_{n;0}^2] < E_{\mu,\sigma_1,\sigma_2}[U_{n;1}^2] < E_{\mu,\sigma_1,\sigma_2}[U_{n;2}^2] < \sigma_2^2,\] \(\text{(3.1.8)}\)

as \(n \to \infty\).

**Proof.** Note that as \(n \to \infty\), we first have:

\[E_{\mu,\sigma_1,\sigma_2}[U_{n;0}^2] = (1 - \epsilon)\sigma_1^2 + \epsilon \sigma_2^2.\] \(\text{(3.1.9)}\)

Also, we obtain:

\[\begin{align*}
U_{n;1}^2 &\xrightarrow{P_{\mu,\sigma_1,\sigma_2}} \frac{\pi}{4} E_{\mu,\sigma_1,\sigma_2}^2[|X_1 - X_2|] = \left[(1 - \epsilon)^2 \sigma_1 + \epsilon (1 - \epsilon) \sqrt{2(\sigma_1^2 + \sigma_2^2) + \epsilon^2 \sigma_2}\right]^2; \\
U_{n;2}^2 &\xrightarrow{P_{\mu,\sigma_1,\sigma_2}} \frac{\pi}{2} E_{\mu,\sigma_1,\sigma_2}^2[|X_1 - \mu|] = [(1 - \epsilon) \sigma_1 + \epsilon \sigma_2]^2.
\end{align*}\] \(\text{(3.1.10)}\)

Next, it follows that

\[\left[(1 - \epsilon)^2 \sigma_1 + \epsilon (1 - \epsilon) \sqrt{2(\sigma_1^2 + \sigma_2^2) + \epsilon^2 \sigma_2}\right]^2 < (1 - \epsilon)^2 \sigma_1^2 + \epsilon (1 - \epsilon) (\sigma_1^2 + \sigma_2^2) + \epsilon^2 \sigma_2^2 = (1 - \epsilon)\sigma_1^2 + \epsilon \sigma_2^2,\] \(\text{(3.1.11)}\)
and also,

\[
\left[ (1 - \epsilon)^2 \sigma_1 + \epsilon (1 - \epsilon) \sqrt{2(\sigma_1^2 + \sigma_2^2) + \epsilon^2 \sigma_2^2} \right]^2 \\
> \left[ (1 - \epsilon)^2 \sigma_1 + \epsilon (1 - \epsilon) (\sigma_1 + \sigma_2) + \epsilon^2 \sigma_2 \right]^2 = [(1 - \epsilon) \sigma_1 + \epsilon \sigma_2]^2. \quad (3.1.12)
\]

Based on the proof of Lemma 3.1.3, the dominated convergence theorem could be applied on \( U_{n,1}^2 \) and \( U_{n,2}^2 \). Thus, combining (3.1.9)-(3.1.12), the desired result follows.

Now, we may summarize by writing that mild robustness of the procedure (3.1.1)-(3.1.2) is felt through the spirit of taking into account what our Theorem 3.1.4 has to say. It indicates that when a population along the line of (3.1.7) is more prone to produce outliers, but yet we continue to assume that the population behaves like normal with variance \( \sigma_1^2 \), the estimators \( U_{n,1}^2, U_{n,2}^2, \) and \( U_{n,0}^2 \) tend to overestimate \( \sigma_1^2 \). As a result, they are no longer unbiased or even consistent for \( \sigma_1^2 \). Additionally, Theorem 3.1.4 also indicates that under procedures (3.1.1)-(3.1.2) the final sample sizes \( N_1, N_2 \) based on \( U_{n,1}^2, U_{n,2}^2 \) on an average could asymptotically fall below the average \( N_0 \) based on \( U_{n,0}^2 \). In other words, \( N_1, N_2 \) could be expected to stay nearer to the correspondingly presumed optimal fixed sample size, \( C^* = a^2 \sigma_1^2 / d^2 \), on an average than \( N_0 \) would.

### 3.1.4. Simulations and Data Analysis

In this section, we summarize performances of the purely sequential FCIE strategies \((N_i, J_{N_i}), i = 0, 1, 2\) from (3.1.1)-(3.1.2) respectively when sample sizes were small (\( \leq 50 \)), moderate (100) or large (200). Throughout this section, we use the following system codes in (3.1.13) for identifying each specific methodology under implementation. We present summary performances in both normal cases and mixture normal cases (see 3.1.7) obtained
via simulations.

\[ i = 0 \quad : \text{Sample variance-based methodology (3.1.1)-(3.1.2)} \]
\[ i = 1 \quad : \text{GMD-based methodology (3.1.1)-(3.1.2)} \] (3.1.13)
\[ i = 2 \quad : \text{MAD-based methodology (3.1.1)-(3.1.2)} \]

3.1.4.1. Samples from a Normal Population

We generated pseudo random samples from a \( N(5, 4) \) population, that is, we had fixed \( \mu = 5 \) and \( \sigma = 2 \). We also fixed \( \alpha = 0.01, 0.05, 0.10, m = 10, 15, 20 \), the optimal fixed sample size \( C \) over a wide range of values including 50, 100, 200, and then determined the values of \( d \) accordingly by referring to (2.1.3). Table 3.1 represents this situation. One may however critique that we should have alternatively first fixed \( d(= 0.2, 0.4, 0.6) \) and determined \( C \) accordingly. Table 3.2 represents that situation.

In Tables 3.1-3.2, we respectively present selected summaries in the two situations that we have mentioned when \( \alpha = 0.05 \) and \( m = 10, 15 \) for brevity alone even though we ran simulations with other values of \( C, \alpha, \) and \( m \).

Using \( T (= 1000, \text{say}) \) independent replications or runs, we estimated the average final sample size \( \bar{n} \) with its estimated standard error \( s(\bar{n}) \), the minimum \( \min(n) \), the lower quartile \( Q_L(n) \), the median \( \text{med}(n) \), the upper quartile \( Q_U(n) \), and the maximum \( \max(n) \) obtained from 1000 observed sample sizes. We also looked at the estimated coverage probability \( \bar{p} \) with its estimated standard error \( s(\bar{p}) \) along with the observed values of both \( \bar{n}/C \) and \( \bar{n} - C \).

More specifically, in Tables 3.1-3.2, as well as in other tables which will follow, we have used the following notation:
\[ n_t : \text{sample size in } t^{\text{th}} \text{ run}; \]

\[ \pi = T^{-1} \sum_{t=1}^{T} n_t : \text{should estimate } C; \]

\[ s(\pi) = \left\{ (T^2 - T)^{-1} \sum_{t=1}^{T} (n_t - \pi)^2 \right\}^{1/2} : \text{estimated standard error of } \pi; \]

\[ \bar{x}_{nt} : \text{sample mean in } t^{\text{th}} \text{ run;} \]

\[ p_t : 1 \text{ if interval } J_{nt} \text{ covers } \mu \text{ (or 0 otherwise)} \]

\[ p = T^{-1} \sum_{t=1}^{T} p_t : \text{should compare with } 1 - \alpha; \]

\[ s(\bar{p}) = \left\{ T^{-1} \bar{p} (1 - \bar{p}) \right\}^{1/2} : \text{estimated standard error of } \bar{p}. \]

(3.1.14)

As a typical case, for example, we prefixed the optimal sample size \( C = 50 \) (small), \( 100 \) (moderate), and \( 200 \) (large) with different pilot sample sizes \( m = 10 \) and \( 15 \) respectively when \( \alpha = 0.05 \). Thus, with small, moderate and large optimal fixed sample sizes, the corresponding fixed-width of the confidence intervals was \( d = 0.5544, 0.3920, \) and \( 0.2772 \). Overall performances did not differ over other choices of \( C, \alpha \) or \( m \). In summary, we discover that the procedures (3.1.1)-(3.1.2) based on either GMD or MAD as the estimator of \( \sigma^2 \) tend to perform as well as the procedure based on the sample variance.

Looking at column 4, it is fair to say that the average sample sizes (\( \pi \)) derived were close to each other, all staying a little smaller than \( C \) no matter whether \( C \) was small, moderate, or large. From column 9, we can see that the first-order term \( \pi/C \) is close to 1, whereas larger the sample sizes were, the closer was \( \pi/C \) to 1.

A similar conclusion could be in order while looking at the estimated second-order values \( \pi - C \) which look rather close to the theoretical asymptotic value \(-1.18\) (see Mukhopadhyay and de silva 2009, eq. 6.2.34, p. 119) when \( i = 0 \). In some cases, the average sample sizes
under procedure (3.1.1)-(3.1.2) with \( i = 0 \) turned out to be smaller than those of other two procedures \((i = 1, 2)\), however the average sample sizes of GMD-based or MAD-based methodologies were often bit smaller.

Next, we focus on columns 5–7, the five number summary \((\min(n), \text{\( Q_L(n) \)}, \text{\( \text{med}(n) \)}, \text{\( Q_U(n) \)}, \text{and} \max(n))\) from 1000 observed \( n \) values in each row. From Tables 3.1-3.2, one may sense that these quantities reflect a right-skewness property of the stopping times, while at the same time, the difference between \( \bar{n} \) and \( \text{med}(n) \) was largely within one observation. For small sample sizes, the MAD-based procedure seemed to be more right-skewed when \( m \) became larger. For moderate and large sample sizes, while the values of \( Q_L(n) \), \( \text{med}(n) \), and \( Q_U(n) \) look close to each other, we noticed some apparent differences in the magnitude of \( \max(n) - \min(n) \) when different procedures were implemented.

Now let us attend to column 8, that is \( \bar{p} \), the estimated coverage probability, and \( s(\bar{p}) \). In a large majority of cases, the interval of \([\bar{p} \pm 2s(\bar{p})]\) did cover our nominal target \( 1 - \alpha \), however sometimes \([\bar{p} \pm 2s(\bar{p})]\) did not cover \( 1 - \alpha \). This feature should not be surprising, since Theorem 3.1.3 shows only asymptotic consistency property.

In summary, the purely sequential procedure (3.1.1)-(3.1.2) has performed remarkably well under the normality assumption for the parent population. We found very little to no significant difference in the overall performances between the existing purely sequential procedure (3.1.1)-(3.1.2) with \( i = 0 \) and our newly developed purely sequential procedures (3.1.1)-(3.1.2) with \( i = 1, 2 \) based on GMD or MAD across the large range of choices of \( m, C \), and \( \alpha \). Recall that Tables 3.1-3.2 highlight selected summaries.

3.1.4.2. Samples from an \( \epsilon \)-Contaminated Normal Population (3.1.7)

Next, we drew pseudo random samples from a mixture normal population (3.1.7),
namely, a population distribution given by \((1 - \epsilon)N(5, 4) + \epsilon N(5, 9)\) to investigate performances of the purely sequential fixed-width confidence interval estimation strategies \((N_i, J_{N_i})\) from (3.1.1)-(3.1.2) with \(i = 0, 1, 2\) respectively when suspect outliers may occur. We fixed \(\alpha = 0.01, 0.05, 0.10, m = 10, 15, 20, C = 50, 100, 200\), and considered \(\epsilon = 0.1\) (10\% contamination), 0.2 (20\% contamination).

We implemented the purely sequential procedures similarly to what we had explained in connection with Table 3.1. Tables 3.3-3.4 show findings from 1000 runs when \(m = 10, 15, C = 50, 100, 200, \alpha = 0.05,\) and \(\epsilon = 0.10, 0.20\) in the spirit of Table 3.1. In our larger simulation runs, we had additionally included the choices (i) \(m = 20, C = 50, 100, 200, \alpha = 0.10, 0.01, \epsilon = 0.10, 0.20\) and (ii) \(m = 10, 15, 20, d = 0.6, 0.4, 0.2, \alpha = 0.10, 0.01, \epsilon = 0.1, 0.2\). However, we exhibit only a brief summary especially since we found very little to no significant difference in the overall performances. In particular, we include no tables here in the spirit of Table 3.2.

At the same time, one will find a number of additional entries in Tables 3.3-3.4 compared with those shown in Tables 3.1-3.2. Let us explain the extra notation.

\[
x_{tj} : j^{th} \text{ observation in } t^{th} \text{ run;}
\]

\[
\gamma_{tj} : = 1 \text{ if } x_{tj} \text{ came from } N(5, 4)
\]

or 0 if \(x_{tj}\) was from \(N(5, 9)\);

\[
\gamma_t = \sum_{j=1}^{n_t} \gamma_{tj} : \text{ proportion of observations sampled from } N(5, 4) \text{ in } t^{th} \text{ run;}
\]

\[
\bar{\gamma} = \frac{\sum_{t=1}^{T} \gamma_t}{T} : \text{ should estimate } \gamma(= 1 - \epsilon);
\]

\[
s(\bar{\gamma}) = \left\{ T^{-1} \bar{\gamma} (1 - \bar{\gamma}) \right\}^{1/2} : \text{ estimated standard error of } \bar{\gamma};
\]
KS_t : = 1 if K-S test’s p-value $\geq 0.05$ in $t^{th}$ run or 0 otherwise; 

$$KS = \sum_{t=1}^{T} KS_t : \text{proportion of data from } T \text{ runs that passed 5\% level K-S normality test.}$$  

We denoted $\gamma = 1 - \epsilon$ and estimated the average sampling from $N(5, 4)$ population proportion $\overline{\gamma}$ and its standard error $s(\overline{\gamma})$. We observe column 10 that $\overline{\gamma}$ is close to the target $\gamma = 90\%-80\%$ in Table 3.3(3.4) while its standard error $s(\overline{\gamma})$ hovers around 0.01. We also used Kolmogorov-Smirnov test to check the normality of our samples within each run and we recorded in column 11 the proportion $KS$, that passed the Kolmogorov-Smirnov test at 0.05 significance level among all 1000 runs. From column 11 in Tables 3.3-3.4, we get an impression that our observed data looked like normal between 89.7\% to 95\% of the time among 1000 runs even though in the background the data came from (3.1.7)!

Again, we gather data purely sequentially from a population described by (3.1.7), but pretend that the population is indeed $N(\mu = 5, \sigma^2 = 4)$ for all practical purpose. In column 1 we show the $C$-values (in the top) computed using (2.1.3) replacing $\sigma^2 = \sigma^2_I = 4$. Obviously, this $C$ is a rather "fake" optimal fixed sample size, and this is clear from the fact that the $\overline{\pi}$ values appear off from such $C$-values shown in the top in column 1. But, in column 1, we also show the bold $C$-values computed from (2.1.3) using the correct $\sigma^2$ value from the population (3.1.7). Clearly, the $\overline{\pi}$ values appear much closer to the bold $C$-values shown in column 1.

3.1.4.3. Some Overall Sentiments

From Tables 3.3-3.4 and many other simulations that we had run, we may summarize our sentiments as follows when a population may be reasonably treated as a normal universe for
practical purposes (as validated by Kolmogorov-Smirnov test), but it is mildly contaminated behind the scene:

(a) All three methodologies oversampled which is clear by comparing the $\bar{n}$ values with $C$-values or $C$-values.

(b) Pilot size $m$ has not impacted overall performances in a significant way.

(c) Average sample sizes of GMD-based and MAD-based purely sequential procedures (3.1.1)-(3.1.2) with $i = 1, 2$ were generally smaller than those associated with the sample variance-based procedure (3.1.1)-(3.1.2) with $i = 0$. We interpret this feature as a reasonable validation of mild robustness of the procedures (3.1.1)-(3.1.2) with $i = 1, 2$ over the procedure (3.1.1)-(3.1.2) with $i = 0$.

(d) MAD-based purely sequential procedure generally came up with a smaller average sample size than that associated with GMD-based procedure.

(e) Average sample sizes from the customary procedure (3.1.1)-(3.1.2) with $i = 0$ came bit closer to the prefixed $C$-values. This is also reflected in the feature that first-order terms $(\pi/C)$ for the purely sequential procedures (3.1.1)-(3.1.2) with $i = 1, 2$ were bit lower than those of the procedure (3.1.1)-(3.1.2) with $i = 0$.

(f) From columns 5–7, we feel that the distributions of all three stopping times were skewed to the right which is validated by the feature that med$(n)$ turned out to be a little larger than $\bar{n}$, but they were largely within one observation of each other. Also, med$(n)$ seen from procedures (3.1.1)-(3.1.2) with $i = 1, 2$ kept being smaller than those derived from procedure (3.1.1)-(3.1.2) with $i = 0$.

(g) MAD-based procedure gave the smallest med$(n)$ and $\bar{n}$ values. Similar conclusions generally hold for the other two quartiles ($Q_L(n), Q_U(n)$).
Nearly all coverage probability estimates ($\bar{p}$) fell slightly below the target confidence level of 0.95. We suspect that this may be partly due to contamination.

By combining our findings from Tables 3.1-3.4 with those obtained from numerous other sets of extensive simulations, we empirically feel confident that our newly developed GMD-based or MAD-based methodologies (3.1.1)-(3.1.2) with $i = 1, 2$ are more robust for practical purposes when we compare them with the sample variance-based methodology (3.1.1)-(3.1.2) with $i = 0$. At the same time, however, the GMD-based or MAD-based methodologies from (3.1.1)-(3.1.2) with $i = 1, 2$ are not very sharply divided. Hence, we suggest implementing either GMD-based or MAD-based methodologies from (3.1.1)-(3.1.2) with $i = 1, 2$ with confidence when sampling from a normal universe, especially when suspect outliers may be expected.

### 3.2. Minimum Risk Point Estimation

In this section, we estimate the unknown $\sigma$ unbiasedly by the estimators from (2.3.4), respectively, based on the sample standard deviation, GMD, and MAD. The corresponding purely sequential procedures will be proposed to construct minimum risk point estimators for the unknown mean $\mu$ under the formulation (2.2.1)-(2.2.3). Again, we begin with pilot data $X_1, \ldots, X_m$ of size $m$, $m \geq 2$, and record one additional observation at-a-time successively as needed until we stop according to the following stopping rule:

$$N_i \equiv N_i(c) = \inf\{n \geq m : n \geq \sqrt{\Lambda/c} \left( V_{n;i} + n^{-\lambda} \right) \},$$  \hspace{1cm} (3.2.1)

where $\lambda(>\frac{1}{2})$ is held fixed and $\sigma$ is replaced with $V_{n;i}$, $i = 0, 1, 2$, in defining the boundary condition as we estimate $n^* \equiv n^*(c)$ from (2.2.3). Observe that the index $i$ corresponds
to the sample standard deviation-based, GMD-based or the MAD-based methodology with

\[ V_{n;i}, i = 0, 1, 2, \] defined via (2.3.1) and (2.3.4).

The stopping rule (3.2.1) looks much like the representations in Chow and Robbins (1965) and Ghosh et al. (1997, equation 9.2.30, p. 274). But, in comparison with (3.1.1), the present boundary condition in (3.2.1) looks slightly different because it includes an additional term, \( n^{-\lambda} \). This is done in order to (i) avoid very early stopping and (ii) facilitate derivations of various convergences with “desired” rates.

With regard to implementation of (3.2.1), for \( i = 0, 1, 2 \), if \( m \geq \sqrt{\frac{A}{C}} (V_{m;i} + m^{-\lambda}) \) is satisfied, we do not take any additional observation and the sample size is \( N_i = m \). Otherwise, we record one more observation and obtain updated \( V_{m+1;i} \) to check with the stopping rule (3.2.1). We terminate sampling at the first time \( N_i = n(\geq m) \) such that \( n \geq \sqrt{\frac{A}{C}} (V_{n;i} + n^{-\lambda}) \) occurs. Finally, having observed the dataset

\[ \{N_i, X_1, ..., X_m, X_{m+1}, ..., X_{N_i}\}, \]

we construct the minimum risk point estimator for \( \mu \) as follows:

\[ \overline{X}_{N_i} \equiv N_i^{-1} \sum_{j=1}^{N_i} X_j, i = 0, 1, 2. \] (3.2.2)

Clearly, \( P_{\mu,\sigma}\{N_i < \infty\} = 1 \) and \( N_i \uparrow \infty \) w.p.1 as \( c \downarrow 0, i = 0, 1, 2. \)

3.2.1. Asymptotic First-Order Properties of the Purely Sequential Methodology (3.2.1)-(3.2.2)

In the spirit of our previous Lemma 3.1.4 and Theorem 3.1.1, modified very mildly, we first summarize the following properties for the purely sequential estimation strategy
(3.2.1)-(3.2.2):

(i) \( E_{\mu,\sigma}[N_i] < \infty, V_{\mu,\sigma}[N_i] < \infty; \)

(ii) \( E_{\mu,\sigma}[\bar{X}_{N_i}] = \mu, V_{\mu,\sigma}[\bar{X}_{N_i}] = \sigma^2 E_{\mu,\sigma}(N_{i}^{-1}); \) \hspace{1cm} (3.2.3)

(iii) \( \sqrt{N_i} (\bar{X}_{N_i} - \mu) / \sigma \sim N(0,1); \)

for all fixed \( \mu, \sigma, A, c, \) and \( i = 0, 1, 2. \)

In the present situation, upon termination according to (3.2.1), we can express the risk associated with the purely sequential estimation strategy \((N_i, \bar{X}_{N_i})\) defined via (3.2.1)-(3.2.2) as follows:

\[
R_{N_i}(c) \equiv E_{\mu,\sigma}[L_{N_i} (\mu, \bar{X}_{N_i})] = A \sigma^2 E_{\mu,\sigma}[N_{i}^{-1}] + c E_{\mu,\sigma}[N_i], i = 0, 1, 2, \hspace{1cm} (3.2.4)
\]

since \( \bar{X}_n \) and \( I\{N_i = n\} \) are independent for each fixed \( n(\geq m) \). The minimum fixed-sample-size risk can be expressed as:

\[
R_{n^*}(c) \equiv A \sigma^2 n^{*-1} + cn^* = 2cn^*, \hspace{1cm} (3.2.5)
\]

with \( n^* \equiv n^*(c) \) from (2.2.3).

Robbins (1959) and Starr (1966) formulated two crucial notions, namely, the risk efficiency and regret originally developed in the context of comparing \( R_{N_0}(c) \) associated with (3.2.4) and \( R_{n^*}(c) \). We implement those notions for comparing \( R_{N_i}(c) \) and \( R_{n^*}(c) \):

\[
\text{Risk Efficiency: } \xi_i(c) \equiv \frac{R_{N_i}(c)}{R_{n^*}(c)} = \frac{1}{2} E_{\mu,\sigma}[N_i/n^*] + \frac{1}{2} E_{\mu,\sigma}[n^*/N_i]; \hspace{1cm} (3.2.6)
\]

\[
\text{Regret: } \omega_i(c) \equiv R_{N_i}(c) - R_{n^*}(c) = c E_{\mu,\sigma}[(N_i - n^*)^2/N_i],
\]

30
for $i = 0, 1, 2$.

Clearly, the following property will also hold:

$$ N_i/n^* \xrightarrow{P_{\mu,\sigma}} 1 \text{ as } c \to 0, i = 0, 1, 2, $$

(3.2.7)

which results from the customary basic inequality and the facts that $N_i \to \infty$ and $\bar{X}_{N_i} \to \mu$ w.p.1($P_{\mu,\sigma}$) as $c \to 0$.

Next, in view of the first part of (3.2.7) and then slightly improvising the proof shown in the case of our earlier Theorem 3.1.2, we can immediately claim the following property:

**Theorem 3.2.1.** For the stopping time $N_i$ defined by the purely sequential estimation strategy (3.2.1)-(3.2.2), for all fixed $\mu, \sigma$, and $A$, we have:

$$ \lim_{c \to 0} E_{\mu,\sigma}[N_i/n^*] = 1, i = 0, 1, 2 \text{ [Asymptotic First-Order Efficiency]}, \quad (3.2.8) $$

with $n^* \equiv n^*(c)$ from (2.2.3).

Next, we show the following lemma, which will lead to the asymptotic first-order risk efficiency of the methodology (3.2.1)-(3.2.2).

**Lemma 3.2.1.** For the stopping time $N_i$ defined by the purely sequential estimation strategy (3.2.1)-(3.2.2), for all fixed $\mu, \sigma$, and $A$, we have: with any arbitrary $0 < \eta < 1$ and $r \geq 2$,

$$ P_{\mu,\sigma}\{N_i \leq \eta n^*\} = O\left(n^{*r/(2(1+\lambda))}\right), i = 0, 1, 2. \quad (3.2.9) $$

**Proof:** Let $|w|$ denote the largest integer that is smaller ($<$) than $w$ and we define:
\[
\begin{align*}
n_{1c} &= \left\lfloor \frac{(A/c)^{1/(1+\lambda)}}{2(1+\lambda)} \right\rfloor = O(c^{-\frac{1}{2(1+\lambda)}}) \text{ and } n_{2c} = \eta n^* = \eta \sigma \sqrt{A/c}. \tag{3.2.10}
\end{align*}
\]

From the definition of \( N_i \) in (3.2.1), it should be clear that \( N_i \geq n_{1c} \) w.p.1(\( P_{\mu,\sigma} \), \( i = 0, 1, 2 \).

Next, in view of (3.2.7), it is obvious that \( P_{\mu,\sigma}\{N_i \leq \eta n^*\} \) ought to converge to zero as \( c \to 0 \). Hence, we set out to obtain the rate at which \( P_{\mu,\sigma}\{N_i \leq \eta n^*\} \) may converge to zero for small \( c \).

**Case** \( i = 1 \):

We write for small \( c \), with \( k_1 = \frac{2(1-\eta)}{\sqrt{n}} (> 0) \):

\[
P_{\mu,\sigma}\{N_1 \leq \eta n^*\}
\leq P_{\mu,\sigma} \left\{ \frac{1}{\sqrt{n}} G_n \leq \eta \sigma \right\} \text{ for some } n \text{ such that } n_{1c} \leq n \leq n_{2c} \\
\leq P_{\mu,\sigma} \left\{ \left| G_n - \frac{2}{\sqrt{n}} \sigma \right| \geq k_1 \sigma \right\} \text{ for some } n \text{ such that } n_{1c} \leq n \leq n_{2c} \\
\leq P_{\mu,\sigma} \left\{ \max_{n_{1c} \leq n \leq n_{2c}} \left| G_n - \frac{2}{\sqrt{n}} \sigma \right| \geq k_1 \sigma \right\}. \tag{3.2.11}
\]

We know that \( G_n \) is a \( U \)-statistic with kernel \(|X_{1i} - X_{2j}|\) and \( E_{\mu,\sigma}[G_n] = \frac{2}{\sqrt{n}} \sigma \). As a result, \( \{G_n; n \geq m\} \) is a reverse martingale adapted to the \( \sigma \)-fields \( \mathcal{F}_n = \sigma \{G_n, G_{n+1}, \ldots\} \), and so is also \( \{G_n - \frac{2}{\sqrt{n}} \sigma; n \geq m\} \). Note the fact that all positive moments of \(|X_{1i} - X_{2j}|\) are finite. Hence, by Kolmogorov’s inequality for reverse martingales, the lemma from Sen and Ghosh (1981), in view of (3.2.10) we obtain from (3.2.11) with \( r \geq 2 \):

\[
P_{\mu,\sigma}\{N_1 \leq \eta n^*\} \leq (k_1 \sigma)^{-r} E_{\mu,\sigma} \left[ \left| G_{n_{1c}} - \frac{2}{\sqrt{n}} \sigma \right|^r \right] = O \left( n_{1c}^{-r/2} \right) = O \left( n^{1-r/(2(1+\lambda))} \right). \tag{3.2.12}
\]

See also Ghosh et al. (1997, Lemma 9.2.3, pp. 275-276).
Case \( i = 0 \):

Observing that in (2.3.4) \( \frac{\Gamma\left(\frac{n+1}{2}\right)\sqrt{n-1}}{\sqrt{2\pi}} \geq 1 \) as \( E_{\mu,\sigma}[S_n] \leq \sqrt{E_{\mu,\sigma}[S^2_n]} = \sigma \), we again write for small \( c \), with \( k_0 = 1 - \eta (0) \):

\[
P_{\mu,\sigma}\{N_0 \leq \eta n^*\} \\
\leq P_{\mu,\sigma}\{S_n \leq \eta \sigma \text{ for some } n \text{ such that } n_1c \leq n \leq n_2c\} \tag{3.2.13} \\
\leq P_{\mu,\sigma}\{|S_n - \sigma| \geq k_0\sigma, \text{ for some } n \text{ such that } n_1c \leq n \leq n_2c\} \\
\leq P_{\mu,\sigma}\left\{\max_{n_1c \leq n \leq n_2c} |S^2_n - \sigma^2| \geq k_0\sigma^2\right\}.
\]

Note that \( S^2_n \) is a common \( U \)-statistic with all positive moments finite, with \( r \geq 2 \), we claim:

\[
P_{\mu,\sigma}\{N_0 \leq \eta n^*\} = O\left(n_i^{-r/2}\right) = O\left(n_i^{* - r/(2(1+\lambda))}\right), \tag{3.2.14}
\]

in the same way we had concluded (3.2.12) from (3.2.11).

Case \( i = 2 \):

We note that

\[
M_n = n^{-1}\sum_{i=1}^{n} |X_i - \mu| - |\bar{X}_n - \mu| = M_{1n} - M_{2n}, \text{ say,}
\]

\( \text{w.p.1}(P_{\mu,\sigma}) \). Similarly, observing that in (2.3.4) \( \sqrt{\frac{\pi n}{2(n-1)}} \geq \sqrt{\frac{\pi}{2}} \), we again write for small \( c \), with \( k_2 = \frac{2(1-\eta)}{\sqrt{2\pi}} \):

\[
P_{\mu,\sigma}\{N_2 \leq \eta n^*\} \\
\leq P_{\mu,\sigma}\left\{\sqrt{\frac{\pi}{2}}M_n \leq \eta \sigma, \text{ for some } n \text{ such that } n_1c \leq n \leq n_2c\right\} \\
\leq P_{\mu,\sigma}\left\{|M_{1n} - \sqrt{\frac{\pi}{2}}\sigma| + |M_{2n}| \geq k_2\sigma, \text{ for some } n \text{ such that } n_1c \leq n \leq n_2c\right\}
\]
\[
\leq P_{\mu,\sigma} \left\{ \max_{n_1 \leq n \leq n_2} \left| M_{1n} - \frac{\sqrt{2}}{\pi} \sigma \right| \geq \frac{1}{2} k_2 \sigma \right\} + P_{\mu,\sigma} \left\{ \max_{n_1 \leq n \leq n_2} |M_{2n}| \geq \frac{1}{2} k_2 \sigma \right\}.
\]

(3.2.15)

Since \(M_{1n}\) and \(X_n\) are both \(U\)-statistics with all positive moments finite, with \(r \geq 2\), we claim:

\[
P_{\mu,\sigma} \{N_2 \leq \eta n^*\} = O\left(n^{-r/2}\right) = O\left(n^{s-r/(2(1+\lambda))}\right).
\]

(3.2.16)

The proof of the lemma is complete that \(P_{\mu,\sigma} \{N_i \leq \eta n^*\} = O\left(n^{s-r/(2(1+\lambda))}\right)\) holds for \(i = 0, 1, 2\).

Now we are in a position to state and prove the main result in this section.

**Theorem 3.2.2.** For the stopping time \(N_i\) defined by the purely sequential estimation strategy (3.2.1)-(3.2.2), for all fixed \(\mu, \sigma\), and \(A\), we have:

\[
\lim_{c \to 0} \xi_i(c) = 1, \ i = 0, 1, 2 \ [Asymptotic \ First-Order \ Risk \ Efficiency],
\]

(3.2.17)

where the term \(\xi_i(c)\) comes from (3.2.6).

**Proof:** In view of (3.2.6) and (3.2.8), it will suffice to prove the following result:

\[
\lim_{c \to 0} E_{\mu,\sigma}[n^*/N_i] = 1, \ i = 0, 1, 2.
\]

(3.2.18)

With \(\eta = \frac{1}{2}\) in Lemma 3.2.1, we express:

\[
E_{\mu,\sigma}[n^*/N_i] = E_{\mu,\sigma}\left[\frac{n^*}{N_i} I\left(N_i \leq \frac{1}{2} n^*\right)\right] + E_{\mu,\sigma}\left[\frac{n^*}{N_i} I\left(N_i > \frac{1}{2} n^*\right)\right]
\]

\[
= E_{\mu,\sigma}[W_1] + E_{\mu,\sigma}[W_2], \text{ say, } i = 0, 1, 2.
\]

(3.2.19)

But, \(W_2\) is bounded and hence \(W_2\) is uniformly integrable. Also, \(W_2 \xrightarrow{P} 1\) as \(c \to 0\) so
that \( \lim_{c \to 0} E_{\mu,\sigma} [W_2] = 1 \). In view of (3.2.9) and (3.2.10), we can write:

\[
E_{\mu,\sigma} [W_1] \leq E_{\mu,\sigma} \left[ \frac{n^*}{n_1} I \left( N_i \leq \frac{1}{2} n^* \right) \right] = O \left( n^* \right) O \left( n^{* - 1/(1+\lambda)} \right) O \left( n^{* - r/(2(1+\lambda))} \right)
\]

\[
= O \left( n^{* \frac{1}{1+\lambda} - \frac{r}{2(1+\lambda)}} \right) = O \left( n^{* \frac{2\lambda - r}{2(1+\lambda)}} \right) \to 0,
\]

as long as we pick some \( r > \max\{2,2\lambda\} \). Thus, from (3.2.18), it follows that (3.2.17) is justified.

\[\blacksquare\]

### 3.2.2. Simulations and Data Analysis

Along the lines of what we did in Section 3.1.4, we had analogously implemented purely sequential minimum risk point estimation strategies \((N_i, \bar{X}_{N_i})\), \(i = 0,1,2\) from (3.2.1)-(3.2.2) under both normal cases and mixture-normal cases (3.1.7). In all tables to follow, we continue to use the index \(i\) along the lines of (3.1.13), but define more precisely as follows:

\[
i = 0 \quad : \text{Sample standard deviation-based methodology (3.2.1)-(3.2.2) with } \lambda = 2
\]

\[
i = 1 \quad : \text{GMD-based methodology (3.2.1)-(3.2.2) with } \lambda = 2
\]

\[
i = 2 \quad : \text{MAD-based methodology (3.2.1)-(3.2.2) with } \lambda = 2
\]

(3.2.20)

as well as other notation used in the spirit of (3.1.14) obtained on the basis of \(T (= 1000, \text{say})\) replications:

\[
n_t \quad : \text{sample size in } t^{th} \text{ run;}
\]

\[
\bar{n} = \frac{1}{T} \sum_{t=1}^{T} n_t \quad : \text{should estimate } n^*;
\]

\[
s(\bar{n}) \quad : \text{estimated standard error (s.e.) of } \bar{n};
\]

\[
R_{n_t} = As^2_{n_t} n_t^{-1} + cn_t \quad : \text{the method of calculation of risk}
\]

using sample variance, \(s^2_{n_t}\), in \(t^{th}\) run;

35
\[ \bar{R}_n = \frac{1}{T} \sum_{t=1}^{T} R_{n_t} \] should estimate \( R_{n^*}(c) \);

\[ s(\bar{R}_n) \] estimated s.e. of \( \bar{R}_n \);

\[ \hat{\xi} = \frac{\bar{R}_n}{R_{n^*}(c)} \] should estimate \( \xi_i(c), i = 0, 1, 2 \)

\[ \hat{\omega} = c \frac{1}{T} \sum_{t=1}^{T} \left( \frac{n_t - n^*}{n_t} \right)^2 \] should estimate \( \omega_i(c), i = 0, 1, 2 \);

\[ \delta_i^2 c \] a theoretical approximation of \( \omega_i(c), i = 0, 1, 2 \).

### 3.2.2.1. Samples from a Normal Population

We generated pseudo random samples from a \( N(5, 4) \) population, that is, we had fixed \( \mu = 5 \) and \( \sigma = 2 \). We also fixed \( A = 100 \) and \( m = 10, 15, 20 \). In addition, we varied the optimal fixed sample size \( n^*(c) \) over a wide range of values including 50, 100, 200, and then determined the values of \( c(= 0.16, 0.04, 0.01) \) accordingly by referring to (2.2.3). Table 3.5 represents this situation.

Given the set of notation defined in (3.2.21), we present selected summaries based on Table 3.5 when \( m = 10, 15 \) and \( n^*(c) = 50, 100, 200 \) for brevity alone even though we ran simulations with other values of \( n^*(c), A, c, m \). Using \( R = 1000 \) independent replications or runs, we estimated the average final sample size \( \bar{n} \) with its estimated standard error \( s(\bar{n}) \), the minimum \( \min(n) \), the lower quartile \( Q_L(n) \), the median \( \text{med}(n) \), the upper quartile \( Q_U(n) \), and the maximum \( \max(n) \) obtained from 1000 observed sample sizes.

We also looked at the rather natural estimated value, \( \bar{R}_n \), the terminal risk along with its estimated standard error \( s(\bar{R}_n) \) in column 8, \( \hat{\xi} \), the risk efficiency and \( \hat{\omega} \), the regret both in column 9. Considering the the values of \( \bar{R}_n \) along with the estimated standard error, we feel tempted to suggest that the number is generally consistent with Remark 3.1 given at the end of Section 3.2.
3.2.2.2. Samples from an $\epsilon$-Contaminated Normal Population (3.1.7)

This section is similar in principle with what we had reported in Section 3.1.4.2. We generated pseudo random samples from a mixture-normal population distribution given by 

$$(1 - \epsilon)N(5, 4) + \epsilon N(5, 9)$$

with $\epsilon = 0.1$, $\epsilon = 0.2$, and $\lambda = 2$. Our Tables 3.6-3.7 show brief summaries of our findings, parallel to those shown in Tables 3.3-3.4. A majority of the entries in Tables 3.6-3.7 ought to be interpreted as those in Tables 3.3-3.4, respectively. The columns showing $\widehat{R}_n$, the terminal risk, $\widehat{\xi}$, the risk efficiency, and $\widehat{\omega}$, the regret, should be interpreted just like those in Table 3.5.

3.2.2.3. Some Overall Sentiments

From Tables 3.5-3.7 and many other simulations that we had run, we may summarize our overall sentiments as follows when a population may be reasonably treated as a normal universe for practical purposes (as validated by Kolmogorov-Smirnov test), but it is mildly contaminated behind the scene:

We empirically feel confident that our newly developed GMD-based or MAD-based methodologies (3.2.1)-(3.2.2) with $i = 1, 2$ are more robust for practical purposes when we compare them with the sample standard deviation-based methodology (3.2.1)-(3.2.2) with $i = 0$. At the same time, however, the GMD-based or MAD-based methodologies from (3.2.1)-(3.2.2) with $i = 1, 2$ are not very sharply divided. Hence, we suggest implementing either GMD-based or MAD-based methodologies from (3.2.1)-(3.2.2) with $i = 1, 2$ with confidence when sampling from a normal universe, especially when suspect outliers may be expected.

Remark 3.1. For all fixed $\mu, \sigma, c$ and $A$, we can check easily that $\xi_i(c) > 1$, $i = 0, 1, 2$. This is consistent with the observations noted generally by Ghosh and Mukhopadhyay (1976).
Remark 3.2. A more original customary purely sequential minimum risk point estimation methodology excludes the term $n^{-\lambda}$ in the stopping rule from (3.2.1), which is given by

$$N'_0 \equiv N'_0(c) = \inf\{n \geq m : n \geq S_n \sqrt{A}/c\},$$

instead. The conclusion from Theorems 3.2.3-3.2.3 also holds for the associated strategy $\left(N'_0, \bar{X}_{N'_0}\right)$ when $m \geq 3$. One may refer to Ghosh et al. (1997, pp. 174-175), Mukhopadhyay and de Silva (2009, pp. 146-147) or look at another source.

Remark 3.3. Besides the asymptotic first-order properties given in Section 3.2.1, the purely sequential MRPE methodology (3.2.1)-(3.2.2) further enjoys asymptotic second-order risk efficiency. Stronger results will be provided in Chapter 5.

3.3. Illustrations with Real Datasets

In this section, we include illustrations of both FCIE and MRPE methodologies for the mean ($\mu$) using two separate real datasets.

First, in Section 3.3.1, we will include data analysis associated with our proposed methodologies using net profit or loss percentages from 352 department stores with sales less than one million dollars in 1925 referred to as the “sales data” in the sequel. The original data can be found from McNair (1930) or it can be downloaded from the Miscellaneous Datasets page of Winner’s personal website (http://www.stat.ufl.edu/~winner/).

Next, in Section 3.3.2, we will include illustrations using the real data associated with the number of days that the seeds of marigold varieties 2 and 3 needed to flower referred to as the “horticulture data” in the sequel. The data were collected and recorded by Mukhopadhyay et al. (2004b).
3.3.1. Illustrations Using Sales Data

The sales data on net profit or loss percentages for 352 department stores came with min = -14.21 to max = 10.31. The whole dataset looked bit skewed to the left with 5 outlying observations, all on the left, namely -14.21, -12.61, -11.86, -11.04, and -9.97. These were clear from the histogram and boxplot which are not included here for brevity. The whole data (with n = 352) gave the following descriptive statistics:

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>$\bar{x}$</th>
<th>s</th>
<th>min</th>
<th>$QL$</th>
<th>med</th>
<th>$Q_U$</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>352</td>
<td>0.261</td>
<td>3.854</td>
<td>-14.21</td>
<td>-2.1325</td>
<td>0.360</td>
<td>2.9300</td>
<td>10.31</td>
</tr>
</tbody>
</table>

Normality of the data was violated with the p-value 0.008611 associated with the Shapiro-Wilk test on the whole data of size 352. We decided to remove the most extreme outlier, -14.21, and then the normality of the data of size 351 gave the p-value 0.09033 associated with the Shapiro-Wilk test on the data of size 351. The revised descriptive statistics from the data (with n = 351) after removing the most extreme outlier, -14.21, were:

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>$\bar{x}$</th>
<th>s</th>
<th>min</th>
<th>$QL$</th>
<th>med</th>
<th>$Q_U$</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>351</td>
<td>0.302</td>
<td>3.781</td>
<td>-12.61</td>
<td>-2.1250</td>
<td>0.360</td>
<td>2.9400</td>
<td>10.31</td>
</tr>
</tbody>
</table>

For purposes of illustrations, we treated the dataset of size 351 as our population after removing the most extreme outlier, -14.21. We drew random samples from this population separately under simple random sampling with replacement (SRSWR) as well as under simple random sampling without replacement (SRSWOR). They did not give appreciably different results and hence we summarize our findings obtained under SRSWOR only.

We began with a pilot sample of size $m = 20$ on each occasion and implemented
the fixed-width 95% confidence interval methodologies using purely sequential strategies (3.1.1)-(3.1.2) based on the sample variance \( i = 0 \), GMD \( i = 1 \), and MAD \( i = 2 \), respectively. Analogously, we implemented the minimum risk point estimation methodologies using purely sequential strategies (3.2.1)-(3.2.2).

After we began each strategy with \( m \) pilot observations, we checked with all three stopping rules simultaneously based whether to terminate or not based on the same sequence of observations. That is, as one strategy terminated, we continued with the other two based on data accrued thus far. Then, eventually, all three strategies based on rather comparable sets of observations from the same population terminated at some stage depending on the sample path. Thus, one will find descriptive statistics presented in Tables 3.8-3.9 for the observed data \( (x) \) are comparable for \( i = 0, 1, 2 \) within a fixed set of configuration. These are summarized in Tables 3.8-3.9 where the choices of respective design constants are made explicit as needed.

One should reiterate the point that each row in Tables 3.8-3.9 is obtained upon termination from a single run of an appropriate estimation methodology under consideration. We emphasize that we had repeated similar illustrations with all 352 observations included in the population, but we saw no appreciably different results compared with those shown in Tables 3.8-3.9. Hence, we leave them out for brevity.

3.3.2. Illustrations Using Horticulture Data

Here, we utilized the horticulture data collected and recorded by Mukhopadhyay et al. (2004b) on the number of days that the seeds of marigold varieties 2 and 3 needed to flower. We treated the real datasets on variety 2 and variety 3 which seemed not to contradict normal distributions confirmed via Shapiro-Wilk test with associated \( p \)-values 0.1473 and
0.4101 for variety 2 and variety 3 respectively.

However, we saw suspect outliers, for example, the observation 22.0 from variety 2 and observations 22.5, 23.0, 33.5, 34.5 from variety 3. The simple descriptive statistics from full datasets on varieties 2 and 3 are summarized as follows:

<table>
<thead>
<tr>
<th>Variety</th>
<th>n</th>
<th>$\bar{x}$</th>
<th>s</th>
<th>min</th>
<th>$Q_L$</th>
<th>med</th>
<th>$Q_U$</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>460</td>
<td>35.13</td>
<td>4.03</td>
<td>22.0</td>
<td>32.0</td>
<td>35.0</td>
<td>38.0</td>
<td>46.5</td>
</tr>
<tr>
<td>3</td>
<td>162</td>
<td>28.20</td>
<td>2.17</td>
<td>22.5</td>
<td>27.0</td>
<td>28.5</td>
<td>29.5</td>
<td>34.5</td>
</tr>
</tbody>
</table>

For purposes of illustrations, we treated these datasets of sizes 460 and 162 respectively as our populations. We drew random samples from these populations in the spirit of Section 3.3.1 under SRSWR as well as SRSWOR. But, SRSWOR and SRSWR did not produce appreciably different results and hence we summarize our findings obtained under SRSWOR only.

As in Section 3.1.1, again we began with a pilot sample of size $m = 20$ on each occasion and implemented the fixed-width 95% confidence interval methodologies using purely sequential strategies (3.1.1)-(3.1.2) based on the sample variance ($i = 0$), GMD ($i = 1$), and MAD ($i = 2$), respectively. Analogously, we implemented the minimum risk point estimation methodologies using purely sequential strategies (3.2.1)-(3.2.2).

After we began each strategy with $m$ pilot observations, we checked with all three stopping rules simultaneously based whether to terminate or not based on the same sequence of observations. That is, as one strategy terminated, we continued with the other two based on data accrued thus far. Then, eventually, all three strategies based on rather comparable sets of observations from the same population terminated at some stage depending on the sample path.
Thus, one will find descriptive statistics presented in Tables 3.10-3.11 for the observed data \((x)\) are comparable for \(i = 0,1,2\) within a fixed set of configuration. These are summarized in Tables 3.10-3.11 where the choices of respective design constants are made explicit as needed. One should reiterate the point that each row in Tables 3.10-3.11 is obtained upon termination from a single run of an appropriate estimation methodology under consideration. We leave out many details for brevity.

### 3.4. Overall Concluding Thoughts

In summary, the purely sequential procedures (3.1.1)-(3.1.2) and (3.2.1)-(3.2.2) have performed remarkably well, whether based on GMD or MAD, under the normality assumption for the parent population. We saw very little to no significant differences in the overall performances between the three purely sequential procedures, that is (3.1.1)-(3.1.2) and (3.2.1)-(3.2.2), across the board when the population can be reasonably assumed normal.

We empirically feel strongly confident that our newly developed GMD-based or MAD-based methodologies are more robust for practical purposes when we compare them with the customary methodologies based on the sample variance or sample standard deviation, respectively. Hence, we enthusiastically suggest implementing either GMD-based or MAD-based methodologies with confidence when sampling from a normal universe, especially when suspect outliers may be expected.
Table 3.1. Simulations under $\alpha = 0.05$ with 1000 runs implementing methods (3.1.1)-(3.1.2) with $i$ from (3.1.13)

<table>
<thead>
<tr>
<th>$C$</th>
<th>$\pi$</th>
<th>min($n$)</th>
<th>$Q_L(n)$</th>
<th>$\bar{p}$</th>
<th>$\pi/C$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>0.5544</td>
<td>47.823</td>
<td>40.00</td>
<td>48.0</td>
<td>0.925</td>
</tr>
<tr>
<td>0.3920</td>
<td>0.374</td>
<td>82</td>
<td>56.00</td>
<td>0.008</td>
<td>-2.177</td>
</tr>
<tr>
<td>1</td>
<td>47.917</td>
<td>10</td>
<td>41.00</td>
<td>49.0</td>
<td>0.939</td>
</tr>
<tr>
<td>2</td>
<td>0.370</td>
<td>84</td>
<td>56.00</td>
<td>0.008</td>
<td>-2.083</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>47.926</td>
<td>15</td>
<td>41.00</td>
<td>0.926</td>
</tr>
<tr>
<td>1</td>
<td>0.352</td>
<td>78</td>
<td>55.00</td>
<td>0.008</td>
<td>-2.074</td>
</tr>
<tr>
<td>2</td>
<td>48.269</td>
<td>15</td>
<td>42.00</td>
<td>49.0</td>
<td>0.932</td>
</tr>
<tr>
<td>2</td>
<td>0.355</td>
<td>79</td>
<td>56.00</td>
<td>0.008</td>
<td>-1.731</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>98.226</td>
<td>10</td>
<td>88.00</td>
<td>99.0</td>
</tr>
<tr>
<td>0.2772</td>
<td>0.513</td>
<td>145</td>
<td>109.00</td>
<td>0.007</td>
<td>-1.774</td>
</tr>
<tr>
<td>1</td>
<td>98.411</td>
<td>13</td>
<td>88.00</td>
<td>99.0</td>
<td>0.929</td>
</tr>
<tr>
<td>2</td>
<td>0.509</td>
<td>145</td>
<td>109.00</td>
<td>0.008</td>
<td>-1.589</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>98.368</td>
<td>10</td>
<td>88.00</td>
<td>99.0</td>
</tr>
<tr>
<td>2</td>
<td>0.519</td>
<td>149</td>
<td>110.00</td>
<td>0.007</td>
<td>-1.632</td>
</tr>
<tr>
<td>200</td>
<td>0</td>
<td>98.465</td>
<td>24</td>
<td>89.00</td>
<td>99.0</td>
</tr>
<tr>
<td>0.2772</td>
<td>0.497</td>
<td>141</td>
<td>109.00</td>
<td>0.007</td>
<td>-1.435</td>
</tr>
<tr>
<td>1</td>
<td>98.397</td>
<td>38</td>
<td>89.00</td>
<td>99.0</td>
<td>0.952</td>
</tr>
<tr>
<td>2</td>
<td>0.492</td>
<td>149</td>
<td>109.00</td>
<td>0.007</td>
<td>-1.554</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>98.780</td>
<td>30</td>
<td>88.00</td>
<td>99.0</td>
</tr>
<tr>
<td>2</td>
<td>0.515</td>
<td>158</td>
<td>110.00</td>
<td>0.007</td>
<td>-1.220</td>
</tr>
<tr>
<td>200</td>
<td>0</td>
<td>198.998</td>
<td>138</td>
<td>185.00</td>
<td>199.0</td>
</tr>
<tr>
<td>0.2772</td>
<td>0.652</td>
<td>264</td>
<td>212.25</td>
<td>0.007</td>
<td>-1.002</td>
</tr>
<tr>
<td>1</td>
<td>198.681</td>
<td>123</td>
<td>185.00</td>
<td>199.0</td>
<td>0.931</td>
</tr>
<tr>
<td>2</td>
<td>0.652</td>
<td>272</td>
<td>213.00</td>
<td>0.008</td>
<td>-1.319</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>198.573</td>
<td>10</td>
<td>183.00</td>
<td>199.0</td>
</tr>
<tr>
<td>2</td>
<td>0.712</td>
<td>267</td>
<td>213.00</td>
<td>0.007</td>
<td>-1.427</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>198.662</td>
<td>120</td>
<td>185.00</td>
<td>199.0</td>
</tr>
<tr>
<td>1</td>
<td>0.651</td>
<td>266</td>
<td>212.00</td>
<td>0.007</td>
<td>-1.338</td>
</tr>
<tr>
<td>2</td>
<td>198.644</td>
<td>92</td>
<td>185.00</td>
<td>199.0</td>
<td>0.945</td>
</tr>
<tr>
<td>2</td>
<td>0.677</td>
<td>276</td>
<td>213.00</td>
<td>0.007</td>
<td>-1.356</td>
</tr>
<tr>
<td>0.707</td>
<td>277</td>
<td>214.00</td>
<td>0.008</td>
<td>-1.351</td>
<td></td>
</tr>
</tbody>
</table>
Table 3.2. Simulations under $\alpha = 0.05$ with 1000 runs implementing methods (3.1.1)-(3.1.2) picking $d$ first with $i$ from (3.1.13)

<table>
<thead>
<tr>
<th>$C$</th>
<th>$\bar{i}$</th>
<th>$\min(n)$</th>
<th>$Q_L(n)$</th>
<th>$\bar{p}$</th>
<th>$\bar{i}/C$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d$</td>
<td>$m$</td>
<td>$i$</td>
<td>$s(\bar{i})$</td>
<td>$\max(n)$</td>
<td>$Q_U(n)$</td>
</tr>
<tr>
<td>42.683</td>
<td>10</td>
<td>0</td>
<td>41.262</td>
<td>10</td>
<td>35.00</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td></td>
<td>0.338</td>
<td>71</td>
<td>49.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>41.281</td>
<td>10</td>
<td>35.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.341</td>
<td>68</td>
<td>49.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>41.221</td>
<td>10</td>
<td>34.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.357</td>
<td>72</td>
<td>49.00</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>0</td>
<td>41.257</td>
<td>15</td>
<td>34.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.324</td>
<td>73</td>
<td>49.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>41.469</td>
<td>15</td>
<td>35.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.338</td>
<td>74</td>
<td>49.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>41.266</td>
<td>15</td>
<td>34.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.347</td>
<td>71</td>
<td>49.00</td>
</tr>
<tr>
<td>96.036</td>
<td>10</td>
<td>0</td>
<td>94.584</td>
<td>10</td>
<td>85.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.486</td>
<td>140</td>
<td>105.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>94.837</td>
<td>45</td>
<td>85.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.481</td>
<td>141</td>
<td>105.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>95.035</td>
<td>11</td>
<td>85.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.509</td>
<td>138</td>
<td>106.00</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>0</td>
<td>95.093</td>
<td>46</td>
<td>86.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.475</td>
<td>138</td>
<td>105.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>94.947</td>
<td>19</td>
<td>85.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.496</td>
<td>141</td>
<td>106.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>94.753</td>
<td>23</td>
<td>84.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.513</td>
<td>137</td>
<td>106.00</td>
</tr>
<tr>
<td>384.146</td>
<td>10</td>
<td>0</td>
<td>382.746</td>
<td>277</td>
<td>364.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.928</td>
<td>470</td>
<td>403.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>383.010</td>
<td>267</td>
<td>363.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.904</td>
<td>470</td>
<td>402.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>382.866</td>
<td>287</td>
<td>361.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.961</td>
<td>467</td>
<td>404.00</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>0</td>
<td>382.887</td>
<td>274</td>
<td>363.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.910</td>
<td>470</td>
<td>403.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>383.104</td>
<td>259</td>
<td>365.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.924</td>
<td>470</td>
<td>402.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>382.751</td>
<td>269</td>
<td>363.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.977</td>
<td>474</td>
<td>405.00</td>
</tr>
</tbody>
</table>
Table 3.3. Simulations under $\alpha = 0.05$ and $\epsilon = 0.10$ in (3.1.7) with 1000 runs implementing methods (3.1.1)-(3.1.2) with $i$ from (3.1.13)

<table>
<thead>
<tr>
<th>$C$</th>
<th>$d$</th>
<th>$m$</th>
<th>$i$</th>
<th>$\pi$</th>
<th>$\min$</th>
<th>$Q_L$</th>
<th>$\bar{p} - Q_U$</th>
<th>$\bar{p} - C$</th>
<th>$\bar{q}$</th>
<th>$\bar{q}$</th>
<th>$K_S$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>10</td>
<td>0</td>
<td>54.127</td>
<td>10</td>
<td>47.00</td>
<td>55.0</td>
<td>0.934</td>
<td>1.083</td>
<td>0.901</td>
<td>0.926</td>
<td></td>
</tr>
<tr>
<td>56.25</td>
<td>0.391</td>
<td>95</td>
<td>62.00</td>
<td>0.008</td>
<td>4.127</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5544</td>
<td>1</td>
<td>53.367</td>
<td>10</td>
<td>46.00</td>
<td>54.0</td>
<td>0.922</td>
<td>1.067</td>
<td>0.901</td>
<td>0.930</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.394</td>
<td>86</td>
<td>62.00</td>
<td>0.008</td>
<td>3.367</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>52.595</td>
<td>10</td>
<td>44.00</td>
<td>53.0</td>
<td>0.927</td>
<td>1.052</td>
<td>0.902</td>
<td>0.930</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.420</td>
<td>95</td>
<td>62.00</td>
<td>0.008</td>
<td>2.595</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>54.176</td>
<td>15</td>
<td>46.00</td>
<td>55.0</td>
<td>0.952</td>
<td>1.084</td>
<td>0.900</td>
<td>0.946</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.395</td>
<td>92</td>
<td>62.25</td>
<td>0.007</td>
<td>4.176</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>53.636</td>
<td>15</td>
<td>46.00</td>
<td>54.0</td>
<td>0.946</td>
<td>1.073</td>
<td>0.900</td>
<td>0.937</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.378</td>
<td>90</td>
<td>61.00</td>
<td>0.007</td>
<td>3.636</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>53.028</td>
<td>15</td>
<td>45.00</td>
<td>54.0</td>
<td>0.939</td>
<td>1.061</td>
<td>0.899</td>
<td>0.936</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.402</td>
<td>87</td>
<td>61.25</td>
<td>0.008</td>
<td>3.028</td>
<td>0.010</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>10</td>
<td>0</td>
<td>109.924</td>
<td>11</td>
<td>99.00</td>
<td>110.5</td>
<td>0.947</td>
<td>1.099</td>
<td>0.901</td>
<td>0.940</td>
<td></td>
</tr>
<tr>
<td>112.5</td>
<td>0.541</td>
<td>165</td>
<td>121.00</td>
<td>0.007</td>
<td>9.924</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.3920</td>
<td>1</td>
<td>109.314</td>
<td>48</td>
<td>99.00</td>
<td>109.0</td>
<td>0.930</td>
<td>1.093</td>
<td>0.901</td>
<td>0.935</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.511</td>
<td>158</td>
<td>121.00</td>
<td>0.008</td>
<td>9.314</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>107.829</td>
<td>56</td>
<td>97.00</td>
<td>108.0</td>
<td>0.930</td>
<td>1.078</td>
<td>0.901</td>
<td>0.933</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.539</td>
<td>159</td>
<td>120.00</td>
<td>0.008</td>
<td>7.829</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>110.357</td>
<td>53</td>
<td>100.00</td>
<td>111.0</td>
<td>0.944</td>
<td>1.104</td>
<td>0.901</td>
<td>0.940</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.520</td>
<td>155</td>
<td>121.00</td>
<td>0.007</td>
<td>10.357</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>109.284</td>
<td>60</td>
<td>98.00</td>
<td>109.5</td>
<td>0.950</td>
<td>1.093</td>
<td>0.902</td>
<td>0.950</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.504</td>
<td>153</td>
<td>120.00</td>
<td>0.007</td>
<td>9.284</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>107.985</td>
<td>44</td>
<td>97.00</td>
<td>108.5</td>
<td>0.933</td>
<td>1.080</td>
<td>0.901</td>
<td>0.939</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.541</td>
<td>158</td>
<td>120.00</td>
<td>0.008</td>
<td>7.985</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>10</td>
<td>0</td>
<td>222.790</td>
<td>150</td>
<td>208.00</td>
<td>224.0</td>
<td>0.929</td>
<td>1.114</td>
<td>0.901</td>
<td>0.938</td>
<td></td>
</tr>
<tr>
<td>225</td>
<td>0.713</td>
<td>293</td>
<td>238.00</td>
<td>0.008</td>
<td>22.790</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2772</td>
<td>1</td>
<td>219.633</td>
<td>143</td>
<td>204.00</td>
<td>220.0</td>
<td>0.932</td>
<td>1.098</td>
<td>0.901</td>
<td>0.941</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.725</td>
<td>292</td>
<td>235.00</td>
<td>0.008</td>
<td>19.633</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>218.486</td>
<td>135</td>
<td>202.00</td>
<td>219.0</td>
<td>0.940</td>
<td>1.092</td>
<td>0.900</td>
<td>0.939</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.727</td>
<td>289</td>
<td>234.00</td>
<td>0.008</td>
<td>18.486</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>222.277</td>
<td>153</td>
<td>207.00</td>
<td>223.0</td>
<td>0.929</td>
<td>1.111</td>
<td>0.901</td>
<td>0.937</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.721</td>
<td>294</td>
<td>238.00</td>
<td>0.008</td>
<td>22.277</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>219.796</td>
<td>153</td>
<td>205.00</td>
<td>220.0</td>
<td>0.925</td>
<td>1.099</td>
<td>0.901</td>
<td>0.927</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.695</td>
<td>276</td>
<td>235.00</td>
<td>0.008</td>
<td>19.796</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>218.059</td>
<td>144</td>
<td>202.00</td>
<td>218.0</td>
<td>0.937</td>
<td>1.090</td>
<td>0.900</td>
<td>0.936</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.744</td>
<td>289</td>
<td>234.00</td>
<td>0.008</td>
<td>18.059</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3.4. Simulations under $\alpha = 0.05$ and $\epsilon = 0.20$ in (3.1.7) with 1000 runs implementing methods (3.1.1)-(3.1.2) with $i$ from (3.1.13)

| C   | d    | m | i | $\bar{\pi}$ | $s(\bar{\pi})$ | min | $Q_L$ | $Q_U$ | med | $\bar{p}$ | $s(\bar{p})$ | $\bar{\pi}/C$ | $s(\bar{\pi})$ | $\bar{\gamma}$ | $s(\bar{\gamma})$ | KS |
|-----|------|---|---|-------------|----------------|-----|-------|-------|-----|----------|-------------|--------------|----------------|----------------|----------------|----------------|------------|
| 50  | 10   | 0 | 59.114 | 10 | 51.75 | 60.0 | **0.931** | 1.182 | 0.805 | 0.921 |
| 62.5| 0.443| 101| 68.00 | 0.008 | **9.114** | 0.013 |
| 0.5544 | 1 | 58.879 | 10 | 51.00 | 59.0 | **0.938** | 1.178 | 0.803 | 0.937 |
|     | 0.409| 96 | 68.00 | 0.008 | **8.879** | 0.013 |
|     | 2 | 58.155 | 10 | 50.00 | 59.0 | **0.935** | 1.163 | 0.803 | 0.938 |
|     | 0.426| 96 | 67.00 | 0.008 | **8.155** | 0.013 |
|     | 15 | 0 | 60.021 | 15 | 51.00 | 61.0 | **0.940** | 1.200 | 0.803 | 0.923 |
|     | 0.431| 96 | 69.00 | 0.008 | **10.021** | 0.013 |
|     | 1 | 59.358 | 15 | 51.00 | 60.0 | **0.930** | 1.187 | 0.802 | 0.930 |
|     | 0.395| 93 | 68.00 | 0.008 | **9.358** | 0.013 |
|     | 2 | 58.139 | 15 | 50.00 | 59.0 | **0.933** | 1.163 | 0.804 | 0.932 |
|     | 0.431| 104 | 67.00 | 0.008 | **8.139** | 0.013 |
| 100 | 10   | 0 | 122.534 | 13 | 111.00 | 123.0 | **0.934** | 1.225 | 0.801 | 0.925 |
| 125 | 0.570| 171 | 135.00 | 0.008 | **22.534** | 0.013 |
| 0.3920 | 1 | 120.809 | 30 | 109.75 | 122.0 | **0.935** | 1.208 | 0.801 | 0.930 |
|     | 0.538| 180 | 132.00 | 0.008 | **20.809** | 0.013 |
|     | 2 | 118.579 | 10 | 107.00 | 119.0 | **0.945** | 1.186 | 0.801 | 0.941 |
|     | 0.568| 172 | 130.00 | 0.007 | **18.579** | 0.013 |
|     | 15 | 0 | 122.738 | 68 | 110.00 | 123.0 | **0.939** | 1.227 | 0.801 | 0.932 |
|     | 0.553| 181 | 135.00 | 0.008 | **22.738** | 0.013 |
|     | 1 | 121.306 | 53 | 111.00 | 122.0 | **0.944** | 1.213 | 0.801 | 0.929 |
|     | 0.539| 172 | 133.00 | 0.007 | **21.306** | 0.013 |
|     | 2 | 118.898 | 66 | 107.00 | 120.0 | **0.937** | 1.189 | 0.801 | 0.933 |
|     | 0.573| 173 | 131.00 | 0.008 | **18.898** | 0.013 |
| 200 | 10   | 0 | 247.049 | 166 | 230.00 | 248.0 | **0.936** | 1.235 | 0.802 | 0.897 |
| 250 | 0.790| 318 | 264.00 | 0.008 | **47.049** | 0.013 |
| 0.2772 | 1 | 242.648 | 165 | 226.00 | 242.5 | **0.935** | 1.213 | 0.801 | 0.914 |
|     | 0.756| 311 | 259.00 | 0.008 | **42.648** | 0.013 |
|     | 2 | 239.725 | 150 | 224.00 | 240.0 | **0.932** | 1.199 | 0.801 | 0.917 |
|     | 0.750| 318 | 255.00 | 0.008 | **39.725** | 0.013 |
|     | 15 | 0 | 246.803 | 168 | 231.00 | 248.0 | **0.936** | 1.234 | 0.801 | 0.912 |
|     | 0.752| 303 | 263.00 | 0.008 | **46.803** | 0.013 |
|     | 1 | 243.253 | 173 | 228.00 | 244.0 | **0.938** | 1.216 | 0.801 | 0.909 |
|     | 0.743| 308 | 259.00 | 0.008 | **43.253** | 0.013 |
|     | 2 | 239.931 | 149 | 224.00 | 240.0 | **0.935** | 1.200 | 0.801 | 0.907 |
|     | 0.765| 310 | 256.00 | 0.008 | **39.931** | 0.013 |
Table 3.5. Simulations under $A = 100$ with 1000 runs implementing methods (3.2.1)-(3.2.2) with $i$ from (3.2.20) and $\lambda = 2$

<table>
<thead>
<tr>
<th>$n^*$</th>
<th>$c$</th>
<th>$m$</th>
<th>$i$</th>
<th>$\pi$</th>
<th>$s(\pi)$</th>
<th>min(n)</th>
<th>$Q_L(n)$</th>
<th>max(n)</th>
<th>$Q_V(n)$</th>
<th>med(n)</th>
<th>$\bar{R}_n$</th>
<th>$s(\bar{R}_n)$</th>
<th>$\tilde{\xi}$</th>
<th>$\tilde{\omega}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>10</td>
<td>0</td>
<td>50.199</td>
<td>22</td>
<td>47.00</td>
<td>50.0</td>
<td>15.777</td>
<td>0.055</td>
<td>0.104</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.16</td>
<td>50.313</td>
<td>30</td>
<td>47.00</td>
<td>51.0</td>
<td>15.806</td>
<td>0.054</td>
<td>0.098</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>50.259</td>
<td>29</td>
<td>46.00</td>
<td>51.0</td>
<td>15.795</td>
<td>0.053</td>
<td>0.107</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>50.247</td>
<td>28</td>
<td>47.00</td>
<td>50.0</td>
<td>15.801</td>
<td>0.053</td>
<td>0.093</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.168</td>
<td>50.460</td>
<td>33</td>
<td>47.00</td>
<td>51.0</td>
<td>15.839</td>
<td>0.053</td>
<td>0.095</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.171</td>
<td>50.457</td>
<td>31</td>
<td>47.00</td>
<td>51.0</td>
<td>15.843</td>
<td>0.053</td>
<td>0.098</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>10</td>
<td>0</td>
<td>100.107</td>
<td>72</td>
<td>95.00</td>
<td>100.0</td>
<td>7.937</td>
<td>0.992</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.04</td>
<td>0.238</td>
<td>123</td>
<td>106.00</td>
<td>0.019</td>
<td>2.33 x 10^{-2}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>100.335</td>
<td>70</td>
<td>95.00</td>
<td>100.0</td>
<td>7.954</td>
<td>0.994</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.235</td>
<td>121</td>
<td>106.00</td>
<td>0.018</td>
<td>2.25 x 10^{-2}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>100.159</td>
<td>76</td>
<td>95.00</td>
<td>101.0</td>
<td>7.941</td>
<td>0.993</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.229</td>
<td>124</td>
<td>105.00</td>
<td>0.018</td>
<td>2.15 x 10^{-2}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.233</td>
<td>123</td>
<td>105.25</td>
<td>0.018</td>
<td>2.23 x 10^{-2}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>10</td>
<td>0</td>
<td>200.132</td>
<td>165</td>
<td>194.00</td>
<td>201.0</td>
<td>3.985</td>
<td>0.996</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>0.333</td>
<td>231</td>
<td>207.00</td>
<td>0.007</td>
<td>5.64 x 10^{-3}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>200.255</td>
<td>167</td>
<td>193.00</td>
<td>201.0</td>
<td>3.988</td>
<td>0.997</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.338</td>
<td>230</td>
<td>207.00</td>
<td>0.007</td>
<td>5.80 x 10^{-3}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>200.026</td>
<td>162</td>
<td>193.00</td>
<td>200.0</td>
<td>3.985</td>
<td>0.996</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.356</td>
<td>230</td>
<td>208.00</td>
<td>0.007</td>
<td>6.44 x 10^{-3}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>200.199</td>
<td>163</td>
<td>193.00</td>
<td>201.0</td>
<td>3.986</td>
<td>0.997</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.328</td>
<td>229</td>
<td>207.25</td>
<td>0.007</td>
<td>5.46 x 10^{-3}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>200.201</td>
<td>160</td>
<td>193.00</td>
<td>200.0</td>
<td>3.987</td>
<td>0.997</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>200.103</td>
<td>151</td>
<td>193.00</td>
<td>201.0</td>
<td>3.989</td>
<td>0.997</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.346</td>
<td>229</td>
<td>208.00</td>
<td>0.007</td>
<td>6.16 x 10^{-3}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3.6. Simulations under $A = 100$ and $\epsilon = 0.10$ in (3.1.7) with 1000 runs implementing methods (3.2.1)-(3.2.2) with $i$ from (3.2.20) and $\lambda = 2$

<table>
<thead>
<tr>
<th>$c$</th>
<th>$m$</th>
<th>$i$</th>
<th>$\pi$</th>
<th>min $Q_L$</th>
<th>$Q_U$</th>
<th>med $\bar{R}_n$</th>
<th>$\widehat{\xi}$</th>
<th>$\overline{\xi} \ (\pi)$</th>
<th>$KS$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>10</td>
<td>0</td>
<td>53.191</td>
<td>26</td>
<td>50.00</td>
<td>53.0 16.744</td>
<td>1.046</td>
<td>0.900</td>
<td>0.938</td>
</tr>
<tr>
<td>53.033</td>
<td>0.183</td>
<td>70</td>
<td>57.00</td>
<td>0.058</td>
<td>12.7</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>52.973</td>
<td>30</td>
<td>49.00</td>
<td>53.0 16.749</td>
<td>1.047</td>
<td>0.901</td>
<td>0.931</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.16</td>
<td>0.178</td>
<td>69</td>
<td>57.00</td>
<td>0.057</td>
<td>11.8</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>52.756</td>
<td>33</td>
<td>49.00</td>
<td>53.0 16.751</td>
<td>1.047</td>
<td>0.901</td>
<td>0.948</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.189</td>
<td>71</td>
<td>57.00</td>
<td>0.059</td>
<td>12.5</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>53.059</td>
<td>28</td>
<td>49.00</td>
<td>53.0 16.699</td>
<td>1.044</td>
<td>0.902</td>
<td>0.948</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.182</td>
<td>70</td>
<td>57.00</td>
<td>0.058</td>
<td>12.2</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>53.152</td>
<td>29</td>
<td>50.00</td>
<td>53.0 16.809</td>
<td>1.051</td>
<td>0.900</td>
<td>0.949</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.171</td>
<td>71</td>
<td>57.00</td>
<td>0.056</td>
<td>11.3</td>
<td>0.010</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>52.917</td>
<td>35</td>
<td>49.00</td>
<td>53.0 16.787</td>
<td>1.049</td>
<td>0.901</td>
<td>0.941</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.177</td>
<td>70</td>
<td>56.00</td>
<td>0.055</td>
<td>11.5</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>10</td>
<td>0</td>
<td>106.393</td>
<td>84</td>
<td>101.00</td>
<td>106.0</td>
<td>8.439</td>
<td>1.055</td>
<td>0.901</td>
</tr>
<tr>
<td>106.066</td>
<td>0.250</td>
<td>130</td>
<td>112.00</td>
<td>0.020</td>
<td>3.64</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.04</td>
<td>1</td>
<td>105.719</td>
<td>72</td>
<td>100.00</td>
<td>105.0</td>
<td>8.427</td>
<td>1.053</td>
<td>0.901</td>
<td>0.959</td>
</tr>
<tr>
<td></td>
<td>0.249</td>
<td>130</td>
<td>111.00</td>
<td>0.021</td>
<td>3.41</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>105.254</td>
<td>75</td>
<td>100.00</td>
<td>105.0</td>
<td>8.424</td>
<td>1.053</td>
<td>0.900</td>
<td>0.939</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.260</td>
<td>129</td>
<td>111.00</td>
<td>0.021</td>
<td>3.42</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>106.611</td>
<td>84</td>
<td>102.00</td>
<td>107.0</td>
<td>8.456</td>
<td>1.057</td>
<td>0.900</td>
<td>0.947</td>
</tr>
<tr>
<td></td>
<td>0.242</td>
<td>131</td>
<td>112.00</td>
<td>0.019</td>
<td>3.62</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>105.974</td>
<td>79</td>
<td>101.00</td>
<td>106.0</td>
<td>8.446</td>
<td>1.056</td>
<td>0.901</td>
<td>0.961</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.239</td>
<td>128</td>
<td>111.00</td>
<td>0.019</td>
<td>3.31</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>105.353</td>
<td>79</td>
<td>100.00</td>
<td>105.0</td>
<td>8.434</td>
<td>1.054</td>
<td>0.901</td>
<td>0.945</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.253</td>
<td>128</td>
<td>110.00</td>
<td>0.020</td>
<td>3.34</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>10</td>
<td>0</td>
<td>211.699</td>
<td>175</td>
<td>204.00</td>
<td>211.0</td>
<td>4.217</td>
<td>1.054</td>
<td>0.901</td>
</tr>
<tr>
<td>212.132</td>
<td>0.348</td>
<td>241</td>
<td>220.00</td>
<td>0.007</td>
<td>0.116</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>1</td>
<td>210.866</td>
<td>178</td>
<td>204.00</td>
<td>211.0</td>
<td>4.221</td>
<td>1.055</td>
<td>0.901</td>
<td>0.953</td>
</tr>
<tr>
<td></td>
<td>0.350</td>
<td>242</td>
<td>218.00</td>
<td>0.007</td>
<td>0.109</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>210.113</td>
<td>164</td>
<td>203.00</td>
<td>210.5</td>
<td>4.224</td>
<td>1.056</td>
<td>0.901</td>
<td>0.947</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.358</td>
<td>244</td>
<td>217.00</td>
<td>0.007</td>
<td>0.105</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>212.014</td>
<td>176</td>
<td>205.00</td>
<td>212.0</td>
<td>4.223</td>
<td>1.056</td>
<td>0.901</td>
<td>0.944</td>
</tr>
<tr>
<td></td>
<td>0.339</td>
<td>248</td>
<td>219.00</td>
<td>0.007</td>
<td>0.117</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>210.620</td>
<td>175</td>
<td>204.00</td>
<td>211.0</td>
<td>4.217</td>
<td>1.054</td>
<td>0.901</td>
<td>0.959</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.347</td>
<td>246</td>
<td>218.00</td>
<td>0.007</td>
<td>0.106</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>210.045</td>
<td>177</td>
<td>203.00</td>
<td>210.0</td>
<td>4.224</td>
<td>1.056</td>
<td>0.901</td>
<td>0.956</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.349</td>
<td>244</td>
<td>217.00</td>
<td>0.007</td>
<td>0.101</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

48
Table 3.7. Simulations under $A = 100$ and $\epsilon = 0.20$ in (3.1.7) with 1000 runs implementing methods (3.2.1)-(3.2.2) with $i$ from (3.2.20) and $\lambda = 2$

<table>
<thead>
<tr>
<th>$n^*$</th>
<th>$c$</th>
<th>$m$</th>
<th>$i$</th>
<th>$\bar{n}$ (max)</th>
<th>$Q_L$</th>
<th>$Q_U$</th>
<th>$\bar{n}_{\text{min}}$ (med)</th>
<th>$100\hat{\omega}$</th>
<th>$\hat{\gamma}$</th>
<th>$s (\hat{\gamma})$</th>
<th>KS</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>10</td>
<td>0</td>
<td>55.759</td>
<td>52.00</td>
<td>56.0</td>
<td>17.552</td>
<td>1.097</td>
<td>0.803</td>
<td>0.920</td>
<td></td>
<td></td>
</tr>
<tr>
<td>55.902</td>
<td>0.197</td>
<td>75</td>
<td>60.00</td>
<td>0.063</td>
<td>0.190</td>
<td>0.013</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.16</td>
<td>1</td>
<td>55.339</td>
<td>33</td>
<td>51.00</td>
<td>56.0</td>
<td>17.554</td>
<td>1.097</td>
<td>0.801</td>
<td>0.929</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>55.075</td>
<td>27</td>
<td>51.00</td>
<td>55.0</td>
<td>17.589</td>
<td>1.099</td>
<td>0.801</td>
<td>0.938</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>55.934</td>
<td>26</td>
<td>52.00</td>
<td>56.0</td>
<td>17.609</td>
<td>1.101</td>
<td>0.801</td>
<td>0.927</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>10</td>
<td>0</td>
<td>111.491</td>
<td>106.00</td>
<td>111.0</td>
<td>8.847</td>
<td>1.106</td>
<td>0.803</td>
<td>0.927</td>
<td></td>
<td></td>
</tr>
<tr>
<td>111.803</td>
<td>0.270</td>
<td>136</td>
<td>117.00</td>
<td>0.022</td>
<td>6.87</td>
<td>0.013</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.04</td>
<td>1</td>
<td>110.892</td>
<td>84</td>
<td>105.00</td>
<td>111.0</td>
<td>8.869</td>
<td>1.109</td>
<td>0.802</td>
<td>0.927</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>110.008</td>
<td>76</td>
<td>104.00</td>
<td>110.0</td>
<td>8.862</td>
<td>1.108</td>
<td>0.802</td>
<td>0.924</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>111.654</td>
<td>81</td>
<td>106.00</td>
<td>111.0</td>
<td>8.861</td>
<td>1.108</td>
<td>0.802</td>
<td>0.926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>110.990</td>
<td>79</td>
<td>106.00</td>
<td>111.0</td>
<td>8.875</td>
<td>1.109</td>
<td>0.802</td>
<td>0.936</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>110.267</td>
<td>84</td>
<td>105.00</td>
<td>110.0</td>
<td>8.877</td>
<td>1.110</td>
<td>0.803</td>
<td>0.926</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>10</td>
<td>0</td>
<td>223.196</td>
<td>216.00</td>
<td>223.0</td>
<td>4.446</td>
<td>1.111</td>
<td>0.800</td>
<td>0.935</td>
<td></td>
<td></td>
</tr>
<tr>
<td>223.607</td>
<td>0.371</td>
<td>262</td>
<td>231.00</td>
<td>0.007</td>
<td>0.291</td>
<td>0.013</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>1</td>
<td>221.380</td>
<td>162</td>
<td>214.00</td>
<td>221.0</td>
<td>4.447</td>
<td>1.112</td>
<td>0.801</td>
<td>0.920</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>219.880</td>
<td>164</td>
<td>212.00</td>
<td>220.0</td>
<td>4.446</td>
<td>1.112</td>
<td>0.801</td>
<td>0.922</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>223.241</td>
<td>175</td>
<td>216.00</td>
<td>223.0</td>
<td>4.447</td>
<td>1.112</td>
<td>0.801</td>
<td>0.924</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>221.250</td>
<td>179</td>
<td>214.00</td>
<td>221.0</td>
<td>4.445</td>
<td>1.111</td>
<td>0.801</td>
<td>0.930</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>219.774</td>
<td>180</td>
<td>213.00</td>
<td>220.0</td>
<td>4.446</td>
<td>1.112</td>
<td>0.801</td>
<td>0.924</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3.8. Purely sequential fixed-width confidence intervals from sales data with −14.21 removed from it under estimation strategies (3.1.1)-(3.1.2) with $m = 20$, $\alpha = 0.05$: Sampling without replacement (SRSWOR) from 351 observations

<table>
<thead>
<tr>
<th>$i$</th>
<th>$n_i$</th>
<th>$\bar{x}_{n_i}$</th>
<th>$s_{n_i}$</th>
<th>min</th>
<th>$Q_L$</th>
<th>med</th>
<th>$Q_U$</th>
<th>max</th>
<th>$\bar{x}_{n_i} - d$</th>
<th>$\bar{x}_{n_i} + d$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$d = 0.50$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>222</td>
<td>0.41</td>
<td>3.79</td>
<td>−12.61</td>
<td>−2.02</td>
<td>0.76</td>
<td>2.97</td>
<td>8.56</td>
<td>−0.09</td>
<td>0.91</td>
</tr>
<tr>
<td>1</td>
<td>204</td>
<td>0.37</td>
<td>3.69</td>
<td>−12.61</td>
<td>−1.89</td>
<td>0.61</td>
<td>2.85</td>
<td>8.56</td>
<td>−0.13</td>
<td>0.87</td>
</tr>
<tr>
<td>2</td>
<td>205</td>
<td>0.36</td>
<td>3.69</td>
<td>−12.61</td>
<td>−1.96</td>
<td>0.57</td>
<td>2.84</td>
<td>8.56</td>
<td>−0.14</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$d = 0.75$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>73</td>
<td>0.24</td>
<td>3.25</td>
<td>−9.97</td>
<td>−1.65</td>
<td>0.21</td>
<td>2.44</td>
<td>7.17</td>
<td>−0.51</td>
<td>0.99</td>
</tr>
<tr>
<td>1</td>
<td>72</td>
<td>0.27</td>
<td>3.26</td>
<td>−9.97</td>
<td>−1.55</td>
<td>0.54</td>
<td>2.48</td>
<td>7.17</td>
<td>−0.48</td>
<td>1.02</td>
</tr>
<tr>
<td>2</td>
<td>73</td>
<td>0.24</td>
<td>3.25</td>
<td>−9.97</td>
<td>−1.65</td>
<td>0.21</td>
<td>2.44</td>
<td>7.17</td>
<td>−0.51</td>
<td>0.99</td>
</tr>
</tbody>
</table>
Table 3.9. Purely sequential minimum risk point estimators from sales data with −14.21 removed from it under estimation strategies (3.2.1)-(3.2.2) with \( m = 20, A = 100, c = 0.1, \lambda = 2 \):

Sampling without replacement (SRSWOR) from 351 observations

<table>
<thead>
<tr>
<th>( i )</th>
<th>( n_i )</th>
<th>( \bar{x}_{n_i} )</th>
<th>( s_{n_i} )</th>
<th>min</th>
<th>( Q_L )</th>
<th>med</th>
<th>( Q_U )</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>117</td>
<td>−0.09</td>
<td>3.69</td>
<td>−12.61</td>
<td>−2.120</td>
<td>0.10</td>
<td>2.520</td>
<td>7.17</td>
</tr>
<tr>
<td>1</td>
<td>114</td>
<td>−0.22</td>
<td>3.65</td>
<td>−12.61</td>
<td>−2.135</td>
<td>0.07</td>
<td>2.148</td>
<td>7.17</td>
</tr>
<tr>
<td>2</td>
<td>113</td>
<td>−0.24</td>
<td>3.66</td>
<td>−12.61</td>
<td>−2.140</td>
<td>0.07</td>
<td>2.140</td>
<td>7.17</td>
</tr>
</tbody>
</table>
Table 3.10. Purely sequential fixed-width confidence intervals from horticulture data under estimation strategies (3.1.1)-(3.1.2) with $m = 20, \alpha = 0.05$: Sampling without replacement (SRSWOR)

<table>
<thead>
<tr>
<th>(3.1.13)</th>
<th>$n_i$</th>
<th>$\bar{x}_{n_i}$</th>
<th>$s_{n_i}$</th>
<th>min</th>
<th>$Q_L$</th>
<th>med</th>
<th>$Q_U$</th>
<th>max</th>
<th>$\bar{x}_{n_i} - d$</th>
<th>$\bar{x}_{n_i} + d$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variety 2 population size 460: $d = 0.75$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>98</td>
<td>35.55</td>
<td>3.78</td>
<td>27.5</td>
<td>33.5</td>
<td>35.5</td>
<td>37.9</td>
<td>46.5</td>
<td>34.80</td>
<td>36.30</td>
</tr>
<tr>
<td>1</td>
<td>94</td>
<td>35.45</td>
<td>3.73</td>
<td>27.5</td>
<td>33.1</td>
<td>35.5</td>
<td>37.5</td>
<td>46.5</td>
<td>34.70</td>
<td>36.20</td>
</tr>
<tr>
<td>2</td>
<td>90</td>
<td>35.41</td>
<td>3.73</td>
<td>27.5</td>
<td>33.1</td>
<td>35.5</td>
<td>37.5</td>
<td>46.5</td>
<td>34.66</td>
<td>36.16</td>
</tr>
<tr>
<td>Variety 3 population size 162: $d = 0.50$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>75</td>
<td>27.73</td>
<td>2.21</td>
<td>23.0</td>
<td>26.5</td>
<td>27.5</td>
<td>29.0</td>
<td>34.5</td>
<td>27.23</td>
<td>28.23</td>
</tr>
<tr>
<td>1</td>
<td>74</td>
<td>27.73</td>
<td>2.22</td>
<td>23.0</td>
<td>26.5</td>
<td>27.5</td>
<td>29.0</td>
<td>34.5</td>
<td>27.23</td>
<td>28.23</td>
</tr>
<tr>
<td>2</td>
<td>75</td>
<td>27.73</td>
<td>2.21</td>
<td>23.0</td>
<td>26.5</td>
<td>27.5</td>
<td>29.0</td>
<td>34.5</td>
<td>27.23</td>
<td>28.23</td>
</tr>
</tbody>
</table>
Table 3.11. Purely sequential minimum risk point estimators from horticulture data under estimation strategies (3.2.1)-(3.2.2) with $m = 20, A = 100, c = 0.1, \lambda = 2$:

Sampling without replacement (SRSWOR)

<table>
<thead>
<tr>
<th>(3.2.29)</th>
<th>Descriptive stats from $n_i$ obs. at Termination</th>
</tr>
</thead>
<tbody>
<tr>
<td>$i$</td>
<td>$n_i$</td>
</tr>
<tr>
<td>Variety 2 population size 460</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>129</td>
</tr>
<tr>
<td>1</td>
<td>130</td>
</tr>
<tr>
<td>2</td>
<td>132</td>
</tr>
<tr>
<td>Variety 3 population size 162</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>75</td>
</tr>
<tr>
<td>1</td>
<td>74</td>
</tr>
<tr>
<td>2</td>
<td>72</td>
</tr>
</tbody>
</table>
Chapter 4

Two-Stage Methodologies
Involving GMD or MAD

In this chapter, we propose the sample variance-based, GMD-based and MAD-based Stein-type (1945, 1949) two-stage estimation strategies when a positive lower bound, denoted $\sigma^2_L$, for the underlying unknown variance $\sigma^2$ is known to us. Both minimum risk point estimation (MRPE) and fixed-width confidence interval estimation (FCIE) problems are investigated. The chapter is based on Mukhopadhyay and Hu (2018).

The structure of this chapter looks similar with that of Chapter 3, but we develop the new Stein-type two-stage estimation strategies for the MRPE problems first in Section 4.1. Section 4.1.1 summarizes the existing two-stage strategies which are relevant in our context, followed by newly developed strategies based on sample standard deviation (indicated by: $i = 0$), GMD (indicated by: $i = 1$), and MAD (indicated by: $i = 2$) in Section 4.1.2. The case "$i = 0$" resembles the well-known methodologies due to Robbins (1959) and Starr (1966).

The main results have been summarized in Theorems 4.1.1-4.1.3. These show that our newly proposed estimation strategies with unbiased estimators of $\sigma$ based on sample standard deviation, GMD, or MAD in defining the stopping boundaries enjoy both asymptotic first-order and second-order efficiency as well as first-order and second-order risk efficiency properties. In Section 4.1.4, simulated performances are briefly discussed.

Section 4.2 develops new two-stage strategies in the context of the FCIE problems.
The main results regarding asymptotic first-order and second-order efficiency and asymptotic consistency properties have been summarized in Theorems 4.2.1-4.2.3, followed with summaries obtained from simulations in Section 4.2.3.

Section 4.3 includes illustrations with real data under both two-stage MRPE and FCIE strategies. In doing so, we again appropriately use the horticulture data recorded and presented by Mukhopadhyay et al. (2004b), which has appeared in Section 3.3.2.

Section 4.4 summarizes some final overall thoughts.

4.1. Minimum Risk Point Estimation

Recall that Robbins (1959) put forward the MRPE problem for a normal mean \( \mu \) having an unknown population variance \( \sigma^2 \). He developed a purely sequential stopping rule and established the original MRPE formulation. Starr (1966) extended his idea by appropriately modifying the loss function.

Here, we revisit Stein (1945,1949) and apply his breakthrough two-stage procedure to address the MRPE problem.

4.1.1. Existing Two-Stage MRPE Methodologies

Under the MRPE formulation (2.2.1)-(2.2.3), \( \sigma \) remains unknown even though an expression of \( n^* \) is known. A straightforward idea is to use the sample standard deviation as an estimator of \( \sigma \) in the expression of \( n^* \) in (2.2.3). With the pilot data of size \( m(\geq m_0 \geq 2) \), in the spirit of Ghosh and Mukhopadhyay (1976), a Stein-type two-stage MRPE procedure is proposed as follows:

\[
N \equiv N(c) = \max \left\{ m, \left[ S_m \sqrt{A/c} \right] + 1 \right\},
\]  

(4.1.1)
where \([w]\) continues to denote the largest integer that is smaller \(<\) than \(w\) as appeared in (3.2.10).

Here, \(m\) and \(m_0\) are fixed positive integers chosen by an experimenter. Note, however, that while \(S^2_m\) is unbiased for \(\sigma^2\), the sample standard deviation \(S_m\) used in (4.1.1) is biased for \(\sigma\). Another sticky point is this: How to determine \(m\), the size of the pilot data? We have more to say shortly.

But, whatever may be a fixed choice of \(m\), if we observe \(N = m\), we do not gather any additional data. Otherwise, that is if \(N > m\), we record additional \((N - m)\) observations in the second stage. Then, having recorded the data \(\{N, X_1, ..., X_N\}\) from both stages combined, the sample mean \(\bar{X}_N\) is used to estimate the unknown normal mean \(\mu\).

Although the Stein-type two-stage strategy (4.1.1) is convenient to implement, the amount of oversampling relative to \(n^*\) becomes bothersome. It does not have the efficiency property, let alone the first-order efficiency or first-order risk efficiency properties. One may refer to Ghosh and Mukhopadhyay (1976).

Mukhopadhyay and Duggan (1997) made the observation that if a positive lower bound for the normal variance \(\sigma^2\) is known to us, say \(0 < \sigma_L^2 \leq \sigma^2\), then an appropriately updated stein-type two-stage strategy would enjoy both asymptotic first-order and second-order efficiency and risk efficiency properties. Their modification of (4.1.1) is summarized as follows: Let

\[
m \equiv m(c) = \max \left\{ m_0, \left\lfloor \frac{\sigma_L \sqrt{A/c}}{c} \right\rfloor + 1 \right\},
\]

(4.1.2)

be our determined pilot sample size instead of an experimenter’s arbitrarily chosen fixed \(m(\geq m_0)\). Observe that the pilot size \(m \equiv m(c) \to \infty\) as \(c \to 0\), and yet \(m \equiv m(c)\) remains small compared with \(n^*(c)\) as \(c \to 0\) in the sense that \(m(c)/n^*(c) < 1\) as \(c \to 0\).
Let the final sample size be given by:

\[ N_0 \equiv N_0(c) = \max \left\{ m, \left[ S_m \sqrt{A/c} \right] + 1 \right\}, \tag{4.1.3} \]

obtained from the pilot data \( \{m, X_1, \ldots, X_m\} \) and then we implement the procedure as we did in the case of (4.1.1). Having obtained the combined data

\[ \{N_0, X_1, \ldots, X_{N_0}\}, \]

the minimum risk point estimator for \( \mu \) is proposed to be \( \bar{X}_{N_0} = N_0^{-1} \sum_{j=1}^{N_0} X_j \).

4.1.2. Substituting Sample Standard Deviation with Alternative Estimators

One should note that in (4.1.3), while \( S_m \) is a consistent estimator for \( \sigma \), it is not unbiased and tends to underestimate \( \sigma \). In view of the inequality: \( \sigma = \sqrt{E_{\mu,\sigma}[S_m^2]} > E_{\mu,\sigma}[S_m] \), it is understood that \( S_m \) tends to underestimate \( \sigma \). Therefore, we should consider an appropriate multiple of \( S_m \), denoted by \( V_{m;0} \), to make it unbiased in the normal case by explicitly defining in the spirit of (2.3.4) such that

\[ V_{m;0} = \Gamma \left( \frac{1}{2}(m - 1) \right) \sqrt{m - 1} \left\{ \sqrt{2} \Gamma \left( \frac{1}{2}m \right) \right\}^{-1} S_m. \]

We also use \( V_{m;0} \) in the sequel to replace \( S_m \) in (4.1.3).

Following (2.3.4), we also propose unbiased estimators for \( \sigma \) based GMD and MAD denoted by \( V_{m;i}, i = 1, 2 \), respectively. The associated estimators are correspondingly expressed as:

(i) \[ V_{m;1} = \frac{\sqrt{\pi}}{2} G_m; \]

(ii) \[ V_{m;2} = \sqrt{\frac{\pi m}{2(m-1)}} M_m, \]
where $G_m$ and $M_m$ come from (2.3.1).

Now we are in a position to propose our new two-stage MRPE strategies and the final sample sizes are precisely given by

$$N_i \equiv N_i(c) = \max \left\{ m, \left\lfloor V_{m;i} \sqrt{A/c} \right\rfloor + 1 \right\}, \quad i = 0, 1, 2,$$

(4.1.4)

where $m$ was defined in (4.1.2). Having obtained the final dataset $\{N_i, X_1, \ldots, X_{N_i}\}$, the associated minimum risk point estimators are then defined as follows:

$$\bar{X}_{N_i} = N_i^{-1} \sum_{j=1}^{N_i} X_j, \quad i = 0, 1, 2.$$

(4.1.5)

### 4.1.3. Properties of the Two-Stage Methodology (4.1.4)-(4.1.5) with $m$ from (4.1.2)

In this section, under the newly developed Stein-type two-stage strategy (4.1.4)-(4.1.5) with $m$ coming from (4.1.2), we first enumerate a number of interesting results summarized by a series of lemmas and theorems.

**Lemma 4.1.1.** For the final sample size $N_i$ defined by the two-stage estimation strategy (4.1.4)-(4.1.5) with $m$ coming from (4.1.2), for all fixed $\mu, \sigma$, and $A$, we have for $r \geq 2$:

$$P_{\mu,\sigma}\{N_i = m\} = O\left(c^{r/4}\right), \quad i = 0, 1, 2,$$

(4.1.6)

for sufficiently small $c$.

**Proof.** We start with $i = 1$. For sufficiently small $c$, we have:
Thus, we can claim in the spirit of the proof of Lemma 3.2.1: by the lemma of Sen and Ghosh (1981),

\[ E_{\mu,\sigma} \left[ |G_m - 2\sigma/\sqrt{\pi}| \right] = O \left( m^{-r/2} \right) = O \left( c^{r/4} \right), \text{ for sufficiently small } c. \tag{4.1.8} \]

Therefore, \( P_{\mu,\sigma} \{ N_1 = m \} = O \left( c^{r/4} \right). \)

Next, for \( i = 2, \) we note (w.p.1):

\[ M_m \geq m^{-1} \sum_{i=1}^{m} |X_i - \mu| - |\overline{X}_m - \mu| = M_{1m} - M_{2m}, \text{ say.} \]

Then, with \( \varepsilon = \sigma - \sigma_L, \) for sufficiently small \( c, \) we can express:

\[
P_{\mu,\sigma} \{ N_2 = m \}
= P_{\mu,\sigma} \{ V_{m:2} \leq \sigma_L \}
\leq P_{\mu,\sigma} \{ M_{1m} - M_{2m} \leq 2\sigma_L/\sqrt{\pi} \}
\leq P_{\mu,\sigma} \{ |M_{1m} - 2\sigma/\sqrt{\pi}| + |M_{2m}| \geq 2\varepsilon/\sqrt{\pi} \}
\leq P_{\mu,\sigma} \{ |M_{1m} - 2\sigma/\sqrt{\pi}| \geq \varepsilon/\sqrt{\pi} \} + P_{\mu,\sigma} \{ \mid \overline{X}_m - \mu \mid \geq \varepsilon/\sqrt{\pi} \}.
\tag{4.1.9}
\]

Thus we can conclude in a similar way that

\[ P_{\mu,\sigma} \{ N_2 = m \} = O \left( c^{r/4} \right). \tag{4.1.10} \]
Finally, when \( i = 0 \), for sufficiently small \( c \), we express:

\[
P_{\mu,\sigma} \{ N_0 = m \} \leq P_{\mu,\sigma} \{ S_m \leq \sigma L \} \leq P_{\mu,\sigma} \{|S_m - \sigma| \geq \sigma - \sigma L \} \leq O \left( m^{-r/2} \right) = O \left( c^{r/4} \right),
\]

(4.1.11)

by Corollary 1.2 in Mukhopadhyay and Vik (1985). The proof is now complete. ■

**Lemma 4.1.2.** For the final sample size \( N_i \) defined by the two-stage estimation strategy (4.1.4)-(4.1.5) with \( m \) coming from (4.1.2), for all fixed \( \mu, \sigma, \) and \( A \), we have:

\[
N_i / n^* \xrightarrow{\text{P}} 1 \text{ as } c \to 0,
\]

(4.1.12)

for \( i = 0, 1, 2 \) where \( n^* \) was defined in (2.2.3).

**Proof.** From (4.1.4), for \( i = 0, 1, 2 \), we have (w.p.1):

\[
V_{m;i} \sqrt{A/c} \leq N_i \leq mI (N_i = m) + V_{m;i} \sqrt{A/c} + 1.
\]

Thus, we can rewrite (w.p.1):

\[
\frac{V_{m;i}}{\sigma} \leq \frac{N_i}{n^*} \leq \frac{mI (N_i = m) \sqrt{c/A} + V_{m;i} + \sqrt{c/A}}{\sigma}.
\]

(4.1.13)

Next by slightly modifying the proof of Lemma 3.1.2, we can show that \( V_{m;i}/\sigma \xrightarrow{\text{P}} 1 \) as \( c \to 0 \). Also, from Lemma 4.1.1 we have:

\[
mI (N_i = m) \sigma^{-1} \sqrt{c/A} = O_{\mu,\sigma} \left( c^{r/4} \right),
\]

so that we can immediately conclude: \( N_i / n^* \xrightarrow{\text{P}} 1 \) as \( c \to 0, i = 0, 1, 2 \). ■
Theorem 4.1.1. For the final sample size $N_i$ defined by the two-stage estimation strategy (4.1.4)-(4.1.5) with $m$ coming from (4.1.2), for all fixed $\mu, \sigma$, and $A$, we have:

$$\lim_{c \to 0} E_{\mu, \sigma}[N_i/n^*] = 1, i = 0, 1, 2 \text{ [Asymptotic First-Order Efficiency]}, \quad (4.1.14)$$

where $n^*$ was defined in (2.2.3).

Proof. For sufficiently small $c$, we observe (w.p.1): for $i = 0, 1, 2$,

$$\frac{mI(N_i = m)}{\sigma \sqrt{A/c}} \leq 1, \quad \frac{V_{m,i}}{\sigma} \leq 2,$$

so that $N_i/n^* \leq 4$. Now, using Lemma 4.1.2, we obtain (4.1.14). \hfill \blacksquare

Theorem 4.1.2. For the final sample size $N_i$ defined by the two-stage estimation strategy (4.1.4)-(4.1.5) with $m$ coming from (4.1.2), for all fixed $\mu, \sigma$, and $A$, we have:

$$\lim_{c \to 0} \xi_i(c) = 1, i = 0, 1, 2 \text{ [Asymptotic First-Order Risk Efficiency]} \quad (4.1.15)$$

with $\xi_i(c)$ defined in (3.2.6).

Proof. In view of Theorem 4.1.1, it will suffice to prove that $\lim_{c \to 0} E_{\mu, \sigma}[n^*/N_i] = 1$ for $i = 0, 1, 2$. Now, we show outlines of the proof in the case of three different choices of $i$.

Observe that

(i) $i = 0$: We have

$$P_{\mu, \sigma} \left\{ N_0 \leq \frac{1}{2} n^* \right\} \leq P_{\mu, \sigma} \left\{ V_{m,0} \sqrt{A/c} \leq \frac{1}{2} \sigma \sqrt{A/c} \right\} \leq P_{\mu, \sigma} \left\{ |S_m - \sigma| \geq \frac{1}{2} \sigma \right\}; \quad (4.1.16)$$
(ii) $i = 1$: We have

$$P_{\mu,\sigma} \left\{ N_1 \leq \frac{1}{2} n^* \right\} \leq P_{\mu,\sigma} \left\{ V_{m;1} \sqrt{A/c} \leq \frac{1}{2} \sigma \sqrt{A/c} \right\} \leq P_{\mu,\sigma} \left( |V_{m;1} - \sigma| \geq \frac{1}{2} \sigma \right); \quad (4.1.17)$$

(iii) $i = 2$: Recall the expressions of $M_{1m}$ and $M_{2m}$ from (4.1.9) and write

$$P_{\mu,\sigma} \left\{ N_2 \leq \frac{1}{2} n^* \right\}$$

$$\leq P_{\mu,\sigma} \left\{ V_{k;2} \leq \frac{1}{2} \sigma \right\} \quad (4.1.18)$$

$$\leq P_{\mu,\sigma} \left\{ M_{1m} - 2\sigma/\sqrt{\pi} - M_{2m} \leq \frac{\sqrt{2}}{2\sqrt{\pi}} \sigma \right\}$$

$$\leq P_{\mu,\sigma} \left\{ |M_{1m} - 2\sigma/\sqrt{\pi}| \geq \frac{4 - \sqrt{2}}{4\sqrt{\pi}} \sigma \right\} + P_{\mu,\sigma} \left\{ |M_{2m}| \geq \frac{4 - \sqrt{2}}{4\sqrt{\pi}} \sigma \right\}.$$

Thus, we can claim $P_{\mu,\sigma} \left\{ N_i \leq \frac{1}{2} n^* \right\} = O \left( e^{r/4} \right)$ for all $r > 2, i = 0, 1, 2$.

At this point, we can conclude that $\lim_{c \to 0} \xi_i(c) = 1, i = 0, 1, 2$ in the same way as we did in the proof of Theorem 3.2.2. $\blacksquare$

**Theorem 4.1.3.** For the final sample size $N_i$ defined by the two-stage estimation strategy (4.1.4)-(4.1.5) with $m$ coming from (4.1.2), for all fixed $\mu, \sigma$, and $A$, we have:

$$\lim_{c \to 0} E_{\mu,\sigma} [N_i - n^*] \text{ is bounded \{} \text{Asymptotic Second-Order Efficiency} \}, \quad (4.1.19)$$

for $i = 0, 1, 2$ where $n^*$ was defined in (2.2.3).

**Proof.** Note that for $i = 0, 1, 2$,

$$(V_{m;i} - \sigma) \sqrt{A/c} \leq N_i - n^* \leq mI (N_i = m) + (V_{m;i} - \sigma) \sqrt{A/c} + 1, \text{ w.p.1.} \quad (4.1.20)$$

We immediately have:
\[ 0 \leq E_{\mu, \sigma}[N_i - n^*] \leq 1 + O\left( c^{\frac{r-2}{4}} \right), \quad (4.1.21) \]

provided that \( r > 2 \). Thus, (4.1.19) holds. \( \blacksquare \)

**Lemma 4.1.3.** For the final sample size \( N_i \) defined by the two-stage estimation strategy (4.1.4)-(4.1.5) with \( m \) coming from (4.1.2), for all \( \mu, \sigma, \) and \( A \), we have:

\[
n^{*-1/2} (N_i - n^*) \xrightarrow{L} N(0, \delta_i^2 \sigma \sigma_L^{-1}) \quad \text{as} \quad c \to 0, i = 0, 1, 2, \quad (4.1.22)
\]

where \( \delta_0^2 = \frac{1}{2}, \delta_1^2 = \frac{1}{3}(\pi + 6\sqrt{3} - 12), \delta_2^2 = \frac{1}{7}(\pi - 2) \) and \( n^* \) was defined in (2.2.3).

**Proof.** From (4.1.20), we can express (w.p.1) for \( i = 0, 1, 2, \)

\[
\sigma^{-1/2} (A/c)^{1/4} (V_{m;i} - \sigma) \leq n^{*-1/2} (N_i - n^*) \leq \sigma^{-1/2} (A/c)^{1/4} (V_{m;i} - \sigma)
\]

\[
+ \sigma^{-1/2} (A/c)^{-1/4} \{mI (N_i = m) + 1\}. \quad (4.1.23)
\]

Now, \( \sigma^{-1/2} (A/c)^{-1/4} \{mI (N_i = m) + 1\} \xrightarrow{P,\mu,\sigma} 0 \) as \( c \to 0 \), so that by Slutsky’s theorem, for sufficiently small \( c \), the asymptotic distribution of \( n^{*-1/2} (N_i - n^*) \) is the same as that of \( (\sigma \sigma_L)^{-1/2} m^{1/2} (V_{m;i} - \sigma) \).

(i) \( i = 0 \): With \( S_m^2 \), we have as \( m \to \infty \):

\[
n^{1/2} \left( S_m^2 - \sigma^2 \right) \xrightarrow{L} N(0, 2\sigma^4),
\]

so that by the delta method, one gets:

\[
n^{1/2} (S_m - \sigma) \xrightarrow{L} N(0, \frac{1}{2}\sigma^2).
\]
Thus,

\[ (\sigma \sigma_L)^{-1/2} m^{1/2} (V_{m;0} - \sigma) \xrightarrow{\mathcal{L}} N(0, \delta_0^2 \sigma \sigma_L^{-1}), \]

as \( c \to 0 \), where \( \delta_0^2 = \frac{1}{2} \).

(ii) \( i = 1 \): Since \( G_m \) is a \( U \)-statistic, from Hoeffding (1948,1961), we have as \( m \to \infty \),

\[ m^{1/2} (G_m - 2\sigma/\sqrt{\pi}) \xrightarrow{\mathcal{L}} N(0, 4\sigma_1^2), \]

where \( \sigma_1^2 = \sigma^2 \left( \frac{1}{3} + \frac{2\sqrt{3} - 4}{\pi} \right) \). Thus, we have:

\[ (\sigma \sigma_L)^{-1/2} m^{1/2} (V_{m;1} - \sigma) \xrightarrow{\mathcal{L}} N(0, \delta_1^2 \sigma \sigma_L^{-1}), \]

as \( c \to 0 \), where \( \delta_1^2 = \frac{1}{3}(\pi + 6\sqrt{3} - 12) \).

(iii) \( i = 2 \): By appealing to Babu and Rao (1992), we have as \( m \to \infty \):

\[ n^{1/2} \left( M_m - \sigma \sqrt{2/\pi} \right) \xrightarrow{\mathcal{L}} N(0, \frac{\pi-2}{\pi} \sigma^2). \]

Thus, we can claim:

\[ (\sigma \sigma_L)^{-1/2} m^{1/2} (V_{m;2} - \sigma) \xrightarrow{\mathcal{L}} N(0, \delta_2^2 \sigma \sigma_L^{-1}), \]

as \( c \to 0 \), where \( \delta_2^2 = \frac{1}{2}(\pi - 2) \).

Lemma 4.1.4. For the final sample size \( N_i \) defined by the two-stage estimation strategy (4.1.4)-(4.1.5) with \( m \) coming from (4.1.2), for all fixed \( \mu, \sigma, \) and \( A \), we have:

\[ n^{\sigma^{-1}} (N_i - n^*)^2, i = 0, 1, 2 \text{ is uniformly integrable,} \]
where \( n^* \) was defined in \((2.2.3)\).

Proof. The necessary steps can be developed by utilizing \((4.1.23)\) repeatedly in the spirits of Mukhopadhyay and Duggan (1997,1999). Details are left out for brevity. ■

Theorem 4.1.4. For the final sample size \( N_i \) defined by the two-stage estimation strategy \((4.1.4)-(4.1.5)\) with \( m \) coming from \((4.1.2)\), for all fixed \( \mu, \sigma, \) and \( A \), we have:

\[
\omega_i (c) = \delta_i^2 \sigma L^{-1} c + o(c), i = 0, 1, 2 \quad [\text{Asymptotic Second-Order Risk Efficiency}],
\]

\[(4.1.24)\]

where \( \omega_i (c) \) was defined in \((3.2.6)\) and \( \delta_0^2 = \frac{1}{2}, \delta_1^2 = \frac{1}{3}(\pi + 6\sqrt{3} - 12), \delta_2^2 = \frac{1}{2}(\pi - 2) \) remain the same as in Lemma 4.1.3.

Proof. We express: for \( i = 0, 1, 2, \)

\[
\omega_i (c) = cE_{\mu, \sigma} \{ N_i^{-1}(N_i - n^*)^2 I (N_i = m) \}
\]

\[(4.1.25)\]

\[
+ cE_{\mu, \sigma} \{ N_i^{-1}(N_i - n^*)^2 I (N_i > m) \}.
\]

On the set \( \{ N_i > m \} \), we observe (w.p.1):

\[
N_i^{-1}(N_i - n^*)^2 = \frac{n^*}{\left( V_{m;i}/\sqrt{A/c} \right) + 1} \frac{(N_i - n^*)^2}{n^*}.
\]

Next, we note (w.p.1):

\[
\frac{\sigma}{V_{m;i}/\sqrt{c/A}} \leq \frac{n^*}{\left( V_{m;i}/\sqrt{A/c} \right) + 1} \leq \frac{\sigma}{V_{m;i}}.
\]

Since both \( \frac{\sigma}{V_{m;i}/\sqrt{c/A}} \) and \( \frac{\sigma}{V_{m;i}} \) converge to 1 in probability\((P_{\mu,\sigma})\) as \( c \to 0 \), we can claim
that for sufficiently small $c$, 
\[ \frac{n^*}{\left\lfloor V_{n^*} \sqrt{A/c} \right\rfloor + 1} \leq 2. \]
Thus, \( N_i^{-1} (N_i - n^*)^2 I (N_i > m) \) is uniformly integrable and then for sufficiently small $c$, we get:

\[ E_{\mu, \sigma} \{ N_i^{-1} (N_i - n^*)^2 I (N_i > m) \} = \delta_i^2 \sigma \sigma_L^{-1}. \tag{4.1.26} \]

Next, on the set \( \{ N_i = m \} \), we can express:

\[ E_{\mu, \sigma} \{ N_i^{-1} (N_i - n^*)^2 I (N_i = m) \} \leq O (m) P_{\mu, \sigma} \{ N_i = m \} = o (1). \tag{4.1.27} \]

Combining (4.1.26) and (4.1.27), the proof of (4.1.24) is complete. 

### 4.1.4. Simulations and Data Analysis

In this section, we summarize simulated performances of the two-stage MRPE methodology \( (N_i, \overline{X}_{N_i}), i = 0, 1, 2 \) from (4.1.4)-(4.1.5) with $m$ coming from (4.1.2), where the index $i$ identifies each strategy under implementation:

\[ i = 0: \text{ Sample standard deviation-based methodology; } \]
\[ i = 1: \text{ GMD-based methodology; } \]
\[ i = 2: \text{ MAD-based methodology. } \tag{4.1.28} \]

We set $\mu = 5$ and $\sigma = 4$, and thus generated pseudorandom samples from a \( N(5, 4^2) \) population. For convenience, we fixed $m_0 = 10$, $A = 100$, and $c = 1.00, 0.16, 0.04, 0.01$ so that the optimal fixed sample size $n^*$ from (2.2.3) would cover a range of values including 40, 100, 200, 400, that is from small to medium to large.

Further, we had assumed that a prior lower bound of the variance was known to us:

We considered $\sigma_L = 1, 2, 3$. Table 4.1 summarizes the findings in these situations obtained
under \( T (= 10000, \text{say}) \) independent replications. We report the average final sample size \( \bar{n} \) along with its estimated standard error \( s(\bar{n}) \), the average risk \( \bar{R}_n \) along with its estimated standard error \( s(\bar{R}_n) \), estimated risk efficiency \( \left( \hat{\xi} \right) \) and regret \( (\hat{\omega}) \) explained in (3.2.40) as well as other entities explained in (4.1.29).

\[ \eta_i c \text{ where } \eta_i = \delta_i^2 \sigma \sigma_L^{-1}, \text{ for } i = 0, 1, 2: \text{ a theoretical approximation of} \]

\[ (4.1.29) \]

the \( \omega_i (c) \).

In Table 4.1, we show 3 blocks corresponding to assigned values of \( \sigma_L \). Within each block, there is little or no significant difference among three procedures \( (i = 0, 1, 2) \), no matter whether the sample sizes were small (40), moderate (100), or large (200, 400). We note from the fifth column that all \( \bar{n} \) values appear a bit over the corresponding optimal fixed sample sizes \( n^* \), but within one observation. These empirically validate Theorem 4.1.3 about the asymptotic second-order efficiency property from equation (4.1.21).

\( \hat{\xi} \) from column 9 obtained using column 7 for \( \bar{R}_n \) reflects the asymptotic first-order risk efficiency property which turns out to remain tightly around 1. Upon comparing the different blocks, it becomes obvious that for larger lower bound \( \sigma_L \) for \( \sigma \), the corresponding \( \hat{\omega} \) values in the last column become smaller and get closer to 0. The estimated regret values stay close to the theoretical ones shown in Theorem 4.1.4 and they compare well with the asymptotic second-order terms, namely, \( \eta_i c = \delta_i^2 \sigma \sigma_L^{-1} c, i = 0, 1, 2 \). All three procedures \( (i = 0, 1, 2) \) appear to perform remarkably well.

4.2. Fixed-Width Confidence Intervals

The fundamental breakthrough ideas of Stein (1945,1949) have been pushed by many
researchers to new heights. Mukhopadhyay and Duggan (1997) and Chattopadhyay and Mukhopadhyay (2013) proceeded to propose new FCIE strategies with practical modifications. We continue to suppose that we want to estimate the normal mean $\mu$ using a confidence interval having a fixed-width $2d$, $d > 0$, with its associated preassigned confidence level at least $(1 - \alpha)$, $0 < \alpha < 1$ under the framework of (2.1.1)-(2.1.3).

4.2.1. A Formulation

While $\sigma^2$ remains unknown in (2.1.3), we use sample variance and GMD-based or MAD-based unbiased estimators of $\sigma^2$, defined in (2.3.1)-(2.3.3) such that with $m(\geq m_0 \geq 2)$:

$$U_{m;i}^2 = \begin{cases} 
S_m^2, & \text{corresponding to } i = 0 \\
\frac{1}{c_{m;1}}G_m^2, & \text{corresponding to } i = 1 \\
\frac{1}{c_{m;2}}M_m^2, & \text{corresponding to } i = 2 
\end{cases} \quad (4.2.1)$$

Thus far, $m$ and $m_0$ have been fixed positive integers chosen arbitrarily by an experimenter. But, how to determine $m$, the size of the pilot data? Under the assumption that $\sigma > \sigma_L > 0$ where $\sigma_L$ is known, we propose the following choice of $m$ and the associated Stein-type two-stage strategies. Let us define:

$$m \equiv m(d) = \max\{m_0, \left\lfloor \frac{a^2\sigma_L^2}{d^2} \right\rfloor + 1\}, \quad (4.2.2)$$

$$N_i \equiv N_i(d) = \max\{m, \left\lfloor \frac{a^2U_{m;i}^2}{d^2} \right\rfloor + 1\},$$

and accordingly construct the fixed-width confidence intervals:

$$J_{N_i} = \left[ \bar{X}_{N_i} \pm d \right], \quad (4.2.3)$$

for the unknown mean $\mu$ based on the combined data $\{N_i, X_1, ..., X_{N_i}\}$, for $i = 0, 1, 2$. 

68
We should mention that for \( i = 0 \), Mukhopadhyay and Duggan (1997) used \( b_{m;0}^2 \), the upper 100 (\( \alpha/2 \) \%) point of the Student’s \( t \) distribution with degrees of freedom \((m - 1)\) instead of \( a^2 \) to define \( N_0 \) in (4.2.3). Also, for \( i = 1, 2 \), Chattopadhyay and Mukhopadhyay (2013) used respectively \( b_{m;1}^2 \) and \( b_{m;2}^2 \), the upper 100 (\( \alpha/2 \) \%) point of the corresponding pivotal distribution of the sample mean standardized appropriately by \( U_{m;1} \) and \( U_{m;2} \), instead.

While our present substitution of \( a^2 \) in the place of \( b_{m;0}^2, b_{m;1}^2, \) and \( b_{m;2}^2 \) may lead to minor loss in the coverage probability, the newly proposed strategies (4.2.2)-(4.2.3) will continue to enjoy attractive asymptotic first-order and second-order properties. The oversampling problem would also be less pronounced. Such performances are summarized in the following sections.

4.2.2. Properties of the Two-Stage Methodology (4.2.2)-(4.2.3)

Under the strategies (4.2.2)-(4.2.3), we have the following result in the spirit of Lemma 4.1.1. We state it without giving its proof.

**Lemma 4.2.1.** For the final sample size \( N_i \) defined by the two-stage estimation strategy (4.2.2)-(4.2.3), for all fixed \( \mu, \sigma, \alpha \) and sufficiently small \( d \), we have:

\[
P_{\mu,\sigma} \{ N_i = m \} = O(d^r), \quad i = 0, 1, 2,
\]

where \( r \geq 2 \).

Now, we move forward to formally summarize the following asymptotic first-order and second-order properties. Some steps in the proofs are omitted for brevity since they may be constructed similarly along the lines of those which were shown in the case of the MRPE problem in Section 4.1.
Theorem 4.2.1. For the final sample size \( N_i \) defined by the two-stage estimation strategy (4.2.2)-(4.2.3), for all fixed \( \mu, \sigma, \alpha \) and sufficiently small \( d \), we have:

\[
\lim_{d \to 0} E_{\mu,\sigma} [N_i/C] = 1, \; i = 0, 1, 2 \quad \text{[Asymptotic First-Order Efficiency]},
\]

where \( C \) was defined in (2.1.3).

Theorem 4.2.2. For the final sample size \( N_i \) defined by the two-stage estimation strategy (4.2.2)-(4.2.3), for all fixed \( \mu, \sigma, \alpha \) and sufficiently small \( d \), we have:

\[
\lim_{d \to 0} P_{\mu,\sigma} \{ \mu \in J_{N_i} \} = 1 - \alpha, \; i = 0, 1, 2 \quad \text{[Asymptotic Consistency]},
\]

where \( J_{N_i} \) was defined in (4.2.3).

Proof. For \( i = 0, 1, 2 \), one should note that for all \( n \geq m \), \( I (N_i = n) \) depends only on \( U_{m;i}^2 \) which are independent of \( \overline{X}_n \). Thus, we can write:

\[
P_{\mu,\sigma} \{ \mu \in J_{N_i} \} = E_{\mu,\sigma} \left[ 2\Phi \left( N_i^{1/2} d/\sigma \right) - 1 \right],
\]

which converges to \( 1 - \alpha \) as \( d \to 0 \) in view of the dominated convergence theorem. \( \blacksquare \)

Theorem 4.2.3. For the final sample size \( N_i \) defined by the two-stage estimation strategy (4.2.2)-(4.2.3), for all fixed \( \mu, \sigma, \alpha \) and sufficiently small \( d \), we have:

(i) \( \lim_{d \to 0} E_{\mu,\sigma} [N_i - C] \) is bounded [Asymptotic Second-Order Efficiency];

(ii) \( C^{-1/2} (N_i - C) \xrightarrow{d} N(0, 4\delta_i^2 \sigma^2 L^{-2}) \) as \( d \to 0 \), where \( \delta_i^2 \) are defined in (3.2.22);

(iii) \( C^{-1} (N_i - C)^2 \) is uniformly integrable;

\[
(4.2.8)
\]
for $i = 0, 1, 2$ where $C$ was defined in (2.1.3).

**Proof.** For $i = 0, 1, 2$, in part (i), for small enough $d$, we note (w.p.1):

$$a^2 \left( U_{m;i}^2 - \sigma^2 \right) / d^2 \leq N_i - C \leq mI(N_i = m) + a^2 \left( U_{m;i}^2 - \sigma^2 \right) / d^2 + 1,$$

(4.2.9)

where $mI(N_i = m) = O(d^{r-2})$ provided that $r > 2$. Thus, as $d \to 0$, we have:

$$0 \leq E_{\mu,\sigma} [N_i - C] \leq 1 + O(d^{r-2}).$$

(4.2.10)

For part (ii), we observe that as $d \to 0$, $C^{-1/2}(N_i - C)$ has the same asymptotic distribution as that of $(\sigma L)^{-1}m^{1/2}\left(U_{m;i}^2 - \sigma^2\right)$. With delta method in place, we can claim Part (ii). Part (iii) can be derived similarly as in Lemma 4.1.4. ■

### 4.2.3. Simulations and Data Analysis

In this section, we summarize performances of the FCIE strategies $(N_i, J_{N_i}), i = 0, 1, 2$ from (4.2.2)-(4.2.3), where the index $i$ is used to identify each procedure under implementation as before. We use the previous settings described in Section 4.1.4 with $\mu = 5$ and $\sigma = 4$, and assume that the known lower bound $\sigma_L$ varies through $\sigma_L = 1, 2, 3$. We also fixed $m_0 = 10$ as what we did in Section 4.1.4.

For brevity, we consider $\alpha = 0.05$ only, and set the optimal fixed sample size $C$ across a range of small, moderate and large values, including 50, 100, 200, 500, and 1000 so that we are able to compare simulated performances under a variety of sample sizes.

The values of $d$ can be determined by referring to (2.1.3). Table 4.2 summarizes the findings in these situations under $T (= 10000, \text{say})$ independent replications. We obtained the average final sample size $\overline{n}$ along with its estimated standard error $s(\overline{n})$, the average
coverage probability $\bar{p}$ along with its estimated standard error $s(\bar{p})$.

The observed values of $n/C$ and $\bar{n} - C$, associated with first-order and second-order terms, are both included. We use the basic notation similar in spirit with those reported in (3.1.14).

In Table 4.2, within each block corresponding to different values of $\sigma_L$, there is little to no significant difference among three procedures, when the sample sizes increased from small (50) to medium (100, 200) to large (500, 1000). We note that our new two-stage strategies do not suffer from noticeable oversampling problem. This is clear since the $\bar{n}$ values from column 5 appear close to the corresponding optimal fixed sample sizes shown in column 2.

The simulations demonstrate that the consistency indeed holds as an asymptotic feature instead of an exact property seen from column 7 shown for $\bar{p}$. The average coverage probabilities appear close to the target level $0.95 (= 1 - \alpha)$ when sample sizes increased. We also note that as $\sigma_L$ increased from 1 to 3, the standard error of $\bar{n}$ dropped and the average coverage probability increased.

4.3. Illustrations with Real Datasets

In this section, we conduct further data analyses and illustrate both MRPE and FCIE methodologies for the normal mean $\mu$ based on methodologies (4.1.4)-(4.1.5) with $m$ from (4.1.2) and (4.2.2)-(4.2.3). Here, we use the horticulture data that comes from Mukhopadhyay et al. (2004b) again, referred to as “horticulture data” recording the number of days that the marigold variety 2 seeds needed to flower. It is assumed that the data followed a normal distribution with a lower bound of standard deviation at least 2.0.

For purposes of illustrations, we consider this data set of size 460 with $\mu = 35.13$ and $\sigma = 4.03$ as a representative of the marigold variety 2 seeds population. In the spirit of
bootstrap, we draw random samples from this “population” under simple random sampling with replacement (SRSWR). Since the population is rather large, simple random sampling without replacement (SRSWOR) would produce comparable results. Fixing \( m_0 = 10 \), we include both MRPE and FCIE methodologies from Sections 4.3.1 and 4.3.2, respectively.

### 4.3.1. MRPE Methodologies

Here, we take advantage of the following bootstrap technique to help us improve the quality of estimators from strategies (4.1.4)-(4.1.5) with \( m \) from (4.1.2). We first fix \( A = 100 \) and \( c = 0.1 \). With the known lower bound of standard deviation \( \sigma_L = 2 \), we are able to determine the pilot sample size \( m \). Then, we applied all three two-stage methodologies (4.1.4)-(4.1.5), namely sample standard deviation-based \((i = 0)\), GMD-based \((i = 1)\), and MAD-based \((i = 2)\), to obtain the final sample size \( N = n \).

Then, we selected extra \((N - m)\) observations if needed (that is, if \( N > m \)) from the population to record our observed sample data to be bootstrapped with \( b = 200 \) replications. With each bootstrap sample, we can calculate the mean and variance. The detailed notation is provided below:

\[
\begin{align*}
B_r & \quad : \text{the } r^{th} \text{ bootstrap sample;} \\
\bar{x}_r & \quad : \text{the bootstrap sample size;} \\
x_{rj} & \quad : \text{the } j^{th} \text{ observation in the } r^{th} \text{ bootstrap sample} \\
\bar{x}_r & = \frac{1}{n} \sum_{j=1}^{n} x_{rj} \quad : \text{the } r^{th} \text{ bootstrap mean estimate;} \\
\hat{\sigma}_r^2 & = \frac{1}{n} \sum_{j=1}^{n} (x_{rj} - \bar{x}_r)^2 \quad : \text{the } r^{th} \text{ bootstrap variance estimate;} \\
\hat{\sigma}_r^2 & = \frac{1}{b} \sum_{r=1}^{b} \hat{\sigma}_r^2 \quad : \text{the bootstrap estimate of variance;} \\
\hat{\sigma}_{BS}^2 & = \frac{1}{b} \sum_{r=1}^{b} \hat{\sigma}_r^2 \quad : \text{the bootstrap estimate of variance;} \\
\hat{\sigma}_{BS} & : \text{the bootstrap estimate of standard deviation.}
\end{align*}
\]
Table 4.3 summarizes the descriptive statistics for $i = 0, 1, 2$, associated with the strategies (4.1.4)-(4.1.5) with $m$ from (4.1.2). The three different procedures yielded results with little to no difference.

### 4.3.2. FCIE Methodologies

Having fixed $\alpha = 0.05$, $d = 0.5$, and $\sigma_L = 2$, we were able to determine the pilot sample size $m$. As in Section 4.3.1, we again applied the bootstrap algorithm. Within each bootstrap, in addition to (4.3.1), we also included the following entities:

$$J_r = [\bar{x}_r \pm d] \quad : \text{the } r^{th} \text{ bootstrap fixed-width confidence interval;}$$

$$p_r : 1 \ (0) \ \text{if} \ \mu \in (\notin) J_r;$$

$$\hat{p}_{BS} = \frac{1}{b} \sum_{r=1}^{b} p_r \quad : \text{should be comparable with } 1 - \alpha;$$

$$s(\hat{p}_{BS}) = \left\{ b^{-1} \hat{p}_{BS} (1 - \hat{p}_{BS}) \right\}^{1/2} \quad : \text{estimated standard error of } \hat{p}_{BS}. \quad (4.3.2)$$

Table 4.4 summarizes the descriptive statistics for $i = 0, 1, 2$, associated with the strategies (4.2.2)-(4.2.3). Again, the three different procedures yielded comparable results with little to no difference.

### 4.4. Overall Concluding Thoughts

In summary, the Stein-type two-stage MRPE and FCIE methodologies from (4.1.4)-(4.1.5) with $m$ from (4.1.2) and (4.2.2)-(4.2.3) have performed remarkably well, when we used unbiased estimators of $\sigma$ (or $\sigma^2$) involving the sample standard deviation (or sample variance), GMD or MAD in defining the stopping boundaries, under the assumption normality. We did not notice appreciable differences in the overall performances of the proposed procedures in the context of either point estimation problem or the interval estimation problem.
Table 4.1. Simulations from $N(5, 4^2)$ under $A = 100, m_0 = 10$
with $T = 10000$ runs implementing methods (4.1.4)-(4.1.5)
with $m$ from (4.1.2) and $i$ from (4.1.28)
\[
\begin{array}{cccccccccc}
\sigma_L & n^* & c & i & \bar{\pi} & s(\bar{\pi}) & \overline{R_n} & s(\overline{R_n}) & \hat{\xi} & \hat{\omega} \\
\hline
1 & 40 & 1 & 0 & 40.452 & 0.096 & 80.091 & 0.101 & 1.001 & 2.456 \\
 & 1 & 40.547 & 0.096 & 80.028 & 0.102 & 1.000 & 2.488 \\
 & 2 & 40.539 & 0.101 & 80.275 & 0.104 & 1.003 & 2.761 \\
 & 100 & 0.16 & 0 & 100.499 & 0.145 & 32.013 & 0.024 & 1.000 & 3.48 \\
 & 1 & 100.332 & 0.146 & 32.037 & 0.023 & 1.001 & 0.350 \\
 & 2 & 100.509 & 0.153 & 32.071 & 0.024 & 1.002 & 0.384 \\
 & 200 & 0.04 & 0 & 200.216 & 0.201 & 15.991 & 0.008 & 0.999 & 0.082 \\
 & 1 & 200.043 & 0.202 & 15.998 & 0.008 & 1.000 & 0.083 \\
 & 2 & 200.266 & 0.215 & 16.011 & 0.008 & 1.001 & 0.093 \\
 & 400 & 0.01 & 0 & 400.691 & 0.283 & 8.004 & 0.003 & 1.000 & 0.021 \\
 & 1 & 400.836 & 0.288 & 8.002 & 0.003 & 1.000 & 0.023 \\
 & 2 & 400.161 & 0.302 & 8.004 & 0.003 & 1.000 & 0.023 \\
\end{array}
\]
\[
\eta_0 = 2, \eta_1 = 2.045, \eta_2 = 2.283
\]
\[
\begin{array}{cccccccccc}
 & 2 & 40 & 1 & 0 & 40.529 & 0.065 & 78.953 & 0.095 & 0.987 & 1.095 \\
 & 1 & 40.538 & 0.066 & 78.900 & 0.095 & 0.986 & 1.115 \\
 & 2 & 40.551 & 0.070 & 78.960 & 0.097 & 0.987 & 1.235 \\
 & 100 & 0.16 & 0 & 100.442 & 0.100 & 31.828 & 0.023 & 0.995 & 0.163 \\
 & 1 & 100.360 & 0.101 & 31.822 & 0.023 & 0.994 & 0.167 \\
 & 2 & 100.543 & 0.107 & 31.881 & 0.023 & 0.996 & 0.184 \\
 & 200 & 0.04 & 0 & 200.390 & 0.143 & 15.954 & 0.008 & 0.997 & 0.041 \\
 & 1 & 200.448 & 0.145 & 15.959 & 0.008 & 0.997 & 0.042 \\
 & 2 & 200.418 & 0.152 & 15.956 & 0.008 & 0.997 & 0.047 \\
 & 400 & 0.01 & 0 & 400.598 & 0.201 & 7.993 & 0.003 & 0.999 & 0.010 \\
 & 1 & 400.316 & 0.202 & 7.990 & 0.003 & 0.999 & 0.010 \\
 & 2 & 400.472 & 0.213 & 7.991 & 0.003 & 0.999 & 0.011 \\
\end{array}
\]
\[
\eta_0 = 1, \eta_1 = 1.023, \eta_2 = 1.142
\]
\[
\begin{array}{cccccccccc}
 & 3 & 40 & 1 & 0 & 40.545 & 0.052 & 78.646 & 0.093 & 0.983 & 0.672 \\
 & 1 & 40.582 & 0.053 & 78.686 & 0.093 & 0.984 & 0.687 \\
 & 2 & 40.596 & 0.055 & 78.849 & 0.092 & 0.986 & 0.739 \\
 & 100 & 0.16 & 0 & 100.562 & 0.083 & 31.813 & 0.023 & 0.994 & 0.110 \\
 & 1 & 100.561 & 0.083 & 31.819 & 0.023 & 0.994 & 0.111 \\
 & 2 & 100.595 & 0.086 & 31.839 & 0.023 & 0.995 & 0.119 \\
 & 200 & 0.04 & 0 & 200.490 & 0.116 & 15.943 & 0.008 & 0.996 & 0.027 \\
 & 1 & 200.451 & 0.116 & 15.943 & 0.008 & 0.996 & 0.027 \\
 & 2 & 200.438 & 0.123 & 15.948 & 0.008 & 0.997 & 0.030 \\
 & 400 & 0.01 & 0 & 400.371 & 0.163 & 7.987 & 0.003 & 0.998 & 0.007 \\
 & 1 & 400.366 & 0.165 & 7.985 & 0.003 & 0.998 & 0.007 \\
 & 2 & 400.413 & 0.174 & 7.985 & 0.003 & 0.998 & 0.008 \\
\end{array}
\]
\[
\eta_0 = 0.667, \eta_1 = 0.682, \eta_2 = 0.761
\]
Table 4.2. Simulations from $N(5, 4^2)$ under $\alpha = 0.05, m_0 = 10$
with $T \equiv 10000$ runs implementing methods (4.2.2)-(4.2.3)

<table>
<thead>
<tr>
<th>$\sigma_L$</th>
<th>$C$</th>
<th>$d$</th>
<th>$i$</th>
<th>$\bar{p}$</th>
<th>$s(\bar{p})$</th>
<th>$\bar{p}$</th>
<th>$s(\bar{p})$</th>
<th>$\bar{p}/C$</th>
<th>$\bar{p} - C$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td>1.1087</td>
<td>0</td>
<td>50.550</td>
<td>0.236</td>
<td>0.919</td>
<td>0.003</td>
<td>1.011</td>
<td>0.550</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>50.297</td>
<td>0.235</td>
<td>0.916</td>
<td>0.003</td>
<td>1.006</td>
<td>0.297</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>50.778</td>
<td>0.253</td>
<td>0.912</td>
<td>0.003</td>
<td>1.016</td>
<td>0.778</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.7840</td>
<td>0</td>
<td>100.733</td>
<td>0.473</td>
<td>0.914</td>
<td>0.003</td>
<td>1.007</td>
<td>0.733</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>100.501</td>
<td>0.483</td>
<td>0.914</td>
<td>0.003</td>
<td>1.005</td>
<td>0.501</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>101.031</td>
<td>0.506</td>
<td>0.912</td>
<td>0.003</td>
<td>1.010</td>
<td>1.031</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>0.5544</td>
<td>0</td>
<td>201.590</td>
<td>0.817</td>
<td>0.928</td>
<td>0.003</td>
<td>1.008</td>
<td>1.590</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>200.736</td>
<td>0.821</td>
<td>0.924</td>
<td>0.003</td>
<td>1.004</td>
<td>0.736</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>200.218</td>
<td>0.860</td>
<td>0.923</td>
<td>0.003</td>
<td>1.001</td>
<td>0.218</td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>0.3506</td>
<td>0</td>
<td>501.255</td>
<td>1.273</td>
<td>0.943</td>
<td>0.002</td>
<td>1.003</td>
<td>1.255</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>502.996</td>
<td>1.287</td>
<td>0.942</td>
<td>0.002</td>
<td>1.006</td>
<td>2.996</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>503.503</td>
<td>1.354</td>
<td>0.944</td>
<td>0.002</td>
<td>1.007</td>
<td>3.503</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>0.2479</td>
<td>0</td>
<td>1000.787</td>
<td>1.785</td>
<td>0.946</td>
<td>0.002</td>
<td>1.001</td>
<td>0.787</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1002.184</td>
<td>1.811</td>
<td>0.948</td>
<td>0.002</td>
<td>1.002</td>
<td>2.184</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1000.102</td>
<td>1.944</td>
<td>0.945</td>
<td>0.002</td>
<td>1.000</td>
<td>0.102</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| 2 | 50 | 1.1087 | 0 | 50.643 | 0.204 | 0.930 | 0.003 | 1.013 | 0.643 |
| | 1 | 50.495 | 0.208 | 0.931 | 0.003 | 1.010 | 0.495 |
| | 2 | 50.370 | 0.217 | 0.927 | 0.003 | 1.007 | 0.370 |
| 100 | 0.7840 | 0 | 100.165 | 0.286 | 0.938 | 0.002 | 1.002 | 0.165 |
| | 1 | 100.559 | 0.292 | 0.936 | 0.002 | 1.006 | 0.559 |
| | 2 | 100.994 | 0.309 | 0.933 | 0.003 | 1.010 | 0.994 |
| 200 | 0.5544 | 0 | 199.635 | 0.399 | 0.943 | 0.002 | 0.998 | -0.366 |
| | 1 | 200.346 | 0.404 | 0.942 | 0.002 | 1.002 | 0.346 |
| | 2 | 200.367 | 0.430 | 0.943 | 0.002 | 1.002 | 0.367 |
| 500 | 0.3506 | 0 | 501.199 | 0.635 | 0.947 | 0.002 | 1.002 | 1.199 |
| | 1 | 499.808 | 0.641 | 0.952 | 0.002 | 1.000 | -0.192 |
| | 2 | 500.253 | 0.676 | 0.950 | 0.002 | 1.001 | 0.253 |
| 1000 | 0.2479 | 0 | 1002.160 | 0.904 | 0.952 | 0.002 | 1.002 | 2.160 |
| | 1 | 1001.465 | 0.917 | 0.951 | 0.002 | 1.001 | 1.465 |
| | 2 | 1001.411 | 0.958 | 0.947 | 0.002 | 1.001 | 1.411 |
Table 4.2 Continued. Simulations from $N(5, 4^2)$ under $\alpha = 0.05$, $m_0 = 10$ with $T = 10000$ runs implementing methods (4.2.2)-(4.2.3)

<table>
<thead>
<tr>
<th>$\sigma_L$</th>
<th>$C$</th>
<th>$d$</th>
<th>$i$</th>
<th>$\bar{n}$</th>
<th>$s(\bar{n})$</th>
<th>$\bar{p}$</th>
<th>$s(\bar{p})$</th>
<th>$\bar{n}/C$</th>
<th>$\bar{n} - C$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>50</td>
<td>1.1087</td>
<td>0</td>
<td>50.562</td>
<td>0.133</td>
<td>0.939</td>
<td>0.002</td>
<td>1.011</td>
<td>0.562</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>50.549</td>
<td>0.133</td>
<td>0.940</td>
<td>0.002</td>
<td>1.011</td>
<td>0.549</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>50.580</td>
<td>0.139</td>
<td>0.941</td>
<td>0.002</td>
<td>1.012</td>
<td>0.580</td>
</tr>
<tr>
<td>100</td>
<td>0.7840</td>
<td>0</td>
<td>100.675</td>
<td>0.189</td>
<td>0.946</td>
<td>0.002</td>
<td>1.007</td>
<td>0.675</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>100.543</td>
<td>0.192</td>
<td>0.945</td>
<td>0.002</td>
<td>1.005</td>
<td>0.543</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>100.476</td>
<td>0.204</td>
<td>0.941</td>
<td>0.002</td>
<td>1.005</td>
<td>0.476</td>
</tr>
<tr>
<td>200</td>
<td>0.5544</td>
<td>0</td>
<td>200.632</td>
<td>0.266</td>
<td>0.951</td>
<td>0.002</td>
<td>1.003</td>
<td>0.632</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>200.242</td>
<td>0.270</td>
<td>0.948</td>
<td>0.002</td>
<td>1.001</td>
<td>0.242</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>200.512</td>
<td>0.286</td>
<td>0.947</td>
<td>0.002</td>
<td>1.003</td>
<td>0.512</td>
</tr>
<tr>
<td>500</td>
<td>0.3506</td>
<td>0</td>
<td>501.100</td>
<td>0.425</td>
<td>0.950</td>
<td>0.002</td>
<td>1.002</td>
<td>1.100</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>500.648</td>
<td>0.428</td>
<td>0.949</td>
<td>0.002</td>
<td>1.001</td>
<td>0.648</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>500.598</td>
<td>0.452</td>
<td>0.949</td>
<td>0.002</td>
<td>1.003</td>
<td>0.512</td>
</tr>
<tr>
<td>1000</td>
<td>0.2479</td>
<td>0</td>
<td>1000.530</td>
<td>0.594</td>
<td>0.948</td>
<td>0.002</td>
<td>1.001</td>
<td>0.530</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1000.842</td>
<td>0.607</td>
<td>0.948</td>
<td>0.002</td>
<td>1.001</td>
<td>0.842</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>1000.426</td>
<td>0.650</td>
<td>0.949</td>
<td>0.002</td>
<td>1.000</td>
<td>0.426</td>
</tr>
</tbody>
</table>
Table 4.3. Point estimators from horticulture data under MRPE methods (4.1.4)-(4.1.5) with $m$ from (4.1.2) with $\sigma_L = 2$

<table>
<thead>
<tr>
<th>$i$</th>
<th>$N$</th>
<th>$\bar{x}_{BS}$</th>
<th>$\bar{\sigma}_{BS}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>112</td>
<td>34.811</td>
<td>4.146</td>
</tr>
<tr>
<td>1</td>
<td>113</td>
<td>34.785</td>
<td>3.884</td>
</tr>
<tr>
<td>2</td>
<td>113</td>
<td>34.785</td>
<td>3.884</td>
</tr>
</tbody>
</table>

$A = 100, c = 0.1,$  
$m_0 = 10, b = 200$
Table 4.4. Confidence intervals from horticulture data under FCIE methods (4.2.2)-(4.2.3) with $\sigma_L = 2$

\begin{tabular}{cccccc}
\hline
$i$ & $N$ & $\bar{x}_{BS}$ & $\bar{\sigma}_{BS}$ & $\bar{\rho}_{BS}$ & $s(\bar{\rho}_{BS})$ \\
\hline
$\alpha = 0.05, d = 0.5, m_0 = 10, b = 200$ \\
0 & 196 & 35.027 & 3.818 & 0.940 & 0.0168 \\
1 & 198 & 35.073 & 3.908 & 0.925 & 0.0186 \\
2 & 199 & 35.087 & 3.979 & 0.930 & 0.0180 \\
\hline
\end{tabular}
Chapter 5

Second-Order Asymptotics in a Class of Purely Sequential MRPE Methodologies

As Mukhopadhyay (1982) had suggested using estimators of $\sigma^2$ other than the sample variance, for instance, Chattopadhyay and Mukhopadhyay (2013) incorporated the GMD, MAD and range to construct their two-stage FCIE methodologies. Following these ideas, we may consider other alternative estimators of $\sigma^2$ (or $\sigma$), propose new stopping boundary conditions, and construct even more robust multi-stage methodologies. Chapter 5 is based on the work Hu and Mukhopadhyay (2018).

In Section 5.1, we propose a new class of purely sequential MRPE methodologies, which involve general robust estimators of $\sigma$ in defining the stopping boundaries. Some appropriate sufficient conditions on such estimators are assumed.

In Section 5.2, we provide both asymptotic first-order and second-order properties enjoyed by the associated methodologies as our main results, summarized in Theorems 5.1.1-5.1.3.

In Section 5.3, we revisit the GMD-based and MAD-based purely sequential MRPE methodologies described earlier in Section 3.2 and develop asymptotic second-order results which are considerably stronger than those reported in Mukhopadhyay and Hu (2017). More specifically, for the purely sequential MRPE methodologies established in (3.2.1)-(3.2.2), we propose to derive second-order approximations of the regret $\omega_i(c), i = 0, 1, 2$. Corresponding simulated performances are provided to validate the second-order approximations of the regret and highlight the usefulness of the methodologies.
Section 5.4 summarizes the chapter with a few overall concluding thoughts.

5.1. A Formulation

Here, we focus on the purely sequential MRPE methodologies alone and take into consideration a class of general robust estimators of $\sigma$, say denoted by $V_n$ based on the observations given by $X_1, X_2, ..., X_n, n \geq 2$ independently from a normal distribution. We aim to find a set of sufficient conditions on $\{V_n, n \geq 2\}$ such that the associated methodologies involving $V_n$ will have a number of asymptotic first-order and second-order properties.

5.1.1. Sufficient Conditions on Robust Estimators

Assuming we have a sequence of independent observations $X_1, X_2, ...$ from a $N(\mu, \sigma^2)$ population with $-\infty < \mu < \infty$ and $0 < \sigma < \infty$, both unknown. Having recorded $X_1, X_2, ..., X_n, n \geq 2$, we use $V_n \equiv V_n(X_1, X_2, ..., X_n)$, assumed positive w.p.1, to denote an appropriate robust estimator of $\sigma$. The sequence of $\{V_n, n \geq 2\}$ satisfies the following set of conditions:

(C1) Independence.

$\overline{X}_n$ and $\{V_k, 2 \leq k \leq n\}$ are distributed independently for all $n \geq 2$.

(C2) Convergence in probability.

$V_n \overset{P}{\rightarrow} \sigma$, as $n \rightarrow \infty$.

(C3) Asymptotic normality.

$\sqrt{n}(\sigma^{-1}V_n - 1) \overset{D}{\rightarrow} N(0, \delta^2)$ for some $\delta^2(>0)$ as $n \rightarrow \infty$.

(C4) Uniform continuity in probability (u.c.i.p). For every $\varepsilon > 0$, there exists a large $\nu \equiv \nu(\varepsilon)$ and small $\gamma > 0$ for which

$P_{\mu, \sigma} \left( \max_{0 \leq k \leq n\gamma} |V_{n+k} - V_n| \geq \varepsilon \right) < \varepsilon$ holds for any $n \geq \nu$. 

81
(C5) Kolmogorov’s inequality. For every $\varepsilon > 0$, and some $n_2 \geq n_1 \geq 2$,

$$P_{\mu,\sigma} \left( \max_{n_1 \leq n \leq n_2} |V_n - \sigma| \geq \varepsilon \right) \leq \varepsilon^{-r} E_{\mu,\sigma} |V_{n_1} - \sigma|^r, \text{ with } r \geq 2.$$  

(C6) Order of central absolute moments. For $n \geq 2$, and $r \geq 2$,

$$E_{\mu,\sigma} [|V_n - \sigma|^r] = O(n^{-r/2}).$$

(C7) Wiener’s condition.

$$E_{\mu,\sigma} \left[ \sup_{n \geq 2} V_n \right] < \infty.$$

**Remark 5.1.** (C4) can be deduced from (C5) and (C6). Note that for any fixed $n_0 \geq \nu$,

$$P_{\mu,\sigma} \left( \max_{0 \leq k \leq n_0 \gamma} |V_{n_0+k} - V_{n_0}| \geq \varepsilon \right) \leq P_{\mu,\sigma} \left( \max_{0 \leq k \leq n_0 \gamma} \{|V_{n_0+k} - \sigma| + |V_{n_0} - \sigma|\} \geq \varepsilon \right) \leq P_{\mu,\sigma} \left( \max_{\nu \leq n \leq n_0 (1+\gamma)} |V_n - \sigma| \geq \varepsilon / 2 \right) + P_{\mu,\sigma} \left( \max_{\nu \leq n \leq n_0} |V_n - \sigma| \geq \varepsilon / 2 \right) \leq 2 (\varepsilon / 2)^{-r} E_{\mu,\sigma} |V_{\nu} - \sigma|^r = O \left( (\nu \varepsilon^2)^{-r/2} \right).$$

By choosing some large enough $\nu$ and appropriate $r$ such that $O \left( (\nu \varepsilon^2)^{-r/2} \right) < \varepsilon$, we can claim that the u.c.i.p property for $\{V_n, n \geq 2\}$ holds.

**5.1.2. Purely Sequential MRPE Methodologies Involving $V_n$**

Now we are in a position to propose a class of purely sequential methodologies based on $V_n$ for an unknown normal mean $\mu$ under the MRPE formulation (2.2.1)-(2.2.3) in the spirit of the methodology (3.2.1)-(3.2.2) in Section 3.2. Beginning with the pilot data $X_1, X_2, ..., X_m$ of size $m \geq 2$, we sample one additional observation at a time sequentially as needed until we terminate according to the following stopping rule:

$$N \equiv N(c) = \inf \left\{ n \geq m : n \geq \sqrt{\frac{1}{c} \left( V_n + n^{-\lambda} \right)} \right\}, \quad (5.1.1)$$
where \( \lambda(> 1/2) \) is held fixed. That is, if we have a pilot sample of size \( m \) such that \( m \geq \sqrt{A/c} (V_m + m^{-\lambda}) \) already holds, no additional observations will be recorded and the final sample size is \( N = m \). Otherwise, we record one more observation at a time and update \( n \) in the stopping rule (5.1.1). We terminate the sampling procedure at the first time that \( N = n (\geq m) \) is observed such that \( n \geq \sqrt{A/c} (V_n + n^{-\lambda}) \). Finally with \( \{N, X_1, ..., X_m, ..., X_N\} \), we establish the minimum risk point estimator for \( \mu \) as follows:

\[
\bar{X}_N \equiv N^{-1} \sum_{i=1}^{N} X_i. \tag{5.1.2}
\]

For the purely sequential MRPE methodology (5.1.1)-(5.1.2), clearly \( P_{\mu, \sigma} (N < \infty) = 1 \) and \( N \uparrow \infty \) w.p.1 as \( c \downarrow 0 \).

5.2. Main Results

In this section, we lay down a number of main lemmas and theorems associated with the purely sequential MRPE methodology given by (5.1.1)-(5.1.2) under conditions (C1)-(C7).

**Theorem 5.1.** For the stopping time \( N \) defined by the purely sequential estimation strategy (5.1.1)-(5.1.2), for all fixed \( \mu, \sigma, \) and \( A \), we have:

\[
\lim_{c \to 0} E_{\mu, \sigma} [N/n^*] = 1 \text{ [Asymptotic First-Order Efficiency].} \tag{5.2.1}
\]

**Proof.** By the stopping rule defined in (5.1.1), we have the following two inequalities (w.p.1):

\[
N \geq \sqrt{A/c} (V_N + N^{-\lambda}),
\]

and
\[ N - 1 < m - 1 + \sqrt{A/c} \left( V_{N-1} + (N-1)^{-\lambda} \right), \]

from which we conclude

\[ \frac{V_N + N^{-\lambda}}{\sigma} \leq \frac{N}{n^*} < \frac{V_{N-1} + (N-1)^{-\lambda}}{\sigma} + \frac{m}{n^*}. \] (5.2.2)

Then, it is clear that as \( c \to 0 \), \( N/n^* \xrightarrow{P_{\mu,\sigma}} 1 \). Next, note that for some sufficiently small \( c \), the right-hand side of (5.2.2) can be bounded as follows:

\[ \frac{N}{n^*} < \sigma^{-1}(\sup_{n \geq 2} V_n + 1) + 1. \]

Thus, under (C7) and by the dominated convergence theorem, \( \lim_{c \to 0} E_{\mu,\sigma} [N/n^*] = 1 \) holds immediately. \( \blacksquare \)

**Lemma 5.1.** For the stopping time \( N \) defined by the purely sequential estimation strategy (5.1.1)-(5.1.2), for all fixed \( \mu, \sigma, \) and \( A \), we have: for any arbitrary \( 0 < \eta < 1 \), with \( r \geq 2 \),

\[ P_{\mu,\sigma} (N \leq \eta n^*) = O \left( n^{* - \frac{r}{2(r+1)}} \right). \] (5.2.3)

**Proof.** Along the lines in the proof of Lemma 3.2.1, let \( \lfloor w \rfloor \) denote the largest integer that is smaller \((<)\) than \( w \) and we define:

\[ n_{1c} = \left\lfloor (A/c)^{\frac{1}{2(r+1)}} \right\rfloor = O(c^{-\frac{1}{2(r+1)}}) \text{ and } n_{2c} = \eta n^* = \eta \sigma \sqrt{A/c}, \] (5.2.4)

again. It should be obvious that \( N \geq n_{1c} \) w.p.1 from the definition of \( N \) in (5.1.1). Next,
we set out to obtain the rate at which \( P_{\mu,\sigma}\{N \leq \eta n^*\} \) may converge to zero for small \( c \):

\[
P_{\mu,\sigma}\{N \leq \eta n^*\} \\
\leq P_{\mu,\sigma}\{V_n \leq \eta \sigma \text{ for some } n \text{ such that } n_{1c} \leq n \leq n_{2c}\} \\
\leq P_{\mu,\sigma}\left\{ \max_{n_{1c} \leq n \leq n_{2c}} |V_n - \sigma| \geq (1 - \eta) \sigma \right\} \\
\leq \{(1 - \eta) \sigma\}^{-r} E_{\mu,\sigma} |V_{n_{1c}} - \sigma|^r = O\left(n_{1c}^{-r/2}\right) = O\left(n^* (n^*/(2(1+\lambda)))\right),
\]

by conditions (C5) and (C6). Thus, (5.2.3) holds. \(\Box\)

**Theorem 5.2.** For the stopping time \( N \) defined by the purely sequential estimation strategy (5.1.1)-(5.1.2), for all fixed \( \mu, \sigma, \text{ and } A \), we have:

\[
\lim_{c \to 0} \xi(c) = 1 \text{ [Asymptotic First-Order Risk Efficiency]}, \tag{5.2.5}
\]

where the term \( \xi(c) \) was defined in (3.2.6).

**Proof.** Note that under (C1), \( \xi(c) = \frac{1}{2} E_{\mu,\sigma}[N/n^*] + \frac{1}{2} E_{\mu,\sigma}[n^*/N] \). Picking \( \eta = \frac{1}{2} \) in Lemma 5.1, (5.2.5) can be justified in the same way as we proved Theorem 3.2.2. \(\Box\)

**Lemma 5.2.** For the stopping time \( N \) defined by the purely sequential estimation strategy (5.1.1)-(5.1.2), for all fixed \( \mu, \sigma, \text{ and } A \), we have:

\[
\frac{(N - n^*)}{N^{1/2}} \overset{\mathcal{L}}{\to} N(0, \sigma^2), \text{ as } c \to 0.
\]

**Proof.** Since \( N/n^* \overset{P_{\mu,\sigma}}{\to} 1 \) as \( c \to 0 \), we have \( \sqrt{n^*}/\sqrt{N} P_{\mu,\sigma} \overset{1}{\to} 1 \) as \( c \to 0 \). Then by Slutsky’s
theorem, it suffices to show

\[
\frac{N_n - n^*}{\sqrt{n}} \xrightarrow{d} N(0, \delta^2) \text{ as } c \to 0.
\]  \hfill (5.2.6)

Note that for sufficiently small \( c \), the following holds w.p.1:

\[
\sqrt{n^*} \left( \frac{V_N}{\sigma} - 1 \right) \leq \frac{N_n - n^*}{\sqrt{n}} \leq \frac{m}{\sqrt{n^*}} + \sqrt{n^*} \left( \frac{V_{N-1}}{\sigma} - 1 \right).
\]  \hfill (5.2.7)

By the conditions (C3) and (C4), as well as Anscombe’s (1952) Random Central Limit Theorem, we have

\[
\sqrt{n^*} \left( \frac{V_N}{\sigma} - 1 \right) \xrightarrow{d} N(0, \delta^2) \text{ and } \frac{m}{\sqrt{n^*}} + \sqrt{n^*} \left( \frac{V_{N-1}}{\sigma} - 1 \right) \xrightarrow{d} N(0, \delta^2),
\]

as \( c \to 0 \). Hence, the lemma follows from (5.2.7).

**Lemma 5.3.** For the stopping time \( N \) defined by the purely sequential estimation strategy (5.1.1)-(5.1.2), for all fixed \( \mu, \sigma, \) and \( A \), we have: \((N - n^*)^2 / N\) is uniformly integrable in \( c \leq c_0 \), for some \( c_0 \).

**Proof.** We prove this lemma in the spirit of Ghosh and Mukhopadhyay (1980) and Ghosh et al. (1997, Lemma 7.2.3, pp. 217-219). First we show \((N - n^*)^2 / n^*\) is uniformly integrable in \( c \leq c_0 \). Write for any \( b > b_0 + 1 \), \( b_0 = \left( \sigma \sqrt{A/c_1} \right)^{-1/2} \), where \( c_1 \) is some appropriate constant such that \( c \leq c_1 \). We have

\[
E_{\mu, \sigma} \left\{ \frac{(N_n - n^*)^2}{n^*} \mathbf{1} \left( \frac{(N_n - n^*)^2}{n^*} > b^2 \right) \right\} = 2 \int_b^\infty xP_{\mu, \sigma} (|N - n^*| > x\sqrt{n^*}) \, dx.
\]  \hfill (5.2.8)
Write $k_1 = \lceil n^* + x\sqrt{n^*} \rceil + 1$, for $x \geq b$. Then,

$$
P_{\mu,\sigma} \left( N > n^* + x\sqrt{n^*} \right)
\leq P_{\mu,\sigma} \left( k_1 - 1 \leq V_{k_1-1}\sqrt{A/c} \right) \leq P_{\mu,\sigma} \left( \frac{V_{k_1-1}}{\sigma} \geq \frac{n^* + x\sqrt{n^*} - 1}{n^*} \right)
\leq P_{\mu,\sigma} \left( \left| \frac{V_{k_1-1}}{\sigma} - 1 \right| \geq \frac{x\sqrt{n^*} - 1}{n^*} \right) \leq \frac{E_{\mu,\sigma} \left( V_{k_1-1} - 1 \right)^{2r_1}}{(x\sqrt{n^*} - 1)^{2r_1}},
$$

(5.2.9)

for some $r_1$. With (C5), we claim that there exists a $\lambda_1 (>) 0$ depending only on $r_1$ such that

$$
P_{\mu,\sigma} \left( N > n^* + x\sqrt{n^*} \right) \leq \lambda_1 (k_1 - 1)^{-r_1} n^{*r_1} \left( x - 1/\sqrt{n^*} \right)^{-2r_1}.
$$

(5.2.10)

Note that $(k_1 - 1)^{-r_1} n^{*r_1} < 1$, for $x \geq b > b_0 + 1 = \left( \sigma\sqrt{A/c_1} \right)^{-1/2} + 1$ and $n^* \geq \sigma\sqrt{A/c_1}$.

It follows that

$$
\int_b^\infty x P_{\mu,\sigma} \left( |N - n^*| > x\sqrt{n^*} \right) dx \leq \lambda_1 \int_b^\infty x (x - b_0)^{-2r_1} dx
= \lambda_1 \frac{(b-b_0)^{1-2r_1}((3-2r_1)b-b_0)}{(1-2r_1)(2-2r_1)} \to 0
$$

(5.2.11)

as $b \to \infty$ by choosing $r_1 > 1$ appropriately.

Next, note that if $\sqrt{n^*} \leq b$,

$$
\int_b^\infty x P_{\mu,\sigma} \left( N - n^* < -x\sqrt{n^*} \right) dx = 0.
$$

(5.2.12)

If $\sqrt{n^*} > b$, there exists some $0 < \gamma < 1$ such that $(1 - \gamma) \sqrt{n^*} > b$, when $c \leq c_2$, for some technically picked $c_2$. Then,

$$
\int_b^\infty x P_{\mu,\sigma} \left( N - n^* < -x\sqrt{n^*} \right) dx
\leq \int_b^{\sqrt{n^*}} x P_{\mu,\sigma} \left( N \leq \gamma n^* \right) dx + \int_b^{(1-\gamma)^{\sqrt{n^*}}} x P_{\mu,\sigma} \left( \gamma n^* < N < n^* - x\sqrt{n^*} \right) dx.
$$

(5.2.13)
By Lemma 5.1, \( P_{\mu, \sigma}(N \leq \gamma n^*) \leq \lambda_2 n^{\gamma - \frac{r_2}{2(1+\lambda)}} \), for some appropriate \( r_2 (> 2 + 2\lambda) \) and \( \lambda_2 (> 0) \) depending on \( r_2 \) alone. Hence,

\[
\int_{b^\gamma n^*}^x P_{\mu, \sigma}(N \leq \gamma n^*) \, dx \leq \lambda_2 b^{2 - \frac{r_2}{2(1+\lambda)}} \to 0 \text{ as } b \to \infty. \tag{5.2.14}
\]

As for \( b \leq x \leq (1 - \gamma) \sqrt{n^*} \), write

\[
k_2 = \lfloor \gamma n^* \rfloor + 1 \text{ and } k_3 = \lceil n^* - x\sqrt{n^*} \rceil.
\]

We obtain that

\[
P_{\mu, \sigma}(\gamma n^* < N < n^* - x\sqrt{n^*})
= P_{\mu, \sigma} \left( \bigcup_{n=k_2}^{k_3} \{ N = n \} \right) \tag{5.2.15}
\leq P_{\mu, \sigma} \left( \bigcup_{n=k_2}^{k_3} \{ \frac{V_n}{\sigma} < \frac{n}{n^*} \} \right).
\]

But for a small \( c \), say \( c \leq c_3 \), for some \( c_3 \),

\[
\frac{n}{n^*} \leq \frac{k_3}{n^*} \leq \frac{n^* - x\sqrt{n^*}}{n^*} = 1 - \frac{x}{\sqrt{n^*}}.
\]

Hence, it follows from (5.2.15) that

\[
P_{\mu, \sigma}(\gamma n^* < N < n^* - x\sqrt{n^*})
\leq P_{\mu, \sigma} \left( \frac{V_n}{\sigma} - 1 < -\frac{x}{\sqrt{n^*}} \right), \text{ for some } k_2 \leq n \leq k_3
\leq P_{\mu, \sigma} \left( \left| \frac{V_n}{\sigma} - 1 \right| > \frac{x}{\sqrt{n^*}} \right), \text{ for some } k_2 \leq n \leq k_3 \tag{5.2.16}
\leq P_{\mu, \sigma} \left( \max_{k_2 \leq n \leq k_3} \left| \frac{V_n}{\sigma} - 1 \right| > \frac{x}{\sqrt{n^*}} \right)
\leq E_{\mu, \sigma} \left| \frac{V_{k_2+1}}{\sigma} - 1 \right|^{2r_3} \leq \lambda_3 k_2^{-r_3} x^{-2r_3} n^{r_3} \leq \lambda_4 x^{-2r_3},
\]

for some appropriate \( \lambda_3 (> 0) \) and \( \lambda_4 (> 0) \), both depending only on \( r_3 \). Choosing \( r_3 > 1 \),

88
we get from (5.2.16) that
\[
\int_b^{(1-\gamma)\sqrt{n^*}} x P_{\mu,\sigma} \left( \gamma n^* < N < n^* - x \sqrt{n^*} \right) \, dx \\
\leq \lambda_4 \int_b^{(1-\gamma)\sqrt{n^*}} x^{-2r_3} \, dx \leq \frac{\lambda_4}{2-2r_3} b^{2-2r_3} \to 0 \text{ as } b \to \infty.
\] (5.2.17)

Now choosing \( c_0 = \min \{c_1, c_2, c_3\} \), with (5.2.11), (5.2.14) and (5.2.17), we can prove the uniform integrability of \((N - n^*)^2/n^*\) in \( c \leq c_0 \). To complete the proof of the lemma, observe that
\[
E_{\mu,\sigma} \left\{ \frac{(N-n^*)^2}{n^*} 1 \left( \frac{(N-n^*)^2}{n^*} > b^2 \right) 1 \left( N > \frac{1}{2} n^* \right) \right\} \\
\leq 2 E_{\mu,\sigma} \left\{ \frac{(N-n^*)^2}{n^*} 1 \left( (N-n^*)^2 > \frac{1}{2} b^2 n^* \right) \right\} \to 0
\] (5.2.18)
as \( b \to \infty \) uniformly in \( c \leq c_0 \). Furthermore, choosing \( b > n^* \), it follows that for \( c \leq c_0 \),
\[
E_{\mu,\sigma} \left\{ \frac{(N-n^*)^2}{n^*} 1 \left( \frac{(N-n^*)^2}{n^*} > b^2 \right) 1 \left( N \leq \frac{1}{2} n^* \right) \right\} \\
\leq m^{-1} n^{2r_4} P_{\mu,\sigma} \left( N \leq \frac{1}{2} n^* \right) \leq \lambda_5 n^{2r_4} \leq \lambda_5 b^{2-r_4} \to 0
\] (5.2.19)
as \( b \to \infty \), with \( r_4 (> 2) \) appropriately chosen and some \( \lambda_5 (> 0) \) depending only on \( r_4 \). In view of (5.2.18) and (5.2.19), the lemma holds. ■

Note that under the condition (C1), \( \omega (c) = c E_{\mu,\sigma} [(N - n^*)^2/N] \). As an immediate result from Lemmas 5.2 and 5.3, the following asymptotic second-order property holds, summarized in Theorem 5.3.

**Theorem 5.3.** For the stopping time \( N \) defined by the purely sequential estimation strategy (5.1.1)-(5.1.2), for all fixed \( \mu, \sigma \), and \( A \), we have:
\[
\omega (c) = \delta^2 c + o (c) \text{ [Asymptotic Second-Order Risk Efficiency]},
\] (5.2.20)
as \( c \to 0 \), where the term \( \omega (c) \) was defined in (3.2.6).
5.3. An Application: Purely Sequential Methodology (3.2.1)-(3.2.2)

In this section, we review the purely sequential methodology (3.2.1)-(3.2.2) as a possible application of the methodology (5.1.1)-(5.1.2) newly proposed in this chapter. By verifying the conditions (C1)-(C7), we further claim the asymptotic second-order risk efficiency of the methodology (3.2.1)-(3.2.2) beyond the asymptotic first-order properties concluded in Section 3.2.1. Simulations are provided as a reasonable validation, as well.

5.3.1. Second-Order Approximations of $\omega_i(c), i = 0, 1, 2$

We give the second-order approximations of the regret $\omega_i(c), i = 0, 1, 2$ from the purely sequential methodology (3.2.1)-(3.2.2) in Corollary 5.1.

**Corollary 5.1.** For the stopping time $N_i$ defined by the purely sequential estimation strategy (3.2.1)-(3.2.2), for all fixed $\mu, \sigma$, and $A$, we have:

$$\omega_i(c) = \delta_i^2 c + o(c), i = 0, 1, 2 \ [\text{Asymptotic Second-Order Risk Efficiency}] \quad (5.3.1)$$

as $c \to 0$, where the term $\omega_i(c)$ was defined in (3.2.6), and $\delta_0^2 = \frac{1}{2}, \delta_1^2 = \frac{1}{3}(\pi + 6\sqrt{3} - 12), \delta_2^2 = \frac{1}{2}(\pi - 2)$ remain the same as in Lemma 4.1.3.

**Proof.** It suffices to show that (C1)-(C7) are satisfied in terms of $V_{n;i}, i = 0, 1, 2$ defined in (2.3.4), respectively. Among the conditions, (C1) and (C7) follow from Mukhopadhyay and Hu (2017, Lemmas 3.1 and 3.3) with minor modifications. (C2) can be simply verified by slightly modifying the proof of Lemma 3.1.2. In the spirit of the proof of Lemma 4.1.3, we have (C3). Both (C5) and (C6) have been indicated in the proof of Lemma 3.2.1. And (C4) comes from combining (C5) and (C6).
Thus, the corollary follows. One may refer to Mukhopadhyay and de Silva (2009), Sen and Ghosh (1981), Mukhopadhyay and Hu (2017,2018) and Hu and Mukhopadhyay (2018) for more details.

5.3.2. Simulations and Data Analysis

Based on what we did in Section 3.2.2, we had considered \( m = 15 \) alone for brevity as we found little to no difference in the overall performances whether the pilot sample size was fixed to be 10 or 15. Besides the notation explained in (3.2.21), we included the following entities in (5.3.2) as well.

\[
\begin{align*}
    s(\hat{\xi}) & : \text{estimated standard error of } \hat{\xi}; \\
    s(\hat{\omega}) & : \text{estimated standard error of } \hat{\omega}; \\
    \delta_i^2 & : \text{a theoretical approximation of } \omega_i(c), i = 0, 1, 2.
\end{align*}
\]

Table 5.1 summarizes selected findings \((m = 15)\) in the spirit of Table 4.1. One should note that within each block of Table 5.1, the theoretical second-order approximations of \( \omega_i(c), i = 0, 1, 2 \) turned out close to the estimated regret values. These empirically validate Corollary 5.1 about the asymptotic second-order risk efficiency property from (5.3.1) of the purely sequential MRPE methodology (3.2.1)-(3.2.2).

5.4. Overall Concluding Thoughts

In summary, the purely sequential MRPE methodology (5.1.1)-(5.1.2) opens the possibility to incorporate a class of robust estimators of \( \sigma \) in defining the stopping boundary conditions without loss of the second-order properties. We have shown that the associated methodology is second-order risk efficient, as long as the robust estimators satisfy the
conditions given by (C1)-(C7). This indicates the purely sequential MRPE methodology (5.1.1)-(5.1.2) will have a much wider application.
Table 5.1. Simulations under $A = 100$, $m = 15$ with 1000 runs implementing methods (3.2.1)-(3.2.2) with $i$ from (3.2.20) with $\lambda = 2$

<table>
<thead>
<tr>
<th>$n^*$</th>
<th>$\bar{n}$</th>
<th>$\bar{R}_n$</th>
<th>$\hat{\xi}$</th>
<th>$\hat{\omega}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c$ i</td>
<td>s ($\bar{n}$)</td>
<td>s ($\bar{R}_n$)</td>
<td>s ($\hat{\xi}$)</td>
<td>s ($\hat{\omega}$)</td>
</tr>
<tr>
<td>$\delta_0^2 c = 8 \times 10^{-2}, \delta_1^2 c = 8.18 \times 10^{-2}, \delta_2^2 c = 9.13 \times 10^{-2}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50 0 15.801 0.988 9.34 $\times$ 10$^{-2}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.16 0.167 0.053 0.033 5.07 $\times$ 10$^{-3}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 50.460 15.839 0.990 9.49 $\times$ 10$^{-2}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 0.168 0.053 0.033 4.65 $\times$ 10$^{-3}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 50.457 15.843 0.990 9.79 $\times$ 10$^{-2}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 0.171 0.053 0.033 5.00 $\times$ 10$^{-3}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta_0^2 c = 2 \times 10^{-2}, \delta_1^2 c = 2.05 \times 10^{-2}, \delta_2^2 c = 2.28 \times 10^{-2}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100 0 100.159 7.941 0.993 2.15 $\times$ 10$^{-2}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.04 0.229 0.018 0.023 1.05 $\times$ 10$^{-3}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 100.239 7.947 0.993 2.23 $\times$ 10$^{-2}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 0.233 0.018 0.023 1.08 $\times$ 10$^{-3}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 100.245 7.944 0.993 2.55 $\times$ 10$^{-2}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 0.250 0.019 0.024 1.27 $\times$ 10$^{-3}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta_0^2 c = 5 \times 10^{-3}, \delta_1^2 c = 5.11 \times 10^{-3}, \delta_2^2 c = 5.71 \times 10^{-3}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200 0 200.199 3.986 0.997 5.46 $\times$ 10$^{-3}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01 0.328 0.007 0.016 2.54 $\times$ 10$^{-4}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 200.201 3.987 0.997 5.80 $\times$ 10$^{-3}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 0.338 0.007 0.017 2.81 $\times$ 10$^{-4}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 200.103 3.989 0.997 6.16 $\times$ 10$^{-3}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 0.346 0.007 0.016 3.27 $\times$ 10$^{-4}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Chapter 6

A Brief Summary

Purely sequential and Stein-type two-stage methodologies involving robust estimators such as GMD and MAD have been developed for both fixed-width confidence interval and minimum risk point estimation problems. With the underlying normal population, we assumed the condition that there exists an known positive lower bound of the variance so that the Stein-type methodologies will then enjoy desirable second-order properties, the purely sequential methodologies seem to automatically satisfy the higher order properties. The first-order properties have been proved, and the second-order properties are in progress.

Compared with the customary multi-stage methodologies based on the sample variance as an unbiased estimator of the population variance, our newly proposed multi-stage methodologies are more robust, when the normality of the underlying population is contaminated mildly. Therefore, our methodologies will have a wider applicability for practical purposes.
Chapter 7

Future Work

We have seen the significance of GMD or MAD, as a more robust alternative estimator than the sample variance. There are a lot of directions that we can follow in the future to make our research more broadly applicable in a number of other substantial inference problems. The potential future research may include the following areas.

7.1. Other Multi-Stage Methodologies

We may develop other multi-stage methodologies based on GMD and MAD, for example three-stage and accelerated sequential sampling, and compare them with the existing multi-stage methodologies and highlight the benefits of one approach over another.

While the Stein’s (1945, 1949) two-stage methodology is quick to terminate, it suffers from the serious oversampling problem, especially when the final sample size is large. Researchers have looked into the sampling designs where an additional stage is imposed so that the methodology will preserve both the first-order and second-order properties. In this spirit of combining procedure efficiency as well as sampling operational savings, the three-stage estimation methodologies have been established.

Mukhopadhyay (1976) first introduced the idea of triple sampling. Later on, Hall (1981) proposed a systematic three-stage fixed-width confidence interval estimation methodology for a normal mean when the population variance is unknown, where sample variance was utilized as an unbiased estimator for the population variance in defining stopping boundaries.
in each stage. He gave a rigorous account of the first-order and second-order properties. Mukhopadhyay (1990) considered more general cases and put forward a unified three-stage methodology. He discussed both the FCIE and MRPE problems with applications. We should mention there is a wide range of literature on three-stage estimation methodologies that one may refer to, including Mukhopadhyay (1985), Mukhopadhyay et al. (1987), Woodroofe (1987), Hamdy (1988), Hamdy et al. (1988), and Liu and Wang (2007).

On the other hand, researchers have been modifying the purely sequential estimation methodology and considering accelerated sequential stopping rules as well. Originally proposed in Hall (1983), the accelerated sequential methodology was further discussed in Mukhopadhyay and Solanky (1991) and Mukhopadhyay (1996).

Here, we may therefore follow this direction and establish both the three-stage and accelerated sequential estimation methodologies based on GMD or MAD correspondingly. We hope to show these methodologies are potentially more robust without loss of the first-order and second-order properties.

7.2. Distribution-Free Methodologies

Recall that Section 3.1.3 only provided the limited robustness of our newly proposed purely sequential estimation methodologies based on GMD or MAD, as we had assumed the population should follow a mixture normal distribution (3.1.7) instead of a pure normal distribution. Mukhopadhyay (1978), Ghosh and Mukhopadhyay (1979), Chow and Yu (1981) and other work set out to study the performances of multi-stage methodologies when the population distribution was unspecified. See Sen (1981,1985), Ghosh et al. (1997), and Mukhopadhyay and de Silva (2009) for a thorough review.

We may continue to discuss the GMD-based or MAD-based multi-stage methodologies
in the distribution-free cases, where stronger robustness is highly likely to be addressed.

7.3. Multivariate Multi-Stage Estimation

Schechtman and Yitzhaki (1987) proposed a new measure of association between two random variables based on GMD, which opened the possibility of extending our new GMD-based multi-stage methodologies (i) from univariate estimation to multivariate estimation or (ii) from one-sample estimation to multi-sample estimation. We may also investigate the MAD case so that our methodologies will have a significantly broader appeal.
# List of Tables

<table>
<thead>
<tr>
<th>Table Number</th>
<th>Page Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 3.1</td>
<td>43</td>
</tr>
<tr>
<td>Table 3.2</td>
<td>44</td>
</tr>
<tr>
<td>Table 3.3</td>
<td>45</td>
</tr>
<tr>
<td>Table 3.4</td>
<td>46</td>
</tr>
<tr>
<td>Table 3.5</td>
<td>47</td>
</tr>
<tr>
<td>Table 3.6</td>
<td>48</td>
</tr>
<tr>
<td>Table 3.7</td>
<td>49</td>
</tr>
<tr>
<td>Table 3.8</td>
<td>50</td>
</tr>
<tr>
<td>Table 3.9</td>
<td>51</td>
</tr>
<tr>
<td>Table 3.10</td>
<td>52</td>
</tr>
<tr>
<td>Table 3.11</td>
<td>53</td>
</tr>
<tr>
<td>Table 4.1</td>
<td>75</td>
</tr>
<tr>
<td>Table 4.2</td>
<td>76</td>
</tr>
<tr>
<td>Table 4.3</td>
<td>78</td>
</tr>
<tr>
<td>Table 4.4</td>
<td>79</td>
</tr>
<tr>
<td>Table 5.1</td>
<td>93</td>
</tr>
</tbody>
</table>
Bibliography


Winner, L. Personal Website http://www.stat.ufl.edu/~winner/data/ret_earn.dat


