7-22-2020

Sequential Estimation Methodologies with Observations Gathered in Groups: Theory, Practice and Data Analysis

Zhe Wang
University of Connecticut - Storrs, zhe.4.wang@uconn.edu

Sequential Estimation Methodologies with Observations Gathered in Groups: Theory, Practice and Data Analysis

Zhe Wang, Ph.D.

University of Connecticut, 2020

ABSTRACT

Purely sequential procedure has been widely studied in different inference problems. However, in purely sequential procedure, only one observation should be taken at-a-time. In real life, packaged items purchased in bulk often cost less per unit sample than the cost of an individual item. This dissertation discussed this situation when observations are gathered in groups. First, two fundamental problems on purely sequential estimation are revisited: (i) the fixed-width confidence interval (FWCI) estimation problem, and (ii) the minimum risk point estimation (MRPE) problem, in the context of estimating an unknown mean (μ) in a normal population having an unknown variance (σ^2). We begin by laying down general frameworks for the second-order asymptotic analyses, in both problems, under sequential sampling of one observation at-a-time. Then, we consider sequentially sampling k observations at-a-time in defining our proposed estimation strategies. In the first attempt, tentative estimators are used to study the feasibility. Then, replace the simple class of estimators with more complicated unbiased and consistent estimators under permutations within each group. These new estimators incorporated in the definition of the stopping boundaries have led to tighter estimation of requisite optimal fixed-sample sizes. In both scenarios, first-order and second-order asymptotic properties have been analyzed under appropriate requirements on the pilot sample size.

Such estimators can also be used in two-sample comparisons. The last part of this dissertation presents the second-order asymptotic properties for comparing treatment means. Two separate situations are considered: (i) $\sigma_1 = \sigma_2 = \sigma$, but σ is assumed unknown, and (ii) $\sigma_1 \neq \sigma_2$ are unknown. For datasets with possible outliers, robust estimators are in use in purely sequential estimation strategies. For each problem, large-scale computer simulations and substantial data analysis have validated corresponding results. The methodologies are illustrated with the help of real-world data.

Sequential Estimation Methodologies with Observations Gathered in Groups: Theory, Practice and Data Analysis

Zhe Wang

B.S., Beijing Normal University, Beijing, China, 2016

M.S., University of Connecticut, CT, USA, 2018

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2020

APPROVAL PAGE

Doctor of Philosophy Dissertation

Sequential Estimation Methodologies with Observations Gathered in Groups: Theory, Practice and Data Analysis

Presented by

Zhe Wang, B.S., M.S.

Major Advisor ______________________________

Nitis Mukhopadhyay

Associate Advisor ______________________________

Joseph Glaz

Associate Advisor ______________________________

Haim Bar

University of Connecticut

2020

ACKNOWLEDGEMENTS

Covid-19 pandemic changes everything in the year 2020. But with the love and support from everyone around me, I feel so blessed to finish my program smoothly during this pandemic. This special experience has made the dissertation such meaningful.

I would like to first express my deepest gratitude to my major advisor, Professor Nitis Mukhopadhyay. I was in his class on my first day at UConn. He is not only an academic advisor, but also a remarkable mentor and devoted friend. I was his "only child" for a long time. As he said, he raised me with tender loving care. I appreciated all the opportunities to learn how to be a qualified researcher and professor. I want to thank him for guiding me through these years of study, and leading me on my way of being a good person.

I am also grateful to Professor Joseph Glaz and Professor Haim Bar for serving on my dissertation committee. I'm so inspired by Professor Glaz's smiling face. Professor Bar was very helpful and supportive of my academic job search. He kept tracking on my progress and always gave practical suggestions. My sincere thanks also go to Professor Vladimir Pozdnyakov and Professor Rick Vitale, for their willingness to serve on my general exam committee, and for their time, support and guidance on my study.

I would like to thank Ann Marie Carroll and Sarah Trallero from the Department of Physics for providing me TAship. My teaching and communication skill was improved significantly during that year. I'm also very thankful to Professor Minghui Chen and Professor Zhiyi Chi for providing me independent teaching opportunities. I found my passion in teaching here. Two years of teaching gives me a sense of fulfillment, and help me a lot in my career.

I extend my special thanks to Professor Malaquías Pena from the School of Engineering. I cherished the opportunity to collaborate with Professor Pena in Eversource Energy Center. He always trusted me and encouraged me to explore new applications. I really appreciate this valuable experience.

None of these would be possible without the support from my family. Many thanks to my beloved parents Xiuhua Zhang and Linghua Wang. I also thank my late grandfather. He planted the faith deep in my heart when I was young. Lastly, I'd like to thank my husband Jinjian Mu, who is my competitor and collaborator in research, as well as best friend and soulmate in life. Our wedding at UConn is my most beautiful memory.

Table of Contents

Chapter 1

Introduction

The concepts of sequential analysis go back many years. Ghosh(1991) traced the conception of this subject back to the sequential sampling inspection procedures introduced by Dodge and Romig (1929), and to the quality control charts developed by Shewhart (1931), and to the two-stage designs of Thompson (1933). Later on, Wald (1948) and Barnard (1946) independently introduced the Sequential Probability Ratio Test (SPRT). Many classical problems came out in the period from 1951 to 1990. Sequential estimation problem stands in the center of all inference problems. The development of sequential estimation provides innovative ideas in many of their applications. In the twenty-first century, new topics and challenges arise including sequential adaptive design and online change point detection. Because of the broad scope of this subject and its vast and multifarious literature, this chapter can only focus on a limited number of topics. One will easily find more details by reviewing the fundamentals from Ghosh (1970), Ghosh and Sen (1991), Ghosh (1997), and Lai (2001).

In Section 1.1, we will revisit two fundamental problems, namely, fixed-width confidence interval problem and minimum risk point estimation problem. Broader literature review can be find in Section 1.2. Sequential analysis on non-Normal distribution and multivariate distribution will be discussed. Then, Section 1.3 gives a summary of nonlinear renewal theory, which will be applied in the following section without proof. Section 1.4 provides the motivation and layout of the thesis.

1.1. Two Fundamental Problems in Sequential Analysis

1.1.1. Fixed-Width Confidence Interval (FWCI) Problem

Having fixed two preassigned numbers, $d(>0)$ and $0 < \alpha < 1$, the problem is estimating μ with a confidence interval J such that (i) the length (or width) of J is $2d$, that is fixed in advance, *and* (ii) the confidence coefficient associated with J, that is, $P_{\boldsymbol{\theta}}\{\mu \in J\} \geq 1-\alpha$ for all fixed $\boldsymbol{\theta}$.

Having recorded $X_1, ..., X_n$ which are i.i.d. random variables from $N(\mu, \sigma^2)$, as a simple-minded approach, we may pretend to record a fixed number of n observations and let

$$J \equiv J_n = [\overline{X}_n - d, \overline{X}_n + d]. \tag{1.1.1}$$

The associated confidence coefficient is expressed as

$$P_\theta\{\mu \in J_n\} = P_\theta\left\{|\overline{X}_n - \mu| \leq d\right\} = 2\Phi\left(\sqrt{n}d/\sigma\right) - 1 \tag{1.1.2}$$

Observe that J_n already has the preassigned width $2d$, but we also require that the associated confidence coefficient be approximately at least $1-\alpha$. Thus, we must have:

$$2\Phi\left(\sqrt{n}d/\sigma\right) - 1 \geq 1 - \alpha = 2\Phi(a) - 1, \text{ say,} \tag{1.1.3}$$

where a is the upper $50\alpha\%$ point of a standard normal distribution.

From (1.1.3), we claim:

$$\begin{gathered}\text{the optimal fixed sample size, had } \sigma^2 \text{ been known, would be} \\ \text{the smallest integer } n \geq a^2\sigma^2 d^{-2} \equiv C_1, \text{ say.}\end{gathered} \tag{1.1.4}$$

However, the magnitude of C_1 remains unknown since σ^2 is unknown. Indeed no fixed-sample-size methodology (that is, fixing n in advance) would be able to come up with a solution for this problem regardless of whether or not the confidence interval is centered (or not centered) at $\overline{X}_n$ or another chosen estimator of μ. Dantzig (1940) proved that fundamental non-existential result. One may also refer to Sen (1981), Woodroofe (1982), Siegmund (1985), Ghosh et al. (1997, Chapter 3), Mukhopadhyay and Solanky (1994), Mukhopadhyay (2000, Chapter 13), Mukhopadhyay et al. (2004), Mukhopadhyay and de Silva (2009, Chapter 2), Zacks (2009,2017) and other sources.

Mahalanobis(1940) first introduced the idea of using pilot sample in large-scale survey to control the margin of error. Stein (1945,1949) developed a pathbreaking two-stage sampling strategy to provide an exact solution for this problem. We begin with the pilot observations $X_1, ..., X_{m_0}$, $m_0 \geq 2$. The final sample size is:

$$N \equiv N(d) = \max\left\{m_0, \left\langle \frac{a^2_{m_0-1} S^2_{m_0}}{d^2} \right\rangle + 1\right\} \tag{1.1.5}$$

where $\langle u \rangle$ is the largest integer smaller than u and a_{m_0-1} is the upper $50\alpha\%$ point of the Student's t distribution with $m_0 - 1$ degree of freedom. In two-stage procedure, we estimate the unknown population variance σ^2 only once. However, one may estimate σ^2 successively in a sequential manner. Anscombe (1952,1953), Ray (1957), Chow and Robbins

(1965), and Starr (1966a) introduced purely sequential sampling strategies by recording one observation at-a-time to achieve the nearly preassigned confidence coefficient $1-\alpha$, but only *asymptotically* (as $d \to 0$). One may further refer to a number of sources including Ghosh et al. (1997, Chapter 6) and Mukhopadhyay and de Silva (2009, Section 6.2). The stopping time is defined as follows:

$$N \equiv N(d) = \inf\{n \geq m_0(\geq 2): \quad n \geq a^2 S_n^2/d^2\} \tag{1.1.6}$$

Mukhopadhyay (1980) developed the modified two-stage procedure, which captured the best properties of both two-stage and sequential procedure. Recall that $C_1 \equiv a^2\sigma^2 d^{-2}$, we first choose and fix $0 < \gamma < \infty$ and define the pilot sample size

$$m_0 = \max\left\{2, \left\langle (a/d)^{2/(1+\gamma)} \right\rangle\right\}$$

Once m_0 is determined, we define the final sample size using (1.1.6). Mukhopadhyay (1976) then introduced the three-stage procedure. Hall (1981) showed that a suitable chosen three-stage procedure can give a fixed-width interval with similar second-order properties in comparison with the purely sequential procedure. Mukhopadhyay (1990) developed a unified framework for general triple sampling procedures.

1.1.2. Minimum Risk Point Estimation(MRPE) Problem

Motivated by a biological application involving estimating the small portion p of some attribute, Haldane (1945) published a seminal paper on bounded-risk sequential estimation.

Robbins (1959) formulated the minimum risk point estimation problem, and proposed the purely sequential procedure. The detailed study of Robbins procedure can be found in Starr (1966). The idea is summarized as follows.

Having recorded $X_1, ..., X_n$ which are i.i.d. random variables from $N(\mu, \sigma^2)$, suppose that the loss function in estimating μ by the sample mean, $\overline{X}_n$, is given by:

$$L_n \equiv L_n(\mu, \overline{X}_n) = A\left(\overline{X}_n - \mu\right)^2 + c_1 n, \tag{1.1.7}$$

where both A and c_1 are known positive constants. Here, c_1 represents the cost per unit observation drawn one at-a-time due to sampling and A, a weight function, representing a kind of rate of exchange between the magnitude of the loss due to estimation error and the cost for making such an error.

The goal is to minimize the associated fixed-sample-size risk function which is expressed as:

$$R_n(c_1) \equiv E_{\boldsymbol{\theta}}\left[L_n(\mu, \overline{X}_n)\right] = A\sigma^2 n^{-1} + c_1 n, \tag{1.1.8}$$

for all $0 < \sigma < \infty$. This risk is (nearly) minimized when we determine the requisite sample size as follows:

$$\begin{gathered} n \text{ is the smallest integer} \geq (A/c_1)^{1/2}\sigma = n_1^*, \text{ say, with associated minimum} \\ \text{fixed-sample-size risk given by } R_{n^*} \equiv R_{n^*}(c_1) = 2c_1 n_1^*. \end{gathered} \tag{1.1.9}$$

However, the magnitude of n_1^* remains unknown since σ^2 is unknown. Indeed no fixed-sample-size methodology (that is, fixing n in advance) would come up with a solution for this problem regardless of whether or not the point estimator involves the sample mean $\overline{X}_n$ or another estimator of choice. Again, Dantzig (1940) proved this fundamental result. One may refer to other sources including Ghosh et al. (1997, Chapter 3), Mukhopadhyay (2000, Chapter 13), and Mukhopadhyay and de Silva (2009, Chapter 2).

Robbins (1959) proposed the following stopping rule for the purely sequential procedure:

$$N \equiv N(c_1) = \inf\left\{n \geq m_0(\geq 2): \quad n \geq (A/c_1)^{1/2} S_n\right\}, \tag{1.1.10}$$

where S_n is the sample standard deviation.

1.2. Broader Literature Review

In Section 1.1, we have already mentioned a number of fundamental references based on normal distribution. In this section, we will give brief broader review on the other area of sequential analysis.

1.2.1. Sequential Analysis on Non-Normal Distribution

The sequential estimation methodology can be applied in a broad range of distributions. For example, negative exponential distribution, referred to as $NExpo(\theta, \sigma)$, and assume that θ and σ are unknown parameters. This distribution is widely used to model the failure times of complex equipment. The $NExpo(\theta, \sigma)$ model has also been used in areas of soil science and weed propagation. To obtain the fixed-width confidence interval estimation for θ, Ghurye (1958) developed a two-stage procedure using the maximum likelihood estimator $\widehat{\theta}_{MLE} = X_{n:1}$. Mukhopadhyay (1974) proposed the purely sequential procedure alone the lines of Chow and Robbins (1965). Costanza et al. (1986) developed two-stage procedure to estimate the common location parameters of several negative exponential distributions.

Exponential distribution, referred to as $Expo(\theta)$ where θ is the unknown mean of the distribution. θ is interpreted as the *mean time to failure* (MTTF). This distribution has been used extensively to model survival times. Mukhopadhyay and Datta (1996) developed a purely sequential fixed-width confidence interval procedure for θ and gave its asymptotic second-order properties. More generally, in a one-parameter exponential family, Woodroofe (1987) discussed a three-stage procedure for estimating the mean. In a two-parameter exponential family, Bose and Mukhopadhyay (1995) investigated a sequential interval estimation procedure with proportional closeness via piecewise stopping time.

1.2.2. Sequential Analysis on Multivariate Distribution

This subject can be easily extended to a multivariate situation. Historically, following the spirits of Stein (1945,1949) and Chow and Robbins (1965), Mukhopadhyay (1975) considered a bivariate scenario, while later on, Ghosh et al. (1976) introduced a purely sequential methodology in the most general case. Mukhopadhyay and Al-Mousawi (1986) developed the purely sequential procedure to construct the fixed-size elliptic confidence region for mean vector in a multivariate normal distribution when the covariance matrix $\Sigma = \sigma^2 H$ where $0 < \sigma^2 < \infty$ is assumed unknown but H is a known positive definite matrix. Srivastava (1967) looked into another scenario when Σ was totally unknown. Under the same setup, Ghosh et al. (1997) also developed three-stage procedure and accelerated sequential procedure.

Another popular case is estimating the parameters in a linear model. Healy (1956) and Chatterjee (1962) developed two-stage fixed-size confidence region for the regression coefficients β under normal errors in the spirit of Stein (1945). Later on Albert (1966) and Srivastava (1967) developed the purely sequential fixed-size confidence region for β under non-normal errors in the spirit of Chow and Robbins (1965). Mukhopadhyay (1974) introduced the minimum point estimation problem for β. The papers of Mukhopadhyay (1991,1993) as well as Ghosh et al (1997) may also be helpful for reviewing this vast area.

1.3. Motivation and Layout

In this dissertation, we discuss the possibility to gather a group of k observations at-a-time in purely sequential procedure. Why should one consider sampling k observations in groups at-a-time? Our motivation is simple: We assume that the observations may keep arriving in groups of size k at-a-time. In many practical scenarios, sampling in group can reduce the overall cost. For example, items such as batteries, pens are customarily sold, for example, in packages of 6, 10 or 12 whereas fruit and vegetables are frequently sold in cases with 12 items or more. From Section 1.1.2, recall that each item individually may cost c_1 unit whereas when one purchases k items at-a-time, suppose that each item cost $c_k(>0)$ unit. Practically, c_k will be substantially smaller than c_1. This amounts to customary wisdom: The cost $(=kc_k)$ of purchasing a package made up of k items will be substantially smaller than the cost $(=kc_1)$ of purchasing k items individually. So, from the outset, we may safely assume:

$$c_k \leq c_{k-1} \leq ... \leq c_1 \text{ with arbitrary but fixed } k \geq 1. \tag{1.3.1}$$

Certainly, it will be more reasonable to assume that these inequalities are strict, however, our new methodologies developed subsequently go through just as well with possible "equalities" included in (1.3.1).

In chapter 2, we revisit two fundamental problems discussed in chapter 1.1. We begin by laying down general frameworks for the second-order asymptotic analyses, in both problems, under sequential sampling of one observation at-a-time. Then, instead of gathering one observation at-a-time, we consider sequentially sampling k observations at-a-time in defining our proposed estimation strategies. We replace the customary sample standard deviation as an estimator for σ with a number of other pertinent estimators to come up with new and more appropriate stopping rules to suit the occasion. This part come from the publication Mukhopadhyay and Wang (2020).

In chapter 3, we discuss potential drawbacks of the estimation strategy developed in Mukhopadhyay and Wang(2020), and propose a new class of estimators. These new estimators incorporated in the definition of the stopping boundaries have led to tighter estimation of requisite optimal fixed-sample sizes. We have analyzed the first-order and second-order asymptotic properties under appropriate requirements on the pilot size. Large-scale computer simulations and substantial data analysis have validated such first-order and second-order results. The methodologies are illustrated with the help of time-series data on offshore wind energy.

In chapter 4, two-sample comparison problem will be discussed. We consider two situation (i) when the two sample share common but unknown variance, and (ii) when the variances are unknown and unequal. We set out to replace the multiples of sample standard deviations used in defining requisite boundary crossing conditions with *Gini's mean difference* (GMD), *mean absolute deviation* (MAD), along with a number of combinations of the sample standard deviations, the GMD's, and the MAD's. Our theory and methodologies are amply supported by both large-scale simulations and the illustrations with real data.

Chapter 5 provides the ideas for future work. Chapter 6 summarizes the thesis in a practical way.

Chapter 2

Purely Sequential FWCI and MRPE Problems for the Mean of a Normal Distribution When Observations Gathered in Groups

2.1. Introduction and Layout

In this chapter, we revisit two fundamental problems on sequential estimation in the light of Ghosh and Mukhopadhyay (1976). More specifically, we go back to (i) the *fixed-width confidence interval* (FWCI) estimation problem, and (ii) the *minimum risk point estimation* (MRPE) problem, both in the context of estimating an unknown mean μ in a $N(\mu, \sigma^2)$ population where σ^2 is also assumed unknown. We denote $\boldsymbol{\theta} = (\mu, \sigma^2)$ and consider the corresponding parameter space $\Re \times \Re^+$.

Having recorded $X_1, ..., X_n$ which are *independent and identically distributed* (i.i.d.) as $N(\mu, \sigma^2)$, suppose that we denote:

$$\overline{X}_n = n^{-1} \sum_{i=1}^{n} X_i \text{ and } S_n^2 = (n-1)^{-1} \sum_{i=1}^{n} \left(X_i - \overline{X}_n\right)^2,\ n \geq 2, \tag{2.1.1}$$

standing for the sample mean and the sample variance respectively.

2.1.1. Fixed-Width Confidence Interval (FWCI) Estimation Problem

Having fixed two preassigned numbers, $d(> 0)$ and $0 < \alpha < 1$, the problem is estimating μ

with a confidence interval J such that (i) the length (or width) of J is $2d$, that is fixed in advance, *and* (ii) the confidence coefficient associated with J, that is, $P_{\boldsymbol{\theta}}\{\mu \in J\} \geq 1-\alpha$ for all fixed $\boldsymbol{\theta}$.

As a simple-minded approach, we may pretend to record a fixed number of n observations and let

$$J \equiv J_n = [\overline{X}_n - d, \overline{X}_n + d]. \tag{2.1.2}$$

The associated confidence coefficient is expressed as:

$$P_{\boldsymbol{\theta}}\{\mu \in J_n\} = P_{\boldsymbol{\theta}}\left\{|\overline{X}_n - \mu| \leq d\right\} = 2\Phi\left(\sqrt{n}d/\sigma\right) - 1. \tag{2.1.3}$$

Here, and elsewhere, we denote: $\phi(x) = (2\pi)^{-1/2}\exp(-x^2/2)$ and $\Phi(x) = \int_{y=-\infty}^{x} \phi(y)dy$ for $x \in R$.

Observe that J_n already has the preassigned width $2d$, but we also require that the associated confidence coefficient be approximately at least $1-\alpha$. Thus, we must have:

$$2\Phi\left(\sqrt{n}d/\sigma\right) - 1 \geq 1 - \alpha = 2\Phi\left(a\right) - 1, \text{ say,} \tag{2.1.4}$$

where a is the upper $50\alpha\%$ point of a standard normal distribution.

From (2.1.4), we claim:

$$\begin{gathered}\text{the optimal fixed sample size, had } \sigma^2 \text{ been known, would be} \\ \text{the smallest integer } n \geq a^2\sigma^2 d^{-2} \equiv C_1, \text{ say.}\end{gathered} \tag{2.1.5}$$

However, the magnitude of C_1 remains unknown since σ^2 is unknown. Indeed no fixed-sample-size methodology (that is, fixing n in advance) would be able to come up with a solution for this problem regardless of whether or not the confidence interval is centered (or not centered) at $\overline{X}_n$ or another chosen estimator of μ. Dantzig (1940) proved that fundamental non-existential result. One may also refer to Sen (1981), Woodroofe (1982), Siegmund (1985), Ghosh et al. (1997, Chapter 3), Mukhopadhyay and Solanky (1994), Mukhopadhyay (2000, Chapter 13), Mukhopadhyay et al. (2004), Mukhopadhyay and de Silva (2009, Chapter 2), Zacks (2009,2017) and other sources.

Stein (1945,1949) developed a pathbreaking two-stage sampling strategy to provide an exact solution for this problem. Anscombe (1952,1953), Ray (1957), Chow and Robbins (1965), and Starr (1966a) introduced purely sequential sampling strategies by recording one observation at-a-time to achieve the nearly preassigned confidence coefficient $1-\alpha$, but only *asymptotically* (as $d \to 0$). One may further refer to a number of sources including Ghosh et al. (1997, Chapter 6) and Mukhopadhyay and de Silva (2009, Section 6.2).

Instead of gathering one observation at-a-time, in this paper we propose sequentially sampling k observations at-a-time in defining our proposed methodology. We would replace the customary sample standard deviation with other pertinent estimators of σ by appropriately defining the newly associated stopping rules.

2.1.2. Minimum Risk Point Estimation (MRPE) Problem

Having recorded $X_1, ..., X_n$ which are i.i.d. random variables from $N(\mu, \sigma^2)$, suppose that the loss function in estimating μ by the sample mean, $\overline{X}_n$, is given by:

$$L_n \equiv L_n(\mu, \overline{X}_n) = A\left(\overline{X}_n - \mu\right)^2 + c_1 n, \tag{2.1.6}$$

where both A and c_1 are known positive constants. Here, c_1 represents the cost per unit observation drawn one at-a-time due to sampling and A, a weight function, representing a kind of rate of exchange between the magnitude of the loss due to estimation error and the cost for making such an error.

The goal is to minimize the associated fixed-sample-size risk function which is expressed as:

$$R_n(c_1) \equiv E_{\boldsymbol{\theta}}\left[L_n(\mu, \overline{X}_n)\right] = A\sigma^2 n^{-1} + c_1 n, \tag{2.1.7}$$

for all $0 < \sigma < \infty$. This risk is (nearly) minimized when we determine the requisite sample size as follows:

$$\begin{gathered} n \text{ is the smallest integer} \geq (A/c_1)^{1/2}\sigma = n_1^*, \text{ say, with associated minimum} \\ \text{fixed-sample-size risk given by } R_{n_1^*} \equiv R_{n_1^*}(c_1) = 2c_1 n_1^*. \end{gathered} \tag{2.1.8}$$

However, the magnitude of n_1^* remains unknown since σ^2 is unknown. Indeed no fixed-sample-size methodology (that is, fixing n in advance) would come up with a solution for this problem regardless of whether or not the point estimator involves the sample mean $\overline{X}_n$ or another estimator of choice. Again, Dantzig (1940) proved this fundamental result. One may refer to other sources including Ghosh et al. (1997, Chapter 3), Mukhopadhyay (2000, Chapter 13), and Mukhopadhyay and de Silva (2009, Chapter 2).

Robbins (1959) first developed this groundbreaking formulation and the ensuing purely sequential sampling strategy along with useful metrics to quantify performances of his proposed sequential estimation methodology. This approach was followed by Starr (1966b), Starr and Woodroofe (1969,1972), Ghosh et al. (1976), and Ghosh and Mukhopadhyay (1976) among others. They introduced purely sequential sampling strategies by recording one observation at-a-time to achieve the nearly minimum risk, but only asymptotically (as $c_1 \to 0$).

Instead of gathering one observation at-a-time, in this paper we propose sequentially sampling k observations at-a-time in defining our proposed methodology. We would replace the customary sample standard deviation with other pertinent estimators of σ by appropriately defining the newly stopping rules. In order to look into the status of research in either problem, the cited books and the monographs of Sen (1981), Woodroofe (1982), Siegmund (1985), Ghosh and Sen (1991), Mukhopadhyay and Solanky (1994), Mukhopadhyay et al. (2004), and Zacks (2009,2017) may guide one to gain a broader view.

2.1.3. A Summary from Nonlinear Renewal Theory

For completeness, we summarize an outline of the nonlinear renewal theory in sequential analysis in the spirits of Woodroofe (1977), Lai and Siegmund (1977,1979), and Mukhopadhyay (1988). In order to work through one cohesive set of notation, we recall the structure defined in Mukhopadhyay and Solanky (1994, pp. 48-50).

Suppose that $W_1, W_2, ...$ are i.i.d. positive (w.p.1) and continuous random variables having *all* positive moments finite and define a generic stopping time:

$$R_{h^*} \equiv R = \inf \left\{ n \geq n_0; \;\; \Sigma_{i=1}^{n} W_i \leq h^* n^{\delta} l(n) \right\}, \tag{2.1.9}$$

where $\delta > 1$, $h^* > 0$, $n_0 \geq 1$ and $l(n) = 1 + l_0 n^{-1} + o(n^{-1})$ as $n \to \infty$ with $-\infty < l_0 < \infty$.

Assume that the distribution function of W_1 satisfies the condition:

$$P\{W_1 \leq u\} \leq B^* u^b, \tag{2.1.10}$$

for all $u > 0$, with some $b > 0$, $B^* > 0$. We may denote:

$$\lambda = E\left[W_1\right], \tau^2 = V[W_1], \beta^* = (\delta - 1)^{-1}, n^* = (\lambda h^{*-1})^{\beta^*}, p = \beta^{*2}\tau^2\lambda^{-2}, \tag{2.1.11}$$

and further let:

$$\begin{aligned} \nu &= \tfrac{\beta^*}{2\lambda}\left\{(\delta-1)^2\lambda^2 + \tau^2\right\} - \Sigma_{n=1}^{\infty} n^{-1} E\left\{\max\left(0, \Sigma_{i=1}^{n} W_i - n\delta\lambda\right)\right\}; \\ \eta &= \beta^*\lambda^{-1}\nu - \beta^* l_0 - \tfrac{1}{2}\delta\beta^{*2}\tau^2\lambda^{-2}; \text{ and } H = n^{*-1/2}(R_{h^*} - n^*). \end{aligned} \tag{2.1.12}$$

Next, an important set of results is summarized from Woodroofe (1977) and Lai and Siegmund (1977,1979) for immediate applications. One may also refer to Mukhopadhyay and Solanky (1994, Theorem 2.4.8). We may point out that the results quoted in (2.1.13) continue to hold under the finiteness of appropriate positive moments of W. We avoid full generality for the sake of simplicity, especially since all positive moments of W are indeed finite in our applications.

With R_{h^*} defined in (2.1.9), we have:

$$
\begin{array}{ll}
\text{(i)} & P(R_{h^*} \le \varepsilon n^*) = O(h^{*n_0 b}) \text{ if } 0 < \varepsilon < 1,\ n_0 \ge 1; \\
\text{(ii)} & H \stackrel{\pounds}{\to} N(0,p) \text{ if } n_0 \ge 1; \\
\text{(iii)} & |H|^\kappa \text{ is uniformly integrable if } n_0 b > \frac{1}{2}\beta^*\kappa, \text{ where } \kappa \text{ is an arbitrary} \\
& \text{but fixed positive number; and} \\
\text{(iv)} & E\left[R_{h^*} - n^*\right] = \eta + o(1) \text{ if } n_0 b > \beta^* \ [\textit{asymptotic second-order efficiency}];
\end{array}
\qquad (2.1.13)
$$

where b comes from (2.1.10), n^*, β^*, p come from (2.1.11), and H, η come from (2.1.12).

For proofs, one may refer to Woodroofe (1977) and Lai and Siegmund (1977,1979). One could additionally refer to Woodroofe (1982), Siegmund (1985), Mukhopadhyay and Solanky (1994, Theorem 2.4.8), Ghosh et al. (1997), and Mukhopadhyay and de Silva (2009).

Property (ii) customarily follows from Ghosh and Mukhopadhyay's (1975) theorem on asymptotic normality of standardized stopping variables. One may consider looking into a recent communication by Mukhopadhyay and Zhang (2018) in this regard. Property (iv) is referred to as what is known to be the *asymptotic second-order efficiency* in the sense of Ghosh and Mukhopadhyay (1981) which originally transpired from Mukhopadhyay (1980).

2.1.4. A Layout of This Chapter

In the contexts of both FWCI and MRPE problems, Section 2.2 begins where Ghosh and Mukhopadhyay's (1976) paper ended with substantial synthesis by significantly updating it via sequentially drawing one observation at-a-time. First, Section 2.2.1.2 develops a second-order expansion of the coverage probability (Theorem 2.2.1) whereas Section 2.2.2.1 develops a second-order expansion of the regret function (Theorem 2.2.2), both associated with general multi-stage estimation strategies under appropriate sets of broad-ranging assumptions.

Next, we move ahead to emphasize purely sequential methodologies when sampling is carried out by recording $k(\geq 2)$ observations at-a-time in a group. Hence, we consider replacing the customarily used sample standard deviation or its suitable multiple seen in the boundary crossing condition (for example, refer to (2.2.1) and (2.2.27)), with more appropriate and new estimators of σ. These new alternate estimators are introduced in (2.3.2), Section 2.3 along with their preliminary properties.

Section 2.4 builds upon purely sequential FWCI methodologies from Section 2.2.1 when sampling is carried out by recording $k(\geq 2)$ observations at-a-time. Main asymptotic second-order results are summarized by Theorems 2.4.1-2.4.2. Section 2.5 extends purely sequential MRPE methodologies from Section 2.2.2 when sampling is carried out by recording $k(\geq 2)$ observations at-a-time. Main asymptotic first-order and second-order results are summarized by Theorems 2.5.1-2.5.3.

These are followed by extensive sets of carefully laid out data analyses (Tables 2.2-2.6 and Tables 2.8-2.12) assisted by large-scale computer simulations in Section 2.2.6. These provide important data-validation of crucial theoretical conclusions and hence give us much confidence in proposing our methodologies developed here with confidence for smooth implementation in practice when purely sequential sampling strategies are contemplated. These are wrapped up with Section 2.7 where we include illustrations of our proposed FWCI and MRPE methodologies using breast cancer data arising from Dua and Graff (2019).

2.2. Updating the Ghosh-Mukhopadhyay (1976) Paper Under Customary Purely Sequential Sampling Strategies

In the spirit of Ghosh and Mukhopadhyay (1976), we begin by reviewing the purely sequential fixed-width confidence interval estimation problem from Section 2.1.1 for μ now taken up in Section 2.2.1. Section 2.2.1.1 develops a second-order expansion of the confidence coefficient associated with the fixed-width confidence interval problem under considerable generality.

Then, we move to the purely sequential minimum risk point estimation problem from Section 2.1.2 for μ now taken up in Section 2.2.2. Section 2.2.2.1 develops a second-order expansion of the associated regret function under considerable generality.

Initially, however, the methodologies under present scrutiny incorporates purely sequential estimation strategies by recording one additional observation at-a-time as needed by incorporating an associated stopping rule. We highlight both first-order and second-order asymptotic characteristics.

2.2.1. FWCI Estimation Problem

Recall the expression of the optimal fixed sample size $C_1(= a^2\sigma^2/d^2)$, had σ^2 been known, from (2.1.5). We begin with the pilot observations $X_1, ..., X_{m_0}$, $m_0 \geq 2$, and then proceed by recording one additional X at-a-time according to the stopping rules of Anscombe (1952), Ray (1957) and Chow and Robbins (1965) defined as follows:

$$\text{Methodology } \mathcal{P}_{1,1}\text{:} \quad N_{\mathcal{P}_{1,1},d} \equiv N_{\mathcal{P}_{1,1}} = \inf \{ n \geq m_0(\geq 2) : \quad n \geq a^2 S_n^2/d^2 \}. \qquad (2.2.1)$$

This sampling strategy is implemented as follows: We first obtain $S_{m_0}^2$ based on the pilot data and check whether $m_0 \geq a^2 S_{m_0}^2/d^2$. If $m_0 \geq a^2 S_{m_0}^2/d^2$, then sampling terminates right here with the final sample size, $N_{\mathcal{P}_{1,1}} = m_0$. But, if $m_0 < a^2 S_{m_0}^2/d^2$, then we record one more observation X and update the sample variance by obtaining $S^2_{m_0+1}$. Next, we check whether $m_0+1 \geq a^2 S^2_{m_0+1}/d^2$. If $m_0+1 \geq a^2 S^2_{m_0+1}/d^2$, then sampling terminates here with the final sample size, $N_{\mathcal{P}_{1,1}} = m_0+1$. Otherwise, sampling continues until for the first time we arrive at a sample size n, purely sequentially drawing one additional observation at-a-time, such that $n \geq a^2 S_n^2/d^2$ happens. Then, we terminate the sampling strategy.

From Chow and Robbins (1965), we can claim that $P_{\boldsymbol{\theta}}\{N_{\mathcal{P}_{1,1,d}} < \infty\} = 1$. Upon termination of sampling, we will propose the fixed-width confidence interval:

$$J \equiv J_{N_{\mathcal{P}_{1,1,d}}} = [\overline{X}_{N_{\mathcal{P}_{1,1,d}}} - d, \overline{X}_{N_{\mathcal{P}_{1,1,d}}} + d], \tag{2.2.2}$$

based on the finally accrued data $\{N_{\mathcal{P}_{1,1,d}}, X_1, ..., X_{N_{\mathcal{P}_{1,1,d}}}\}$ giving rise to the randomly stopped sample mean, $\overline{X}_{N_{\mathcal{P}_{1,1,d}}}$.

In what follows, a number of interesting *asymptotic first-order* properties are summarized as $d \to 0$:

$$\begin{gathered} N_{\mathcal{P}_{1,1,d}}/C_1 \overset{P_{\boldsymbol{\theta}}}{\to} 1;\ E_{\boldsymbol{\theta}}[N_{\mathcal{P}_{1,1,d}}/C_1] \to 1 \text{ [Asymptotic first-order efficiency];} \\ \text{and } P_{\boldsymbol{\theta}}\{\mu \in J_{N_{\mathcal{P}_{1,1,d}}}\} \to 1-\alpha \text{ [Asymptotic consistency];} \end{gathered} \tag{2.2.3}$$

for every fixed μ, σ, m_0 and α. One may find proofs from combining sources including Chow and Robbins (1965), Ghosh and Mukhopadhyay (1976), Ghosh et al. (1997, Section 8.2), and Mukhopadhyay and de Silva (2009, Section 6.2).

We may rewrite $N_{\mathcal{P}_{1,1,d}} = R_d + 1$ w.p.1 where we denote:

$$R_d \equiv R = \inf\left\{n \geq m_0 - 1;\ \ \Sigma_{i=1}^n W_i \leq C_1^{-1} n^2 (1+n^{-1})\right\}, \tag{2.2.4}$$

with W_i's distributed as i.i.d. χ_1^2 random variables. Clearly, (2.2.4) agrees with (2.1.9), and thus (2.1.10)-(2.1.11) would lead to:

$$\lambda = 1, \tau^2 = 2, h^* = C_1^{-1}, n^* = C_1, \delta = 2, \beta^* = 1, l_0 = 1,$$
$$p = 2,\ n_0 = m_0 - 1 \text{ and } b = \tfrac{1}{2}. \tag{2.2.5}$$

Applying (2.1.12), we immediately obtain:

$$\nu = \tfrac{3}{2} - D \text{ with } D = \Sigma_{n=1}^{\infty} n^{-1} E\left\{\max(0, \chi_n^2 - 2n)\right\} \text{ which implies:}$$
$$\eta = \tfrac{3}{2} - D - 3 = -\tfrac{3}{2} - D. \tag{2.2.6}$$

Now, (2.1.13) in combination with (2.2.5)-(2.2.6) lead to the following significantly sharper results as $d \to 0$:

(i) $P_{\boldsymbol{\theta}}\{N_{\mathcal{P}_{1,1},d} \le \varepsilon C_1\} = O(C_1^{-(m_0-1)/2})$ for fixed $0 < \varepsilon < 1$, if $m_0 \ge 2$;

(ii) $H \equiv C_1^{-1/2}(N_{\mathcal{P}_{1,1},d} - C_1) \xrightarrow{\pounds} N(0, 2)$, if $m_0 \ge 2$;

(iii) $|H|^{\kappa}$ is uniformly integrable if $m_0 > 1 + \kappa$ where κ is an arbitrary but fixed positive number; and (2.2.7)

(iv) $E_{\boldsymbol{\theta}}\left[N_{\mathcal{P}_{1,1},d} - C_1\right] = -\tfrac{1}{2} - D + o(1)$ if $m_0 \ge 4$;

for every fixed μ, σ, m_0 and α. Part (i) in (2.2.7) is obvious since

$$P_{\boldsymbol{\theta}}\{N_{\mathcal{P}_{1,1},d} \le \varepsilon C_1\} = P_{\boldsymbol{\theta}}\{R + 1 \le \varepsilon C_1\} \le P_{\boldsymbol{\theta}}\{R \le \varepsilon C_1\}.$$

We also recall that $N_{\mathcal{P}_{1,1}} = R_d + 1$ w.p.1. Thus, part (iv) in (2.2.7) follows from part (iv) in

(2.1.13), that is,

$$E_{\boldsymbol{\theta}}\left[N_{\mathcal{P}_{1,1},d} - C_1\right] = \eta + 1 + o(1) \text{ if } m_0 \geq 4.$$

2.2.1.1. Estimation of D and η Defined in (2.2.6)

We obtained a fairly accurate estimated value of D via large-scale simulations carried out with the help of Matlab R2019b. Let us define $Y_n = \max(0, \chi_n^2 - 2n)$, and then we have $D = \Sigma_{n=1}^{\infty} n^{-1} E[Y_n]$. To estimate $E[Y_n]$, having fixed n, we had generated 10000 random values of $\max(0, \chi_n^2 - 2n)$ and calculated the associated sample mean $\overline{Y}_n$. We repeated this approach for every fixed $n = 1, 2, ..., 10000$ and estimated D by $\Sigma_{n=1}^{\infty} n^{-1} \overline{Y}_n$. We found:

$$\begin{aligned} D &\approx 0.6839 \text{ with its estimated standard error (s.e.) } 0.00004 \\ \Rightarrow \eta + 1 &= 1 - \tfrac{3}{2} - D \approx -1.1839 \text{ with its estimated s.e. } 0.00004. \end{aligned} \tag{2.2.8}$$

In other words, we should expect:

$$E_{\boldsymbol{\theta}}\left[N_{\mathcal{P}_{1,1},d} - C_1\right] \approx -1.1839. \tag{2.2.9}$$

In all fairness, we should mention that a magnitude of this discrepancy between $E_{\boldsymbol{\theta}}[N_{\mathcal{P}_{1,1},d}]$ and C_1 was reported to be -1.1825 by Mukhopadhyay and de Silva (2009, p. 119) on line 2 from the top. That is quite close to what we are reporting in (2.2.9), both well within the estimated standard errors.

Remark 2.2.1. Since, we have

$$H \equiv C_1^{-1/2}(N_{\mathcal{P}_{1,1},d} - C_1) = C_1^{-1/2}(R_d - C_1) + C_1^{-1/2}, \tag{2.2.10}$$

Slutsky's theorem will resolve part (ii) in (2.2.7) from part (ii) in (2.1.13).

Remark 2.2.2. From (2.2.9), we observe that for sufficiently large C_1, we can claim (w.p.1):

$$\begin{aligned}&\left|C_1^{-1/2}(N_{\mathcal{P}_{1,1,d}} - C_1)\right| \leq \left|C_1^{-1/2}(R_d - C_1)\right| + 1\\ \Rightarrow &\left|C_1^{-1/2}(N_{\mathcal{P}_{1,1,d}} - C_1)\right|^2 \leq \left|C_1^{-1/2}(R_d - C_1)\right|^2 + 2\left|C_1^{-1/2}(R_d - C_1)\right| + 1.\end{aligned} \tag{2.2.11}$$

In view of (2.2.11), we will clearly require uniform integrability of $\left|C_1^{-1/2}(R_d - C_1)\right|^2$ on the right-hand side which will show the uniform integrability of $\left|C_1^{-1/2}(R_d - C_1)\right|$. Thus, part (iii) in (2.2.7) will follow from (2.2.11) in view of part (iii) in (2.1.13).

Indeed parts (iii)-(iv) in (2.2.7) show asymptotic second-order results. Part (iv) shows that one observation at-a-time purely sequential estimation methodology (2.2.1)-(2.2.2) is asymptotically second-order efficient in the sense of Ghosh and Mukhopadhyay (1981).

Woodroofe (1977) gave an elegant asymptotic second-order expansion of the associated confidence coefficient $P_{\boldsymbol{\theta}}\{\mu \in J_{N_{\mathcal{P}_1,d}}\}$, namely, the following:

$$P_{\boldsymbol{\theta}}\{\mu \in J_{N_{\mathcal{P}_{1,1},d}}\} = (1-\alpha) - \tfrac{1}{2}aC_1^{-1}(2 + 2D + a^2)\phi(a) + o(d^2) \text{ if } m_0 \geq 7, \tag{2.2.12}$$

for every fixed μ, σ, and α. Hence, we refrain from rehashing its derivation.

Instead, we move on to the next Section 2.2.1.2 where we provide a justification of a result analogous to that in (2.12) under substantial generality. Such a general result will further help us in developing the subsequent material.

2.2.1.2. A General Asymptotic Second-Order Expansion of the Confidence Coefficient

We begin with a general, but arbitrary, multi-stage sampling strategy leading to a suitably defined arbitrary stopping time, $M \equiv M_d$. We are, however, assured that M_d is "close" to C_1. Let us now list a number of *assumptions* that may be needed in order to draw certain specific conclusion(s):

(A1): For every fixed $n \geq m_0(> 1)$, the pilot size, the event $[M_d = n]$ depends only on the statistic $\mathbf{Z}_n \equiv \{X_1 - X_n, ..., X_{n-1} - X_n\}$;

(A2): $M_d/C_1 \overset{P_{\boldsymbol{\theta}}}{\to} 1$ as $d \to 0$, and also $P_{\boldsymbol{\theta}}\{M_d \leq \varepsilon C_1\} = O(C_1^{-a_1 m_0 + a_2})$ if $0 < \varepsilon < 1$, as $d \to 0$, with some $a_1 > 0$, if $m_0 > a_2/a_1$;

(A3): $H \equiv C_1^{-1/2}(M_d - C_1) \overset{\pounds}{\to} N(0, a_3)$ as $d \to 0$, with some $a_3 > 0$;

(A4): $|H|^2$ is uniformly integrable if $m_0 > a_4(> 0)$; and

(A5): $E_{\boldsymbol{\theta}}[M_d - C_1] = a_5 + o(1)$ as $d \to 0$, if $m_0 > a_6(> 0)$;

(2.2.13)

for every fixed μ, σ, and α.

Such a general multi-stage sampling strategy does not necessarily have to look like those highlighted in either (2.1.9) or (2.2.1). One may contemplate utilizing other suitable strategies including appropriate two-stage, three-stage, or accelerated sequential sampling which will terminate with the finally accrued data $\{M_d, X_1, ..., X_{M_d}\}$ and the associated fixed-width confidence interval:

$$J_{M_d} = [\overline{X}_{M_d} - d, \overline{X}_{M_d} + d] \text{ for } \mu. \tag{2.2.14}$$

Theorem 2.2.1. *Under a general asymptotically second-order efficient multi-stage fixed-width confidence interval methodology (M_d, J_{M_d}) from (2.2.14), under the standing assumptions (A1)-(A5) from (2.2.13), for every fixed μ, σ, and α, we have the following asymptotic* *second-order expansion of the confidence coefficient associated with* J_{M_d} *as* $d \to 0$:

$$P_{\boldsymbol{\theta}}\{\mu \in J_{M_d}\} = (1-\alpha) + a\left\{a_5 - \tfrac{1}{4}a_3(1+a^2)\right\}\phi(a)C_1^{-1} + o(d^{-2}), \text{ if the pilot size, } m_0 > \max\{1, a_6, a_4, \tfrac{1}{2a_1}(5+2a_2)\}.$$

Proof: Under the **assumption (A1)**, we invoke Basu's (1955) theorem in the spirit of Mukhopadhyay (2000, Example 6.6.15) to claim that the random variable $I(M_d = n)$ and the sample mean, $\overline{X}_n$, are independent for every fixed $n \geq m_0$. Thus, we can express:

$$\begin{aligned} P_{\boldsymbol{\theta}}\{\mu \in J_{M_d}\} &= E_{\boldsymbol{\theta}}\left[2\Phi\left(M_d^{1/2}d/\sigma\right) - 1\right] = E_{\boldsymbol{\theta}}\left[2\Phi\left(aM_d^{1/2}C_1^{-1/2}\right) - 1\right] \\ &= E_{\boldsymbol{\theta}}\left[\psi\left(M_d/C_1\right)\right] \text{ where } \psi(x) \equiv 2\Phi\left(ax^{1/2}\right) - 1,\ x > 0. \end{aligned} \tag{2.2.15}$$

Observe that $\frac{d}{dx}\Phi(x) = \phi(x)$ and $\frac{d^2}{dx^2}\Phi(x) = \frac{d}{dx}\phi(x) = -x\phi(x)$. Thus, we may write:

$$\begin{aligned} \psi^{'}(x) \equiv &\quad \tfrac{d}{dx}\psi(x) = ax^{-1/2}\phi\left(ax^{1/2}\right); \\ \psi^{''}(x) \equiv &\quad \tfrac{d^2}{dx^2}\psi(x) = -\tfrac{1}{2}a(x^{-3/2} + a^2x^{-1/2})\phi\left(ax^{1/2}\right). \end{aligned} \tag{2.2.16}$$

Then, applying Taylor's expansion of $\psi(x)$ around $x = 1$, we obtain:

$$P_{\boldsymbol{\theta}}\{\mu \in J_{M_d}\} = \psi(1) + \psi^{'}(1)C_1^{-1}E_{\boldsymbol{\theta}}\left[M_d - C_1\right] + \tfrac{1}{2}C_1^{-1}E_{\boldsymbol{\theta}}\left[H^2\psi^{''}(\xi_d)\right], \tag{2.2.17}$$

with H coming from (2.2.13), and ξ_d being a random variable lying between 1 and M_d/C_1.

The first part of **assumption (A2)** gives: $\xi_d \stackrel{P_{\boldsymbol{\theta}}}{\to} 1$ as $d \to 0$. Next, from **assumption**

(A5), we immediately rewrite the middle term from (2.2.17) as:

$$\psi'(1)C_1^{-1}E_{\boldsymbol{\theta}}\left[M_d - C_1\right] = a_5\psi'(1)C_1^{-1} + o(C_1^{-1}) \text{ if } m_0 > a_6. \tag{2.2.18}$$

Now, we need to handle the last term from (2.2.17) which, in view of (2.2.16), becomes:

$$-\tfrac{1}{4}aC_1^{-1}E_{\boldsymbol{\theta}}\left[H^2\left\{\xi_d^{-3/2} + a^2\xi_d^{-1/2}\right\}\phi\left(a\xi_d^{1/2}\right)\right]. \tag{2.2.19}$$

We use "u" as a generic positive constant not involving d. Now, let us break up the whole space as the union of two disjoint events or sets as follows:

$$[M_d/C_1 > \tfrac{1}{2}] \text{ and } [M_d/C_1 \leq \tfrac{1}{2}]. \tag{2.2.20}$$

We argue that $\xi_d > \frac{1}{2}$ on the set $[M_d/C_1 > \frac{1}{2}]$ so that we can express (w.p.1):

$$\left|H^2\left\{\xi_d^{-3/2} + a^2\xi_d^{-1/2}\right\}\phi\left(a\xi_d^{1/2}\right)I(M_d/C_1 > \tfrac{1}{2})\right| \leq uH^2. \tag{2.2.21}$$

Next, in view of (2.2.21), **assumption (A4)** shows that the term

$$H^2\left\{\xi_d^{-3/2} + a^2\xi_d^{-1/2}\right\}\phi\left(a\xi_d^{1/2}\right)I(M_d/C_1 > \tfrac{1}{2})$$

is uniformly integrable if $m_0 > a_4$.

Now, **assumption (A3)** leads to the following conclusion (as $d \to 0$):

$$H^2 \left\{\xi_d^{-3/2} + a^2 \xi_d^{-1/2}\right\} \phi\left(a\xi_d^{1/2}\right) I(M_d/C_1 > \tfrac{1}{2}) \xrightarrow{\pounds} (1+a^2)\phi(a)\, a_3 \chi_1^2. \tag{2.2.22}$$

Then, since the left-hand side of (2.2.22) is uniformly integrable in view of (A4), we obtain:

$$\begin{aligned} &E_{\boldsymbol{\theta}}\left[H^2 \left\{\xi_d^{-3/2} + a^2 \xi_d^{-1/2}\right\} \phi\left(a\xi_d^{1/2}\right) I(M_d/C_1 > \tfrac{1}{2})\right] \\ &\quad = (1+a^2)\phi(a)\, a_3 + o(1) \text{ if } m_0 > a_4. \end{aligned} \tag{2.2.23}$$

Next, we argue that $\xi_d > M_d/C_1 \geq m_0/C_1$ on the set $[M_d/C_1 \leq \frac{1}{2}]$ so that we can express:

$$\begin{aligned} &E_{\boldsymbol{\theta}}\left[H^2 \left\{\xi_d^{-3/2} + a^2 \xi_d^{-1/2}\right\} \phi\left(a\xi_d^{1/2}\right) I(M_d/C_1 \leq \tfrac{1}{2})\right] \\ &\leq u E_{\boldsymbol{\theta}}\left[\left(\tfrac{M_d^2 + C_1^2}{C_1}\right)\left\{C_1^{3/2} + a^2 C_1^{1/2}\right\} I(M_d/C_1 \leq \tfrac{1}{2})\right] \\ &\leq O(C_1) O(C_1^{3/2}) P_{\boldsymbol{\theta}}\{M_d/C_1 \leq \tfrac{1}{2}\} \\ &= O(C_1^{\frac{5}{2}}) O(C_1^{-a_1 m_0 + a_2}), \text{ by assumption (A2)}, \\ &= o(1) \text{ if } m_0 > \tfrac{1}{2}(5 + 2a_2)/a_1. \end{aligned} \tag{2.2.24}$$

At this point, a combination of (2.2.23)-(2.2.24) leads to the result:

$$\begin{aligned} &E_{\boldsymbol{\theta}}\left[H^2 \left\{\xi_d^{-3/2} + a^2 \xi_d^{-1/2}\right\} \phi\left(a\xi_d^{1/2}\right)\right] = (1+a^2)\phi(a)\, a_3 + o(1), \\ &\qquad \text{if } m_0 > \max\{a_4, \tfrac{1}{2a_1}(5 + 2a_2)\}. \end{aligned} \tag{2.2.25}$$

Observe that $\psi(1) = 1 - \alpha$. Next, we go back to (2.2.17) and combine it with (2.2.18)-

(2.2.19) and (2.2.25) to claim:

$$P_{\boldsymbol{\theta}}\{\mu \in J_{M_d}\} = (1-\alpha) + aC_1^{-1}\left\{a_5 - \tfrac{1}{4}a_3(1+a^2)\right\}\phi(a) + o(d^2), \quad \text{if } m_0 > \max\{a_6, a_4, \tfrac{1}{2a_1}(5+2a_2)\}. \tag{2.2.26}$$

The proof is now complete. ■

Remark 2.2.3. In the context of one observation at-a-time purely sequential estimation methodology $(N_{\mathcal{P}_{1,1},d}, J_{N_{\mathcal{P}_{1,1},d}})$, one needs to examine whether (2.2.12) follows from (2.2.26). First, we identify n_0 with m_0. All **assumptions (A1)-(A5)** from (2.2.13) are seen to hold with $a_1 = a_2 = \frac{1}{2}$, $a_3 = 2$, $a_4 = 3$, $a_5 = -\frac{1}{2} - D$, and $a_6 = 3$. Thus, a sufficient condition on the pilot size for the second-order expansion (2.2.26) now turns out to be $n_0 > 6$, that is $n_0 \geq 7$ and hence (2.2.12) and (2.2.26) agrees on all counts.

2.2.2. MRPE Problem

Recall the expression of the optimal fixed sample size $n_1^*(= \sqrt{A/c_1}\sigma)$ had σ^2 been known from (2.1.8). Here, c_1 is the cost for each observation. We begin with the pilot observations $X_1, ..., X_{m_0}$, $m_0 \geq 2$, and then proceed by recording one additional X at-a-time according to the stopping rules of Anscombe (1952), Ray (1957) and Chow and Robbins (1965) defined as:

$$\text{Methodology } Q_{1,1}\text{:} \quad N_{Q_{1,1},c_1} \equiv N_{Q_{1,1}} = \inf\left\{n \geq m_0(\geq 2): \quad n \geq (A/c_1)^{1/2} G_n\right\}, \tag{2.2.27}$$

where we denote:

$$G_n = a_n S_n \text{ with } a_n = \left\{\tfrac{1}{2}(n-1)\right\}^{1/2} \Gamma(\tfrac{1}{2}(n-1))\{\Gamma(\tfrac{n}{2})\}^{-1}. \tag{2.2.28}$$

This makes G_n *unbiased* for estimating σ.

This sampling strategy is implemented in a same spirit as in Section 2.2.1: We first obtain G_{m_0} based on the pilot data and check whether $m_0 \geq (A/c_1)^{1/2} G_{m_0}$. If $m_0 \geq (A/c_1)^{1/2} G_{m_0}$, then sampling terminates right here and the final sample size $N_{Q_{1,1}} = m_0$. But, if $m_0 < (A/c_1)^{1/2} G_{m_0}$, then we record one more observation X and update the unbiased estimator (of σ) by G_{m_0+1}. Next, we check whether $m_0+1 \geq (A/c_1)^{1/2} G_{m_0+1}$. If $m_0+1 \geq (A/c_1)^{1/2} G_{m_0+1}$, then sampling terminates right here and the final sample size $N_{Q_{1,1}} = m_0 + 1$. Otherwise, sampling continues until for the first time we arrive at a sample size n, purely sequentially drawing one additional observation at-a-time, such that $n \geq (A/c_1)^{1/2} G_n$ happens. Then, we terminate the sampling strategy.

From Chow and Robbins (1965), we can claim that $P_{\boldsymbol{\theta}}\{N_{Q_{1,1,c_1}} < \infty\} = 1$. Thus, upon termination of sampling, we will propose the point estimator $\overline{X}_{Q_{1,1,c_1}}$ for μ based on the finally accrued data, namely, $\{N_{Q_{1,1,c_1}}, X_1, ..., X_{N_{Q_{1,1,c_1}}}\}$. The associated **sequential risk** is given by:

$$R_{N_{Q_{1,1}}}(c_1) = E_{\boldsymbol{\theta}}\left[L_{N_{Q_{1,1,c_1}}}(\mu, \overline{X}_{N_{Q_{1,1,c_1}}})\right] = A\sigma^2 E_{\boldsymbol{\theta}}[N^{-1}_{Q_{1,1,c_1}}] + c_1 E_{\boldsymbol{\theta}}[N_{Q_{1,1,c_1}}]. \tag{2.2.29}$$

Then, the other major requisite entities are given by:

$$\begin{aligned} \text{Risk efficiency:} \quad & \xi_{Q_{1,1}}(c_1) \equiv \frac{R_{N_{Q_{1,1}}}(c_1)}{R_{n_1^*}} = \tfrac{1}{2}E_{\boldsymbol{\theta}}[N_{Q_{1,1},c_1} n_1^{*-1}] + \tfrac{1}{2}E_{\boldsymbol{\theta}}[n_1^* N_{Q_{1,1},c_1}^{-1}], \\ \text{Regret:} \quad & \omega_{Q_{1,1}}(c_1) \equiv R_{N_{Q_{1,1},c_1}}(c_1) - R_{n_1^*} = c_1 E_{\boldsymbol{\theta}}\left[\frac{(N_{Q_{1,1},c_1} - n_1^*)^2}{N_{Q_{1,1},c_1}}\right], \end{aligned} \tag{2.2.30}$$

as defined by Robbins (1959).

Some interesting asymptotic first-order properties may be summarized as follows (as $c_1 \to 0$):

$$\begin{gathered} N_{Q_{1,1},c_1}/n_1^* \stackrel{P_{\boldsymbol{\theta}}}{\to} 1;\ E_{\boldsymbol{\theta}}[N_{Q_{1,1},c_1}/n_1^*] \to 1 \text{ [Asymptotic first-order efficiency]}; \\ \xi_{Q_{1,1}}(c_1) \to 1 \text{ if } m_0 \geq 3 \text{ [Asymptotic risk efficiency]}; \end{gathered} \tag{2.2.31}$$

in the spirit of Starr (1966b), but they fixed $a_n \equiv 1$ for all n. Starr and Woodroofe (1969) and Woodroofe (1977) respectively gave the following asymptotic second-order results (as $c_1 \to 0$):

$$\text{Bounded regret:} \quad \omega_{Q_{1,1}}(c_1) = O(c_1) \text{ if } m_0 \geq 3;$$

$$\text{Expansion of regret:} \quad \omega_{Q_{1,1}}(c_1) = \tfrac{1}{2}c_1 + o(c_1) \text{ if } m_0 \geq 4.$$

Recall a_n from (2.2.28). Next, in order to apply nonlinear renewal theory, we may rewrite the inequality inside the stopping rule defined in (2.2.27) as follows:

$$\begin{gathered} n \geq \sqrt{\tfrac{A}{c_1}} a_n S_n \Leftrightarrow n^2 \geq \tfrac{A}{c_1} a_n^2 S_n^2 \Leftrightarrow \tfrac{n^2(n-1)}{\sigma^2} \geq \tfrac{A}{c_1} a_n^2 \textstyle\sum_{i=1}^{n-1} W_i \\ \Leftrightarrow \textstyle\sum_{i=1}^{n-1} W_i \leq n^2(n-1) a_n^{-2} n_1^{*-2}. \end{gathered} \tag{2.2.32}$$

We let $N_{Q_{1,1,c_1}} = R_{c_1} + 1$ w.p.1 where

$$R_{c_1} \equiv R = \inf\left\{ n \geq m_0 - 1 : \quad \sum_{i=1}^{n} W_i \leq n(n+1)^2 n_1^{*-2} \left[\sqrt{n/2} \frac{\Gamma(n/2)}{\Gamma((n+1)/2)} \right]^{-2} \right\} \quad (2.2.33)$$

Using the expansion of the ratio of gamma functions from Abramowitz and Stegun (1972, 6.1.47, p. 257), namely,

$$z^{b-a}\Gamma(z+a)\left\{\Gamma(z+b)\right\}^{-1} = 1 + \tfrac{1}{2}(a-b)(a+b-1)z^{-1} + o(z^{-1}) \text{ as } z \to \infty,$$

with $z = \frac{1}{2}n, a = 0$ and $b = \frac{1}{2}$, it is easy to show:

$$\sqrt{\tfrac{n}{2}}\Gamma(\tfrac{n}{2})\left\{\Gamma(\tfrac{n+1}{2})\right\}^{-1} = 1 + \tfrac{1}{4}n^{-1} + o(n^{-1}).$$

Thus, we can equivalently express $R_{c_1} \equiv R$ from (2.2.33) as follows:

$$\begin{aligned} &\inf\left\{ n \geq m_0 - 1 : \quad \sum_{i=1}^{n} W_i \leq n_1^{*-2} n^3 (1+n^{-1})^2 \left(1 + \tfrac{1}{4n} + o(n^{-2})\right)^{-2} \right\} \\ &= \inf\left\{ n \geq m_0 - 1 : \quad \sum_{i=1}^{n} W_i \leq n_1^{*-2} n^3 \left(1 + \tfrac{3}{2}n^{-1} + o(n^{-1})\right) \right\}, \end{aligned} \quad (2.2.34)$$

where the W_i's are i.i.d. χ_1^2 random variables.

Now, the representation from (2.2.34) agrees with (2.1.9) where we have:

$$\lambda = 1, \tau^2 = 2, h^* = n_1^{*-2}, n^* = n_1^*, \delta = 3, \beta^* = \tfrac{1}{2}$$

$$l_0 = \tfrac{3}{2}, p = \tfrac{1}{2}, n_0 = m_0 - 1, \text{and } b = \tfrac{1}{2}.$$

Applying (2.1.12), we immediately obtain:

$$\begin{aligned} \nu = \tfrac{3}{2} - D \text{ where } D = \textstyle\sum_{n=1}^{\infty} n^{-1} E\{\max(0, \chi_n^2 - 3n)\} \\ \Rightarrow \eta = -\tfrac{3}{4} - \tfrac{1}{2} D. \end{aligned} \tag{2.2.35}$$

An estimated value of D via large-scale simulations gave

$$D \approx 0.2324 \text{ with its estimated standard error (s.e.) } 0.00035$$

$$\Rightarrow \ \nu \approx 1.2661 \text{ and } \eta \approx -0.8662.$$

Recall that $N_{Q_{1,1}} = R+1$ w.p.1. Then, combining (2.1.11) - (2.1.13) with (2.2.35), we obtain significantly sharper results:

$$\begin{aligned} &\text{(i)} \quad P_\theta\{N_{Q_{1,1}} \le \epsilon n_1^*\} = O(n_1^{*-m_0+1}) \text{ with } 0 < \epsilon < 1 \text{ fixed;} \\ &\text{(ii)} \quad H \equiv n_1^{*-1/2}(N_{Q_{1,1}} - n_1^*) \xrightarrow{\pounds} N(0, \tfrac{1}{2}); \\ &\text{(iii)} \quad |H|^\kappa \text{ is uniformly integrable if } m_0 > 1 + \tfrac{1}{2}\kappa \text{ with } \kappa(>0) \text{ fixed,} \\ &\text{(iv)} \quad E_\theta[N_{Q_{1,1}} - n_1^*] \approx 0.1338 + o(1) \text{ if } m_0 \ge 3. \end{aligned} \tag{2.2.36}$$

2.2.2.1. Asymptotic Second-Order Regret Under Generality

In the spirit of Section 2.2.1.2, we consider a general multi-stage sampling strategy leading to a terminal stopping time $M \equiv M_{c_1}$, defined appropriately, so that we are assured "closeness" between M_{c_1} and n_1^*. Let us now list a number of *assumptions* that may be needed in order to draw certain specific conclusion(s):

(B1): For every fixed $n \geq m_0(>1)$, the pilot size, the event $[M_{c_1} = n]$ depends only on the statistic $\mathbf{Z}_n \equiv \{X_1 - X_n, ..., X_{n-1} - X_n\}$;

(B2): $M_{c_1}/n_1^* \xrightarrow{P_{\boldsymbol{\theta}}} 1$ as $c_1 \to 0$, and also $P_{\boldsymbol{\theta}}\{M_{c_1} \leq \varepsilon n_1^*\} = O(n_1^{*-a_1 m_0 + a_2})$ if $0 < \varepsilon < 1$, as $d \to 0$, with some $a_1 > 0$, $a_2 \geq 0$ if $m_0 > a_2/a_1$;

(B3): $H \equiv n_1^{*-1/2}(M_{c_1} - n_1^*) \xrightarrow{\pounds} N(0, a_3)$ as $c_1 \to 0$, with some $a_3 > 0$;

(B4): $|H|^2$ is uniformly integrable if $m_0 > a_4(>0)$; and

(B5): $E_{\boldsymbol{\theta}}\left[M_{c_1} - n_1^*\right] = a_5 + o(1)$ as $c_1 \to 0$ if $m_0 > a_6(>0)$.

(2.2.37)

for every fixed μ, σ, and α.

Theorem 2.2.2. *Under a general asymptotically second-order efficient multi-stage minimum risk point estimation methodology* $(M_{c_1}, \overline{X}_{M_{c_1}})$, *under the standing assumptions (B1)-(B5) from (2.2.37), we have the following asymptotic second-order expansion of the regret as* $c_1 \to 0$:

$$\omega(c_1) = c_1 E_{\boldsymbol{\theta}}\left[(M_{c_1} - n_1^*)^2 M_{c_1}^{-1}\right] = a_3 c_1 + o(c_1) \text{ if } m_0 > \max\{a_6, a_4, a_1^{-1}(3 + a_2)\}. \tag{2.2.38}$$

Proof: **Assumption (B5)** states that our general multi-stage methodology is asymptotically second-order efficient in the Ghosh-Mukhopadhyay (1981) sense.

Now, we can express

$$E_{\boldsymbol{\theta}}\left[\frac{(M_{c_1}-n_1^*)^2}{M_{c_1}}\right] = E_{\boldsymbol{\theta}}\left[\frac{(M_{c_1}-n_1^*)^2}{M_{c_1}} I(M_{c_1} \leq \tfrac{1}{2}n_1^*)\right] + E_{\boldsymbol{\theta}}\left[\frac{(M_{c_1}-n_1^*)^2}{M_{c_1}} I(M_{c_1} > \tfrac{1}{2}n_1^*)\right]. \tag{2.2.39}$$

Using **assumption (B2)** we obtain:

$$E_{\boldsymbol{\theta}}\left[\frac{(M_{c_1}-n_1^*)^2}{M_{c_1}}I(M_{c_1}\le\tfrac{1}{2}n_1^*)\right]\le\tfrac{5}{4}n_1^{*2}P_{\boldsymbol{\theta}}\{M_{c_1}\le\tfrac{1}{2}n_1^*\}=O\left(n_1^{*2-a_1m_0+a_2}\right)=o(1),\qquad(2.2.40)$$

if we pick $m_0>(2+a_2)/a_1$.

Next, obviously, we can write:

$$\frac{(M_{c_1}-n_1^*)^2}{M_{c_1}}I(M_{c_1}>\tfrac{1}{2}n_1^*)\le 2\frac{(M_{c_1}-n_1^*)^2}{n_1^*}.$$

But, under **assumption (B4),** $\frac{(M_{c_1}-n_1^*)^2}{n_1^*}$ is uniformly integrable so that $\frac{(M_{c_1}-n_1^*)^2}{M_{c_1}}I(M_{c_1}>\frac{1}{2}n_1^*)$ is then uniformly integrable.

Next, **assumption (B3)** provides: $\frac{M_{c_1}-n_1^*}{n_1^{*1/2}}\xrightarrow{\pounds}N(0,a_3)$ so that $\frac{(M_{c_1}-n_1^*)^2}{M_{c_1}}I(M_{c_1}>\frac{1}{2}n_1^*)\xrightarrow{\pounds}a_3\chi_1^2$. Hence, we conclude:

$$E_{\boldsymbol{\theta}}\left[\frac{(M_{c_1}-n_1^*)^2}{M_{c_1}}I(M_{c_1}>\tfrac{1}{2}n_1^*)\right]=a_3+o(1).$$

Now, putting together (2.2.39)-(2.2.41) completes the proof. ■

Clearly, the expansion of the regret function, $\omega_{Q_{1,1}}(c_1)$, follows immediately from Theorem 2.2.2. We leave out other details for brevity.

2.3. Sampling k Observations Gathered At-a-time in A Group With Selected Preliminaries

In the implementation of the purely sequential methodologies $\mathcal{P}_{1,1}$ and $Q_{1,1}$ from (2.2.1) and (2.2.27) respectively, we had emphasized recording one additional observation at-a-time, as needed, until termination. That is indeed customary. But, gathering a group of k observations may be called for in some practical situations. For example, items such as batteries, pens are customarily sold, for example, in packages of 6, 10 or 12 whereas fruit and vegetables are frequently sold in cases with 12 items or more.

From Section 2.1.2, recall that each item individually may cost c_1 unit whereas when one purchases k items at-a-time, suppose that each item cost $c_k(>0)$ unit. In many practical scenarios, c_k will be substantially smaller than c_1. This amounts to customary wisdom: The cost $(= kc_k)$ of purchasing a package made up of k items will be substantially smaller than the cost $(= kc_1)$ of purchasing k items individually. So, from the outset, we may safely assume:

$$c_k \le c_{k-1} \le ... \le c_1 \text{ with arbitrary but fixed } k \ge 1. \tag{2.3.1}$$

Certainly, it will be more reasonable to assume that these inequalities are strict, however, our new methodologies developed subsequently go through just as well with possible "equalities" included in (2.3.1).

Why should one consider sampling k observations in groups at-a-time? Our motivation is simple: We assume that the observations may keep arriving in groups of size k at-a-time. So, there may not be a possibility or practicality of gathering one observation at-a-time.

Sampling in batches of equal size has been successfully explored in a recent paper of Malinovsky and Zacks (2018) handling proportional closeness estimation of probability of contamination under a group testing methodology. Admittedly, our present work has an entirely different flavor from what Malinovsky and Zacks (2018) had considered, but perhaps there is a common but distant thread (sampling in batches) among these areas.

2.3.1. New Classes of Unbiased Estimators of the Scale

Suppose that we are allowed to gather $k(\geq 2)$ observations at-a-time. Now, having recorded n groups of independent observation vectors,

$$(X_{i1}, X_{i2}, ..., X_{ik}), \ i = 1, 2, ..., n,$$

from a $N(\mu, \sigma^2)$ population, the total sample size is obviously kn.

Now, in certain scenarios, especially when there may be a good chance of encountering outlying observations, the sample variance may not be a good estimator of the population variance. We therefore propose using newer *unbiased* and *consistent estimators* of (i) σ^2, namely, $U_{k,n}^{(j)}$, $j = 1, 2$ and (ii) σ, namely, $T_{k,n}^{(i)}$, $i = 1, 2$, to be incorporated in defining a range of alternative purely sequential (i) fixed-width confidence interval methodologies and (ii) the minimum risk point estimation methodologies respectively.

We formally propose to adapt the following *unbiased* estimators:

$$\begin{aligned}
\widehat{\sigma^2} \text{ in FWCI:} \quad & U_{k,n}^{(1)} \equiv \tfrac{1}{n} \textstyle\sum_{i=1}^{n} \tfrac{(k-1)}{k} \left\{ X_{i1} - \tfrac{1}{k-1} \textstyle\sum_{j=2}^{k} X_{ij} \right\}^2; \\
& U_{k,n}^{(2)} \equiv \tfrac{(k-1)\pi}{nk(2n-2+\pi)} \left\{ \textstyle\sum_{i=1}^{n} \left| X_{i1} - \tfrac{1}{k-1} \textstyle\sum_{j=2}^{k} X_{ij} \right| \right\}^2. \\
\widehat{\sigma} \text{ in MRPE:} \quad & T_{k,n}^{(1)} \equiv \Gamma(\tfrac{n}{2}) \left\{ \Gamma(\tfrac{n+1}{2}) \right\}^{-1} \left\{ \tfrac{k-1}{2k} \textstyle\sum_{i=1}^{n} \left(X_{i1} - \tfrac{1}{k-1} \textstyle\sum_{j=2}^{k} X_{ij} \right)^2 \right\}^{1/2}; \\
& T_{k,n}^{(2)} \equiv \tfrac{1}{n} \textstyle\sum_{i=1}^{n} \sqrt{\tfrac{(k-1)\pi}{2k}} \left\{ \left| X_{i1} - \tfrac{1}{k-1} \textstyle\sum_{j=2}^{k} X_{ij} \right| \right\}.
\end{aligned} \tag{2.3.2}$$

Lemma 2.3.1. *For every fixed* $\mu, \sigma, k(> 1)$ *and* n, the *estimators defined in (2.3.2) are unbiased and consistent for the respective parameters.*

Proof: We will treat the cases $i = 1, 2$ and $j = 1, 2$ separately.

Case 1. $E_{\boldsymbol{\theta}}[U_{k,n}^{(1)}] = \sigma^2$: It is easy to see that $X_{i1} - \frac{1}{k-1}\sum_{j=2}^{k} X_{ij}$, $i = 1, ..., n$, are i.i.d. random variables from $N(0, \frac{k}{k-1}\sigma^2)$. Let

$$W_i = \tfrac{k-1}{k\sigma^2}(X_{i1} - \tfrac{1}{k-1}\textstyle\sum_{j=2}^{k} X_{ij})^2. \tag{2.3.3}$$

Then, the W_i's are i.i.d. χ_1^2 random variables. Notice that $U_{k,n}^{(1)} = \frac{1}{n}\sum_{i=1}^{n} \sigma^2 W_i$ where the W_i's come from (3.3). Obviously, we have:

$$E_{\boldsymbol{\theta}}\left[U_{k,n}^{(1)}\right] = E_{\boldsymbol{\theta}}\left[\tfrac{1}{n}\textstyle\sum_{i=1}^{n} \sigma^2 W_i\right] = \sigma^2.$$

Clearly, $U_{k,n}^{(1)} \overset{P_{\boldsymbol{\theta}}}{\to} \sigma^2$ as $n \to \infty$.

Case 2. $E_{\boldsymbol{\theta}}[U_{k,n}^{(2)}] = \sigma^2$: We observe that

$$\begin{aligned}
&\left(\textstyle\sum_{i=1}^{n} |X_{i1} - \tfrac{1}{k-1}{}_{j=2}^{\,k}X_{ij}|\right)^2 \\
&\quad = \textstyle\sum_{i=1}^{n}(X_{i1} - \tfrac{1}{k-1}{}_{j=2}^{\,k}X_{ij})^2 \\
&\qquad + \textstyle\sum\sum_{i \neq i'}^{n} \left\{|X_{i1} - \tfrac{1}{k-1}{}_{j=2}^{\,k}X_{ij}|\right\}\left\{|X_{i'1} - \tfrac{1}{k-1}{}_{j=2}^{\,k}X_{i'j}|\right\}.
\end{aligned}$$

But, since $(X_{i1}, X_{i2}, ..., X_{ik})$ are independent of $(X_{i'1}, X_{i'2}, ..., X_{i'k})$ for $i \neq i'$, we can express:

$$\begin{aligned}
&E_{\boldsymbol{\theta}}\left[\left(\textstyle\sum_{i=1}^{n} |X_{i1} - \tfrac{1}{k-1}{}_{j=2}^{\,k}X_{ij}|\right)^2\right] \\
&\quad = nE_{\boldsymbol{\theta}}\left[(X_{11} - \tfrac{1}{k-1}{}_{j=2}^{\,k}X_{1j})^2\right] \\
&\qquad + n(n-1)E_{\boldsymbol{\theta}}\left[|X_{11} - \tfrac{1}{k-1}{}_{j=2}^{\,k}X_{1j}|\right] E_{\mu,\sigma}\left[|X_{21} - \tfrac{1}{k-1}{}_{j=2}^{\,k}X_{2j}|\right].
\end{aligned}$$

We also know:

$$E_{\boldsymbol{\theta}}\left[(X_{i1}-\tfrac{1}{k-1}{}_{j=2}^{k}X_{ij})^2\right]=\tfrac{k\sigma^2}{k-1};\text{ and }E_{\boldsymbol{\theta}}\left[\left|X_{i1}-\tfrac{1}{k-1}{}_{j=2}^{k}X_{ij}\right|\right]=\sqrt{\tfrac{2k\sigma^2}{(k-1)\pi}}.$$

Hence, we obtain:

$$E_{\boldsymbol{\theta}}\left[\left(\sum_{i=1}^{n}\left|X_{i1}-\tfrac{1}{k-1}{}_{j=2}^{k}X_{ij}\right|\right)^2\right]=\tfrac{nk(2n-2+\pi)}{(k-1)\pi}\sigma^2,$$

which implies: $E_{\boldsymbol{\theta}}\left[U_{k,n}^{(2)}\right]=\sigma^2$. Clearly, $U_{k,n}^{(2)}\overset{P_{\boldsymbol{\theta}}}{\to}\sigma^2$ as $n\to\infty$.

Case 3. $E_{\boldsymbol{\theta}}[T_{k,n}^{(1)}]=\sigma$: With W_i's coming from (2.3.3), we note that $Y={}_{i=1}^{n}W_i\sim\chi_n^2$, which gives:

$$E_{\boldsymbol{\theta}}\left[Y^{1/2}\right]=\frac{2^{(n+1)/2}\Gamma(\frac{n+1}{2})}{2^{n/2}\Gamma(\frac{n}{2})}=\sqrt{2}\Gamma(\tfrac{n+1}{2})\left\{\Gamma(\tfrac{n}{2})\right\}^{-1}.$$

That is, we have:

$$E_{\boldsymbol{\theta}}\left[T_{k,n}^{(1)}\right]=\Gamma(\tfrac{n}{2})\left\{\Gamma(\tfrac{n+1}{2})\right\}^{-1}E_{\boldsymbol{\theta}}\left[\sqrt{\tfrac{\sigma^2}{2}Y}\right]=\sigma.$$

Clearly, $T_{k,n}^{(1)}\overset{P_{\boldsymbol{\theta}}}{\to}\sigma$ as $n\to\infty$.

Case 4. $E_{\boldsymbol{\theta}}[T_{k,n}^{(2)}]=\sigma$: Again, starting with $X_{i1}-\frac{1}{k-1}\sum_{j=2}^{k}X_{ij}$ as in (2.3.3), we can express:

$$E_{\boldsymbol{\theta}}\left[T_{k,n}^{(2)}\right]=\sqrt{\tfrac{\pi}{2}}E_{\boldsymbol{\theta}}\left[\sqrt{\tfrac{(k-1)}{k}}\left|X_{11}-\tfrac{1}{k-1}\sum_{j=2}^{k}X_{1j}\right|\right]=\sigma.$$

Clearly, $T_{k,n}^{(1)}\overset{P_{\boldsymbol{\theta}}}{\to}\sigma$ as $n\to\infty$.

We omit additional details of proofs of the consistency properties, largely relying upon weak law of large numbers, for brevity. Lemma 2.3.1 is now proved. ■

2.4. Purely Sequential FWCI Methodologies Under Sampling in Groups

We consider sampling $k(\geq 2)$ observations at-a-time, recorded in groups of i.i.d. observation vectors:

$$(X_{i1}, X_{i2}, ..., X_{ik}), \ i = 1, 2, ..., n, ...$$

from a $N(\mu, \sigma^2)$ population. Having recorded $(X_{i1}, X_{i2}, ..., X_{ik})$, $i = 1, 2, ..., n$, we denote the sample mean and the associated confidence interval for μ as follows:

$$\overline{X}_{k,n} \equiv (nk)^{-1}{}_{i=1}^{\,n}{}_{j=1}^{\,k} X_{ij} \text{ and } J_{k,n} = [\overline{X}_{k,n} - d, \overline{X}_{k,n} + d]. \tag{2.4.1}$$

We hold k fixed, but otherwise leave it rather arbitrary.

We begin with

$$(X_{11}, X_{12}, ..., X_{1k}), ..., (X_{m_0 1}, X_{m_0 2}, ..., X_{m_0 k}),$$

a pilot data of k-tuples of size m_0 (> 1), and then record additional k-tuples of the X's at-a-time as needed. The confidence coefficient associated with $J_{k,n}$ from (2.4.1) is expressed as:

$$P_{\boldsymbol{\theta}}\{\mu \in J_{k,n}\} = P_{\boldsymbol{\theta}}\left\{|\overline{X}_{k,n} - \mu| \leq d\right\} = 2\Phi\left(\frac{\sqrt{kn}d}{\sigma}\right) - 1. \tag{2.4.2}$$

Observe that $J_{k,n}$ already has the required fixed-width, $2d$. But, we also require that the associated confidence coefficient must be nearly (at least) $1 - \alpha$. Thus, we must have:

$$2\Phi\left(\sqrt{kn}d/\sigma\right) - 1 \geq 1 - \alpha = 2\Phi(a) - 1,$$

where recall that a is the upper $50\alpha\%$ point of a $N(0,1)$ distribution as in (2.1.4). From (2.4.2), we then claim that the required number of k-tuples

$$\text{must be the smallest } n \geq a^2\sigma^2/(d^2k) \equiv C_k, \text{ say.} \tag{2.4.3}$$

This C_k is the optimal required fixed number of k-tuples had σ^2 been known. The magnitude of C_k, however, remains unknown. Hence, we define two seemingly analogous purely sequential stopping times in the spirit of (2.2.1) as:

$$\begin{array}{l}\text{Methodology } \mathcal{P}_{1,k}\text{:} \quad N_{\mathcal{P}_{1,k},d} \equiv N_{\mathcal{P}_{1,k}} = \inf\left\{n \geq m_0(\geq 2); \quad n \geq a^2U_{k,n}^{(1)}/(kd^2)\right\};\\ \text{Methodology } \mathcal{P}_{2,k}\text{:} \quad N_{\mathcal{P}_{2,k},d} \equiv N_{\mathcal{P}_{2,k}} = \inf\left\{n \geq m_0(\geq 2); \quad n \geq a^2U_{k,n}^{(2)}/(kd^2)\right\};\end{array} \tag{2.4.4}$$

with $U_{k,n}^{(i)}$, $i = 1, 2$, coming from (2.3.2). Upon termination, we have the associated final dataset, namely,

$$\left\{N_{\mathcal{P}_{i,k},d}, (X_{j1}, X_{j2}, ..., X_{jk}), j = 1, ..., N_{\mathcal{P}_{i,k},d}\right\}, i = 1, 2.$$

Obviously, $N_{\mathcal{P}_{1,k},d}$ and $N_{\mathcal{P}_{2,k},d}$ are both finite w.p.1 (Chow and Robbins, 1965) for all fixed $\boldsymbol{\theta}, d, \alpha, k$, and m_0.Then, the terminal estimation strategies can be summarized as follows:

$$\left(N_{\mathcal{P}_{i,k},d}, J_{k,N_{\mathcal{P}_{i,k}},d} = \left[\overline{X}_{k,N_{\mathcal{P}_{i,k}},d} \pm d\right]\right) \text{ associated with } \mathcal{P}_{i,k} \text{ from (2.4.4), } i = 1, 2. \tag{2.4.5}$$

2.4.1. Preliminaries and Selected First-Order Properties

We begin by stating a number of interesting results without too many specific details for their proofs in Lemmas 2.4.1-2.4.4. In the case of Lemma 2.4.5, we briefly outline a proof.

Lemma 2.4.1. *Under the stopping rules* $N_{\mathcal{P}_{i,k},d}$ *defined in (2.4.4), for all fixed* $\boldsymbol{\theta}, d, \alpha, k,$ *and* $m_0(\geq 2)$, *the statistics* $\left\{U_{k,n}^{(i)},\ i=1,2\right\}$ *and* $\overline{X}_{k,n}$ *are distributed independently for each fixed* $n \geq m_0, i = 1, 2$.

Lemma 2.4.2. *Under the stopping rules* $N_{\mathcal{P}_{i,k},d}$ *defined in (2.4.4), for all fixed* $\boldsymbol{\theta}, d, \alpha, k,$ *and* $m_0(\geq 2)$, *we have:*

$$E_{\boldsymbol{\theta}}\left[\overline{X}_{k,N_{\mathcal{P}_{i,k}},d}\right] = \mu \text{ and } V_{\boldsymbol{\theta}}\left[\overline{X}_{k,N_{\mathcal{P}_{i,k}},d}\right] = k^{-1}\sigma^2 E_{\boldsymbol{\theta}}\left[N_{\mathcal{P}_{i,k},d}^{-1}\right],$$

$i = 1, 2.$

Lemma 2.4.3. *Define* $H_i \equiv N_{\mathcal{P}_{i,k},d}^{1/2}\left(\overline{X}_{k,N_{\mathcal{P}_{i,k}},d} - \mu\right)/\sigma$ *under the stopping rules* $N_{\mathcal{P}_{i,k},d}$ *from (2.4.4). Then, for all fixed* $\boldsymbol{\theta}, d, \alpha, k,$ *and* $m_0(\geq 2)$, *we have:*

$$H_i \sim N(0,1),\ i = 1, 2.$$

Lemma 2.4.4. *For the stopping time* $N_{\mathcal{P}_{i,k},d}$ *defined in (2.4.4), for all fixed* $\boldsymbol{\theta}, \alpha, k,$ *and* $m_0(\geq 2)$, *we have as* $d \to 0$:

$$\begin{aligned} &(i) \quad N_{\mathcal{P}_{i,k},d}/C_k \xrightarrow{P_{\boldsymbol{\theta}}} 1; \text{ and} \\ &(ii) \quad E_{\boldsymbol{\theta}}\left[N_{\mathcal{P}_{i,k},d}/C_k\right] \to 1 \text{ [Asymptotic first-order efficiency]}; \end{aligned} \tag{2.4.6}$$

$i = 1, 2$, *parallel to (2.2.3), where* $C_k(= a^2\sigma^2/(d^2k))$ *defined in (2.4.3) is the optimal fixed number of groups.*

Lemma 2.4.5. *For the stopping time* $N_{\mathcal{P}_{i,k}}, d$ *from (2.4.4) and the proposed fixed-width confidence interval estimation strategy* $\left(N_{\mathcal{P}_{i,k},d}, J_{k,N_{\mathcal{P}_{i,k},d}} = \left[\overline{X}_{k,N_{\mathcal{P}_{i,k}},d} \pm d\right]\right)$ *from (2.4.5), for all fixed* $\boldsymbol{\theta}, \alpha, k,$ *and* $m_0(\geq 2)$, *we have*:

$$\lim_{d \to 0} P_{\boldsymbol{\theta}} \left\{ \mu \in J_{k,N_{\mathcal{P}_{i,k},d}} \right\} = 1 - \alpha \quad [\textit{Asymptotic consistency}], \tag{2.4.7}$$

$i = 1, 2.$

Proof: We express the coverage probability as:

$$P_{\boldsymbol{\theta}} \left\{ \mu \in J_{k,N_{\mathcal{P}_{i,k},d}} \right\} = E_{\boldsymbol{\theta}} \left\{ 2\Phi \left(\sqrt{kN_{\mathcal{P}_{i,k},d}}\, d/\sigma \right) - 1 \right\}.$$

Now, Lemma 2.4.4, part (i) and dominated convergence theorem together complete the proof.

■

2.4.2. Asymptotic Second-Order Properties

In this section, Theorems 2.4.1-2.4.2 summarize a number of major asymptotic second-order properties associated with the stopping rules $\mathcal{P}_{1,k}$ and $\mathcal{P}_{2,k}$ from (2.4.4) and the associated estimation strategy from (2.4.5). However, before we state these theorems, we begin by laying down a number of essential technical details.

Case 1: $i = 1$

First, we rewrite the stopping time corresponding to $\mathcal{P}_{1,k}$ from (2.4.4) as follows:

$$N_{\mathcal{P}_{1,k},d} \equiv N_{\mathcal{P}_{1,k}} = \inf \left\{ n \geq m_0 : \quad C_k^{-1} n^2 \geq \Sigma_{i=1}^{n} W_i \right\}, \tag{2.4.8}$$

where the W_i's are distributed as i.i.d. χ_1^2. By comparing this representation with that in (2.4.8) and (2.1.9), we immediately have:

$$\lambda = 1, \tau^2 = 2, h^* = C_k^{-1}, n^* = C_k, \delta = 2, \beta^* = 1, l_0 = 0, p = 2, \text{ and } b = \tfrac{1}{2}.$$

We again apply nonlinear renewal theory from Section 2.1.3 to express:

$$\begin{aligned} D_{1,k}^{(1)} &\equiv D_1^{(1)} = \Sigma_{n=1}^{\infty} n^{-1} E\left\{\max(0, \chi_n^2 - 2n)\right\}; \\ \nu_{1,k}^{(1)} &\equiv \nu_1^{(1)} = \tfrac{3}{2} - D_1^{(1)}; \\ \eta_{1,k}^{(1)} &\equiv \eta_1^{(1)} = -\tfrac{1}{2} - D_1^{(1)}. \end{aligned} \tag{2.4.9}$$

Naturally, these correspond to $U_{k,n}^{(1)}$. The superscript (1) corresponds to the FWCI problem.

From (2.2.8), we recall that we had found:

$$D_{1,k}^{(1)} \equiv D_1^{(1)} \approx 0.6839 \text{ with its estimated s.e. } 0.00004.$$

Observe that indeed the three entities $D_{1,k}^{(1)}, \nu_{1,k}^{(1)}, \eta_{1,k}^{(1)}$ are free from k and we have:

$$\nu_{1,k}^{(1)} \equiv \nu_1^{(1)} \approx 0.8161 \text{ and } \eta_{1,k}^{(1)} \equiv \eta_1^{(1)} \approx -1.1839. \tag{2.4.10}$$

Case 2: $i = 2$

Along the lines of (2.4.8), the stopping time $N_{P_{2,k},d}$ associated with $\mathcal{P}_{2,k}$ from (2.4.4) is rewritten as:

$$N_{P_{2,k}} = \inf\left\{n \geq m_0 : \quad \sum_{i=1}^{n} Y_i \leq n^{3/2} C_k^{-1/2}\left(1 + \left(\tfrac{\pi}{4} - \tfrac{1}{2}\right) n^{-1} + o(n^{-1})\right)\right\}, \qquad (2.4.11)$$

where $Y_i = \left(\frac{1}{2}\pi W_i\right)^{1/2}$ and W_i's were defined in (2.3.3).

Next, for arbitrary $u(> 0)$, we obtain:

$$P\left\{Y_1 \leq u\right\} = \int_0^{\pi u^2/2} (2\pi y)^{-1/2} e^{-y/2} dy \leq Bu,$$

with some appropriate $B(> 0)$ not involving u.

Again, by comparing the representation (2.4.11) with that in (2.1.9) or (2.2.4), we immediately have:

$$\lambda = 1, \tau^2 = \tfrac{1}{2}\pi - 1, h^* = C_k^{-1/2}, n^* = C_k, \delta = \tfrac{3}{2}, \beta^* = 2,$$

$$l_0 = \tfrac{1}{4}\pi - \tfrac{1}{2}, p = 2\pi - 4, \text{ and } b = 1.$$

We apply nonlinear renewal theory from Section 2.1.3 to express:

$$\begin{aligned} D_{2,k}^{(1)} &\equiv D_2^{(1)} =_{n=1}^{\infty} n^{-1} E\left\{\max\left(0,_{i=1}^{n} Y_i - \tfrac{3}{2}n\right)\right\}; \\ \nu_{2,k}^{(1)} &\equiv \nu_2^{(1)} = -\tfrac{3}{4} + \tfrac{1}{2}\pi - D_2^{(1)}; \\ \eta_{2,k}^{(1)} &\equiv \eta_2^{(1)} = \tfrac{5}{2} - \pi - 2D_2^{(1)}. \end{aligned} \qquad (2.4.12)$$

Naturally, these correspond to $U_{k,n}^{(2)}$. Again, the superscript (1) corresponds to the FWCI problem.

In the spirit of (2.2.8), we performed large-scale simulations to come up with the following estimated values:

$$D_2^{(1)} \approx 0.3186 \text{ with its estimated s.e. } 0.00004.$$

Observe that indeed $D_{2,k}^{(1)}, \nu_{2,k}^{(1)}, \eta_{2,k}^{(1)}$ are free from k and we obtain:

$$\nu_{2,k}^{(1)} \approx 0.5022 \text{ and } \eta_{2,k}^{(1)} \approx -1.2788 \text{ with estimated s.e. } 0.00008. \tag{2.4.13}$$

Theorem 2.4.1. *For the stopping time $N_{\mathcal{P}_{i,k},d}$ defined by (2.4.4), we denote $H_i = C_k^{-1/2}(N_{\mathcal{P}_{i,k},d} - C_k)$, $i = 1, 2$. Then, for every fixed $\boldsymbol{\theta}, k$ and α, we have the following results as $d \to 0$:*

$$(i) \quad \left.\begin{array}{l} (a)\ P_{\boldsymbol{\theta}}\{N_{\mathcal{P}_{1,k},d} \leq \varepsilon C_k\} = O(C_k^{-m_0/2}) \\ (b)\ P_{\boldsymbol{\theta}}\{N_{\mathcal{P}_{2,k},d} \leq \varepsilon C_k\} = O(C_k^{-m_0}) \end{array}\right\} \text{ with fixed } 0 < \varepsilon < 1 \text{ and } m_0 \geq 2;$$

$$(ii) \quad |H_i|^\kappa \text{ is uniformly integrable if } m_0 > \max\{1, \kappa\} \text{ with fixed } \kappa(>0),\ i = 1, 2;$$

$$(iii) \quad \left.\begin{array}{l} (a)\ H_1 \xrightarrow{\pounds} N(0, 2)\ i = 1 \\ (b)\ H_2 \xrightarrow{\pounds} N(0, 2\pi - 4) \end{array}\right\} \text{ if } m_0 \geq 2;$$

$$(iv) \quad \left.\begin{array}{l} (a)\ E_{\boldsymbol{\theta}}\left[N_{\mathcal{P}_{1,k},d} - C_k\right] = \eta_{1,k}^{(1)} + o(1) \\ (b)\ E_{\boldsymbol{\theta}}\left[N_{\mathcal{P}_{2,k},d} - C_k\right] = \eta_{2,k}^{(1)} + o(1) \end{array}\right\} \text{ if } m_0 \geq 3;$$

with C_k, $\eta_{1,k}^{(1)}(\approx -1.1839)$, and $\eta_{2,k}^{(1)}(\approx -1.2788)$ coming from (2.4.3), (2.4.10), and (2.4.13) respectively.

Proof: Having exhibited the technical details in (2.4.8)-(2.4.13), only a very brief outline is warranted here. One should follow along the basic layout from Section 2.1.3. Analogously, the results from (2.2.4)-(2.2.9) will also be helpful. ■

Theorem 2.4.2. *For the stopping time $N_{\mathcal{P}_{i,k},d}$ defined by (2.4.4) with the proposed fixed-width confidence interval estimation strategy $\left(N_{\mathcal{P}_{i,k},d}, J_{k,N_{\mathcal{P}_{i,k},d}} = \left[\overline{X}_{k,N_{\mathcal{P}_{i,k},d}} \pm d\right]\right)$ from (2.4.5), for every fixed $\boldsymbol{\theta}, k$ and α, we have the following results as $d \to 0$:*

$$P_{\boldsymbol{\theta}}\left\{\mu \in J_{k,N_{\mathcal{P}_{i,k},d}} = \left[\overline{X}_{k,N_{\mathcal{P}_{i,k},d}} \pm d\right]\right\} = (1-\alpha) + \gamma_i C_k^{-1} + o(d^2),$$

[*Asymptotic second-order consistency*],

with $p_1 = 2$, $p_2 = 2\pi - 4$, $\gamma_i = \left\{\eta_{i,k}^{(1)} - \frac{1}{4}p_i\left(a^2+1\right)\right\} a\phi(a)$, for $i = 1, 2$, if $m_0 \geq 6$.

Proof: This result follows immediately from the more general statement given in Theorem 2.2.1 by explicitly matching the results from Theorem 2.4.1 with the assumptions (A1)-(A5) gathered in (2.2.13) upon identifying all requisite terms. We found:

$$\gamma_1 \approx -0.4125 \text{ and } \gamma_2 \approx -0.4630, \tag{2.4.14}$$

numerically. ■

2.5. Purely Sequential MRPE Methodologies Under Sampling in Groups

We again consider sampling $k(\geq 2)$ observations at-a-time, recorded in groups of i.i.d. observation vectors:

$$\left(X_{i1}, X_{i2}, ..., X_{ik}\right),\ i = 1, 2, ..., n, ...$$

from a $N(\mu, \sigma^2)$ population. We hold k fixed, but otherwise leave it rather arbitrary.

Having recorded $(X_{i1}, X_{i2}, ..., X_{ik}), i = 1, 2, ..., n$, we propose to estimate μ with $\overline{X}_{k,n} = \frac{1}{kn}\sum_{i=1}^{n}\sum_{j=1}^{k} X_{ij}$ under a loss function in the spirit of (2.1.6) which is composed of squared error loss plus a linear cost due to sampling as follows:

$$L_{k,n}(\mu, \overline{X}_{k,n}) = A(\overline{X}_{k,n} - \mu)^2 + c_k kn, \tag{2.5.1}$$

where $c_k(>0)$ is the known cost per unit observation and $A(>0)$ is also assumed known.

The associated fixed-sample-size risk function can be expressed as:

$$R_{k,n} = E_{\mu,\sigma}[L_{k,n}(\mu, \overline{X}_{k,n})] = A\sigma^2(kn)^{-1} + kc_k n, \tag{2.5.2}$$

Our goal is to construct the minimum risk point estimator of μ which leads to the optimal fixed number of groups, n_k^*, had σ been known:

$$\begin{gathered} n \text{ is the minimum number of groups } \geq \tfrac{1}{k}\sqrt{A/c_k}\sigma = n_k^*, \text{ say, associated} \\ \text{with the minimum risk given by } R_{k,n_k^*} \equiv R_{n_k^*}(c_k) = 2kc_k n_k^*. \end{gathered} \tag{2.5.3}$$

We start with $(X_{11}, X_{12}, ..., X_{1k}), ..., (X_{m_0 1}, X_{m_0 2}, ..., X_{m_0 k})$, a pilot data of k-tuples of size $m_0(>1)$, and then record additional k-tuples of the X's at-a-time as needed. The stopping rule is defined as:

$$\begin{aligned} &\text{Methodology } Q_{1,k}\text{:} \quad N_{Q_{1,k},c_k} \equiv N_{Q_{1,k}} = \inf\left\{n \geq m_0(\geq 2): \quad n \geq \tfrac{1}{k}\sqrt{A/c_k}T_{k,n}^{(1)}\right\}; \\ &\text{Methodology } Q_{2,k}\text{:} \quad N_{Q_{2,k},c_k} \equiv N_{Q_{2,k}} = \inf\left\{n \geq m_0(\geq 2): \quad n \geq \tfrac{1}{k}\sqrt{A/c_k}T_{k,n}^{(2)}\right\}; \end{aligned} \tag{2.5.4}$$

with $T_{k,n}^{(i)}$, $i = 1, 2$, coming from (2.3.2).

That is, for example, if $m_0 \geq \frac{1}{k}\sqrt{A/c_k}T_{k,m_0}^{(i)}$ is satisfied, we do not take any additional k-tuple and the sampled number of groups is $N_{Q_i,k} = m_0$. Otherwise, we record next k-

tuple $(X_{m_0+1,1}, X_{m_0+1,2}, ..., X_{m_0+1,k})$ and obtain updated $T^{(i)}_{k,m_0+1}$ to check with the stopping rule (2.5.4). We terminate sampling at the first time $N_{Q_{1,k},c_k} = n(\geq m_0)$ such that $n \geq \frac{1}{k}\sqrt{A/c_k}T^{(i)}_{k,n}$ happens. Upon termination, we have the associated final dataset, namely,

$$\left\{N_{Q_{i,k},c_k}, (X_{j1}, X_{j2}, ..., X_{jk}), j = 1, 2, ..., N_{Q_{i,k},c_k}\right\}, i = 1, 2.$$

2.5.1. Preliminaries and Selected First-Order Properties

In the spirit of Section 2.4.1, we begin by stating a number interesting results without outlining too many specific details for their proofs.

Lemma 2.5.1. *Under the stopping time $N_{Q_{i,k},c_k}$ defined in (2.5.4), for all fixed $\boldsymbol{\theta}, c_k, A, k,$ and $m_0(\geq 2)$, the statistics $\left\{T^{(i)}_{k,n},\ i = 1, 2\right\}$ and $\overline{X}_{k,n}$ are distributed independently for each fixed $n \geq m_0, i = 1, 2$.*

Lemma 2.5.2. *Under the stopping time $N_{Q_{i,k},c_k}$ defined in (2.5.4), for all fixed $\boldsymbol{\theta}, c_k, A, k,$ and $m_0(\geq 2)$, we have:*

$$E_{\boldsymbol{\theta}}\left[\overline{X}_{k,N_{Q_{i,k},c_k}}\right] = \mu \text{ and } V_{\boldsymbol{\theta}}\left[\overline{X}_{k,N_{Q_{i,k},c_k}}\right] = k^{-1}\sigma^2 E_{\boldsymbol{\theta}}\left[N^{-1}_{Q_{i,k},c_k}\right],$$

$i = 1, 2.$

Lemma 2.5.3. *Define $H_i \equiv N^{1/2}_{Q_{i,k},c_k}\left(\overline{X}_{k,N_{Q_{i,k},c_k}} - \mu\right)/\sigma$ under the stopping rules $N_{Q_{i,k},c_k}$ from (2.5.4). Then, for all fixed $\boldsymbol{\theta}, c_k, A, k,$ and $m_0(\geq 2)$, we have:*

$$H_i \sim N(0, 1),\ i = 1, 2.$$

Lemma 2.5.4. *For the stopping time $N_{Q_{i,k},c_k}$ defined in (2.5.4), for all fixed $\boldsymbol{\theta}, A, k,$ and $m_0(\geq 2)$, we have as $c_k \to 0$:*

$$
\begin{aligned}
&(i) \quad N_{Q_{i,k},c_k}/n_k^* \xrightarrow{P_{\boldsymbol{\theta}}} 1; \text{ and} \\
&(ii) \quad E_{\boldsymbol{\theta}}\left[N_{Q_{i,k},c_k}/n_k^*\right] \to 1 \text{ [Asymptotic first-order efficiency]};
\end{aligned}
\tag{2.5.5}
$$

$i = 1, 2$, parallel to (2.2.31), where $n_k^*(= \frac{1}{k}\sqrt{A/c_k}\sigma)$ defined in (2.5.3) is the optimal fixed number of groups.

Upon termination, we propose to estimate μ by $\overline{X}_{k,N_{Q_{i,k},c_k}}$ obtained from the fully accrued dataset, namely,

$$
\left\{N_{Q_{i,k},c_k}, (X_{j1}, X_{j2}, ..., X_{jk}), j = 1, 2, ..., N_{Q_{i,k},c_k}\right\}, i = 1, 2.
$$

The risk associated with $\overline{X}_{k,N_{Q_{i,k},c_k}}$ can be expressed as:

$$
\begin{aligned}
R_{k,N_{Q_{i,k},c_k}}(c_k) &\equiv E_{\boldsymbol{\theta}}\left[L_{k,N_{Q_{i,k},c_k}}(\mu, \overline{X}_{k,N_{Q_{i,k},c_k}})\right] = AE_{\boldsymbol{\theta}}\left[(\overline{X}_{k,N_{Q_{i,k},c_k}} - \mu)^2\right] + kc_kE_{\boldsymbol{\theta}}\left[N_{Q_{i,k},c_k}\right] \\
&= A\sigma^2k^{-1}E_{\boldsymbol{\theta}}\left[N_{Q_{i,k},c_k}^{-1}\right] + kc_kE_{\boldsymbol{\theta}}\left[N_{Q_{i,k},c_k}\right], \; i = 1, 2, \text{ from Lemma 2.5.2.}
\end{aligned}
\tag{2.5.6}
$$

Then, the major requisite entities are given by:

$$
\begin{aligned}
\text{Risk efficiency:} \quad \xi_{Q_{i,k}}(c_k) &\equiv \frac{R_{k,N_{Q_{i,k}}}(c_k)}{R_{n_k^*}} = \tfrac{1}{2}E_{\boldsymbol{\theta}}[N_{Q_{i,k},c_k}n_k^{*-1}] + \tfrac{1}{2}E_{\boldsymbol{\theta}}[n_k^*N_{Q_{i,k},c_k}^{-1}], \\
\text{Regret:} \quad \omega_{Q_{i,k}}(c_k) &\equiv R_{k,N_{Q_{i,k}}}(c_k) - R_{n_k^*} = kc_kE_{\boldsymbol{\theta}}\left[\frac{(N_{Q_{i,k},c_k} - n_k^*)^2}{N_{Q_{i,k},c_k}}\right],
\end{aligned}
\tag{2.5.7}
$$

in the spirit of Robbins (1959) as in (2.2.30), $i = 1, 2$.

Next, we state a major result which shows that the methodology $Q_{i,k}$ is *asymptotically* *first-order risk efficient.*

Theorem 2.5.1. *For the stopping rule $Q_{i,k}$ defined in (2.5.4), for all fixed $\boldsymbol{\theta}, A$, and k, we have as $c_k \to 0$:*

$$\xi_{Q_{i,k}}(c_k) \to 1 \ \ [\textit{Asymptotic first-order efficiency}], \tag{2.5.8}$$

if $m_0 \geq 3$, where n_k^* and $\xi_{Q_{i,k}}(c_k)$ come from (2.5.3) and (2.5.7) respectively, $i = 1, 2$.

Proof: Results analogous to those covered by the **Assumptions (B1)-(B5)** as laid down in (2.2.37) hold under the stopping rule $Q_{i,k}$, $i = 1, 2$. Thus, we can verify that $E_{\boldsymbol{\theta}}[n_k^* N_{Q_{i,k},c_k}^{-1}] \to 1$ as $c_k \to 0$, if $m_0 \geq 3$, $i = 1, 2$. From (2.5.5), part (ii), we already know that $E_{\boldsymbol{\theta}}[N_{Q_{i,k},c_k} n_k^{*-1}] \to 1$ as $c_k \to 0$, if $m_0 \geq 2$, $i = 1, 2$. Hence, the result follows. Further details are left out for brevity. ■

2.5.2. Asymptotic Second-Order Properties

In this section, Theorems 2.5.2 - 2.5.3 summarize a number of major asymptotic second-order properties associated with the stopping rules $Q_{1,k}$ and $Q_{2,k}$ from (2.5.4). However, before we tackle these theorems, we begin by laying down a number of essential technical details.

Case 1: $i = 1$

First, we rewrite the stopping time corresponding to $Q_{1,k}$ from (2.5.4) as follows:

$$N_{Q_{1,k},c_k} \equiv N_{Q_{1,k}} = \inf\left\{n \geq m_0 : \quad \textstyle\sum_{i=1}^{n} W_i \leq n^3 n_k^{*-2} l_n\right\} \tag{2.5.9}$$

where W_i's are i.i.d. random variables as defined in (2.3.3) with $l_n = 1 - \frac{1}{2}n^{-1} + o(n^{-1})$.

Next, by comparing this representation with that in (2.1.9), we immediately have:

$$\lambda = 1, \tau^2 = 2, h^* = n_k^{*-2}, n^* = n_k^*, \delta = 3, \beta^* = \tfrac{1}{2}, l_0 = -\tfrac{1}{2}, p = \tfrac{1}{2}, \text{ and } b = \tfrac{1}{2}.$$

Then, we invoke the representation via nonlinear renewal theory from Section 1.3 to express:

$$\begin{aligned} D_{1,k}^{(2)} &\equiv D_1^{(2)} = \textstyle\sum_{n=1}^{\infty} n^{-1} E\left\{\max(0, \chi_n^2 - 3n)\right\}; \\ \nu_{1,k}^{(2)} &\equiv \nu_1^{(2)} = \tfrac{3}{2} - D_1^{(2)}; \\ \eta_{1,k}^{(2)} &\equiv \eta_1^{(2)} = \tfrac{1}{4} - \tfrac{1}{2} D_1^{(2)}. \end{aligned} \tag{2.5.10}$$

Naturally, these correspond to $T_{k,n}^{(1)}$. The superscript (2) corresponds to the MRPE problem.

We found the following estimates:

$$D_{1,k}^{(2)} \equiv D_1^{(2)} \approx 0.2339 \text{ with estimated s.e. } 0.000025,$$

so that we have:

$$\nu_{1,k}^{(2)} \equiv \nu_1^{(2)} \approx 1.2662 \text{ and } \eta_{1,k}^{(2)} \approx 0.1330 \text{ with estimated s.e. } 0.0000125. \tag{2.5.11}$$

Again, observe that indeed the entities $D_{1,k}^{(2)}$, $\nu_{1,k}^{(2)}$,$\eta_{1,k}^{(2)}$ are free from k.

Case 2: $i = 2$

Recall that $Y_i = \left(\frac{1}{2}\pi W_i\right)^{1/2}$ and W_i's were defined in (2.3.3). Then, we rewrite the stopping time corresponding to $Q_{2,k}$ from (2.5.4) as follows:

$$N_{Q_{2,k},c_k} \equiv N_{Q_{2,k}} = \inf\left\{n \geq m_0 : \quad \sum_{i=1}^{n} Y_i \leq n^2 n_k^{*-1} l(n)\right\}, \tag{2.5.12}$$

with $l(n) \equiv 1$ and W_i's are i.i.d. χ_1^2 random variables.

Next, by comparing this presentation with that in (2.1.9), we obtain:

$$\lambda = 1,\ \tau^2 = \tfrac{\pi}{2} - 1,\ h^* = n_k^{*-1},\ n^* = n_k^*,\ \delta = 2,\ \beta^* = 1,\ l_0 = 0,\ p = \tfrac{\pi}{2} - 1, b = 1$$

Then we incorporate nonlinear renewal theorem to express:

$$\begin{aligned} D_{2,k}^{(2)} &\equiv D_2^{(2)} = \textstyle\sum_{n=1}^{\infty} n^{-1} E\left\{\max\left(0, \sum_{i=1}^{n} \sqrt{\tfrac{\pi}{2} W_i} - 2n\right)\right\}; \\ \nu_{2,k}^{(2)} &\equiv \nu_2^{(2)} = \tfrac{1}{4}\pi - D_2^{(2)}; \\ \eta_{2,k}^{(2)} &\equiv \eta_2^{(2)} = 1 - \tfrac{1}{4}\pi - D_2^{(2)}. \end{aligned} \tag{2.5.13}$$

Naturally, these correspond to $T_{k,n}^{(2)}$. The superscript (2) corresponds to the MRPE problem. Using simulations, we have found:

$$D_{2,k}^{(2)} \equiv D_2^{(2)} \approx 0.0815 \text{ with estimated s.e. } 0.000259.$$

Again, observe that indeed the three entities $D_{2,k}^{(2)}, \nu_{2,k}^{(2)}, \eta_{2,k}^{(2)}$ are free from k and we have:

$$\nu_{2,k}^{(2)} \equiv \nu_2^{(2)} \approx 0.7038 \text{ and } \eta_{2,k}^{(2)} \equiv \eta_2^{(2)} \approx 0.1331 \text{ with estimated s.e. } 0.000259. \tag{2.5.14}$$

Theorem 2.5.2. *For the stopping time $N_{Q_{i,k},c_k}$ defined by (2.5.4), we denote $H_i = n_k^{*-1/2}(N_{Q_{i,k},c_k} -$*

$n_k^{*-1/2}$), $i = 1, 2$. *Then, for every fixed* $\boldsymbol{\theta}, c_k$ *and* d, *we have the following results as* $c_k \to 0$:

$$
\begin{array}{ll}
(i) & P_\theta\{N_{Q_{i,k},c_k} \le \varepsilon n_k^*\} = O(n_k^{*-m_0}) \text{ if } 0 < \epsilon < 1 \text{ and } m_0 \ge 2, i = 1, 2; \\
(ii) & |H_i|^\kappa \text{ is uniformly integrable if } m_0 > \max\{1, \tfrac{1}{2}\kappa\}; \\
(iii) & \left.\begin{array}{l} (a)\ H_1 \xrightarrow{\pounds} N(0, \tfrac{1}{2}) \\ (b)\ H_2 \xrightarrow{\pounds} N(0, \tfrac{1}{2}\pi - 1) \end{array}\right\} \text{ if } m_0 \ge 2; \\
(iv) & \left.\begin{array}{l} (a)\ E_\theta[N_{Q_{1,k},c_k} - n_k^*] = \eta_{1,k}^{(2)} + o(1) \\ (b)\ E_\theta[N_{Q_{2,k},c_k} - n_k^*] = \eta_{2,k}^{(2)} + o(1) \end{array}\right\} \text{ if } m_0 \ge 3.
\end{array}
$$

with $\eta_{1,k}^{(2)}(\approx 0.1330)$ *and* $\eta_{2,k}^{(2)}(\approx 0.1331)$ *coming from (2.5.11) and (2.5.14) respectively.*

Proof: Combining Section 2.1.3 with (2.5.10) and (2.5.13), the results follow from the nonlinear renewal theory. Details are omitted for brevity. ■

Theorem 2.5.3. *Given the regret function defined in (2.5.7), we have as* $c_k \to 0$:

$$
\omega_{Q_{i,k}}(c_k) \equiv R_{k,N_{Q_{i,k}}}(c_k) - R_{n_k^*} = kc_k q_i + o(c_k),
$$

with $q_1 = \frac{1}{2}$ *and* $q_2 = \frac{1}{2}\pi - 1$ *if* $m_0 \ge 4$.

Proof: This result follows immediately from the more general statement given in Theorem 2.2.2 by explicitly matching the results from Theorem 2.5.2 with the assumptions (B1)-(B5) gathered in (2.2.37) and upon identifying all requisite terms. Details are omitted for brevity.

■

Remark 2.5.1. This remark is relevant for both FWCI and MRPE problems developed in

Sections 2.4 and 2.5. In all ensuing estimation methodologies, one has surely noted that the random variables $\overline{X}_n$ and $I[N = n]$ are independently distributed for every $n \geq 2$ where we are referring to "N" as a generic notation for the terminal sample size or terminal number of groups. Such a broad ranging statement regarding independence of $\overline{X}_n$ and $I[N = n]$, for every $n \geq 2$, follows from Basu's (1955) theorem.

2.6. Data Analyses From Simulations

In this section, we investigate the performances of both FWCI and MRPE problems via computer simulations when the sample sizes are respectively small or medium to large.

2.6.1. Simulation Study on FWCI Problem

Next, we set out to compare the performances of the purely sequential fixed-width confidence interval estimation strategies defined via (2.2.1) with those under (2.4.4). To be precise, we briefly recall the stopping rules:

$$
\begin{aligned}
&(2.2.1)\ \mathcal{P}_{1,1}: \quad N_{\mathcal{P}_{1,1},d} = \inf\{n \geq m_0(\geq 2): \quad n \geq a^2 S_n^2/d^2\};\\
&(2.4.4)\ \mathcal{P}_{1,k}: \quad N_{\mathcal{P}_{1,k},d} = \inf\left\{n \geq m_0(\geq 2): \quad n \geq a^2 U_{k,n}^{(1)}/(kd^2)\right\};\\
&(2.4.4)\ \mathcal{P}_{2,k}: \quad N_{\mathcal{P}_{2,k},d} = \inf\left\{n \geq m_0(\geq 2): \quad n \geq a^2 U_{k,n}^{(2)}/(kd^2)\right\}.
\end{aligned} \tag{2.6.1}
$$

Having fixed $\alpha = 0.05$, pseudo random samples were generated from a $N(\mu = 5, \sigma^2 = 4)$ population. Recall that m_0 and C_k are respectively the pilot group number and the optimal fixed number of groups consisting of k observations each. For example, if $m_0 = 5$ and $k = 3$, we begin sampling with 5 groups each having in it 3 observations.

We fixed the following choices: $km_0 = 18, 48, 72$, $k = 1, 2, 3$, $kC_k = 90, 240, 600$, and d was determined accordingly from the expression of C_k. The FWCI methodologies from

(2.6.1) were then implemented one by one with $T = 10000$ independent runs under each configuration. In Tables 2.2-2.6, we would use the set of notation defined precisely in Table 2.1.

Table 2.1. The set of notation used in Tables 2.2-2.6: $T = 10000$

n_i	: terminal sample size in i^{th} run;
$\bar{n} = T^{-1}\sum_{i=1}^{T} n_i$	: ave sample size, should compare with C_k;
$s_{\bar{n}} = \left\{(T^2 - T)^{-1}\sum_{i=1}^{T}(n_i - \bar{n})^2\right\}^{1/2}$	: estimated standard error (s.e.) of $\bar{n}$;
$\bar{x}_{n_i}$	: terminal sample mean in i^{th} run;
$\bar{x} = T^{-1}\sum_{i=1}^{T}\bar{x}_{n_i}$	: combined sample ave, should compare with μ;
$s_x = \left\{(T^2 - T)^{-1}\sum_{i=1}^{T}(\bar{x}_{n_i} - \bar{x})^2\right\}^{1/2}$	: estimated s.e. of $\overline{x}$;
$p_i = I\left(\lvert\bar{x}_{n_i} - \mu\rvert \le d\right)$	: 1(or 0) if J_{n_i} covers (or does not cover) μ in i^{th} run;
$\bar{p} = T^{-1}\sum_{i=1}^{T} p_i$	: estimated cov probability, should compare with $1-\alpha$;
$z \equiv \bar{p} - (1-\alpha)$	: should compare with γC_k^{-1} from Theorem 2.4.2;
$s_{\bar{p}} \equiv s_z = \{T^{-1}\bar{p}(1-\bar{p})\}^{1/2}$	: estimated s.e. of $\bar{p}$;
$\bar{n}/C_k$	: should compare with 1;
$\overline{\eta} \equiv \bar{n} - C_k$	: should compare with $\hat{\eta}$;
p-value$_1$	: p-value for testing $E[N - C_k] \approx \eta$;
p-value$_2$	: p-value for testing $E[\overline{p} - (1-\alpha)] \approx \gamma C_k^{-1}$.

Table 2.2 summarizes the simulated performances for the FWCI estimation problem while implementing the purely sequential procedure $\mathcal{P}_{1,1}$ from (2.2.1). Tables 2.3-2.4 present similar summaries for the FWCI estimation problem while implementing our purely sequential procedures $\mathcal{P}_{1,2}, \mathcal{P}_{2,2}$ from (2.4.4) with alternative estimators $U_{2,n}^{(1)}$ and $U_{2,n}^{(2)}$ respectively as we gathered a pair of observations at-a-time (that is, we fixed $k = 2$).

In the spirits of Tables 2.3-2.4, we present similar summaries in Tables 2.5-2.6 for our purely sequential procedures $\mathcal{P}_{1,3}, \mathcal{P}_{2,3}$ from (2.4.4) as we gathered three observations at-a-time (that is, we fixed $k = 3$). Each table also provides the values of the corresponding theoretical second-order values $\hat{\eta}$ in the table's heading and γC_k^{-1} in Column 1.

Columns 3-5 suggest that the average sample sizes ($\overline{n}$) are close to the pre-assigned optimal fixed sample sizes had σ been known. We note that $s_{\overline{n}}$ values were very small throughout this exercise. The first-order efficiency term, $\overline{n}/C_k$, shows values that are very close to 1, with minor undersampling at times. The second-order efficiency term, $\overline{\eta}(=\overline{n}-C_k)$, stay reasonably close to the estimated theoretical value ($\hat{\eta}$).

In Section 2.4, $\hat{\eta}$ values were provided along with their estimated s.e. values, $s_{\hat{\eta}}$. We empirically tested:

$$H_0 : E[N - C_k] = \eta \text{ vs. } H_1 : E[N - C_k] \neq \eta, \tag{2.6.2}$$

formulated along a customary two-sample problem, utilizing the dataset on $(\overline{n} - C, s_{\overline{n}})$ and $(\hat{\eta}, s_{\hat{\eta}})$ based on two independent 10000 runs each. The associated p-values (p-value$_1$) are shown in Column 9. In Tables 2.2-2.6, we note that for moderate and large sample sizes, these p-values are larger than 0.05 which seem to indicate that our simulations validated the second-order efficiency property reasonably well. However, any discrepancy appears very small for all sample sizes, small, medium or large.

Column 7 shows the estimated coverage probabilities $\overline{p}$ which are remarkably close to $1-\alpha$ along with very small $s_{\overline{p}}$ values across all tables. In the spirit of a two-sample comparison highlighted by (2.6.2), we examined Column 8 for its practical validity of the estimated second-order term, $\overline{p} - (1 - \alpha)$, the difference between the estimated coverage probability $\overline{p}$

and the set target $1-\alpha$.

We went ahead and tested if these values reasonably agreed with the theoretical second-order term. The associated p-values (p-value$_2$) are shown in Column 10. In Tables 2.2-2.6, we note that for moderate and large sample sizes, these p-values are larger than 0.05 which seem to indicate that our simulations validated the second-order expansion of our coverage probability reasonably well. However, any discrepancy appears very small for all sample sizes, small, medium or large.

Table 2.2. Simulations for FWCI problem under $\mu = 5, \sigma = 2$ with 10000 runs implementing $N_{\mathcal{P}_{1,1},d}$ from (2.2.1): $k = 1$, $\alpha = 0.05$, $\hat{\eta} = -1.18$, and $P_{\boldsymbol{\theta}}\{\mu \in J_{N_{P_{1,1},d}}\} = 1 - \alpha + \gamma C_1^{-1} + o(C_1^{-1})$ from (2.2.26)

C_1 γC_1^{-1}	m_0	$\overline{n}$ $s_{\overline{n}}$	$\overline{n}/C_1$	$\bar{\eta} =$ $\overline{n} - C_1$	$\overline{x}$ $s_{\overline{x}}$	$\overline{p}$ $s_{\overline{p}}$	z	p-value$_1$	p-value$_2$
90 -0.0046	18	88.61 0.141	0.985	-1.38	5.000 0.0021	0.9448 0.0023	-0.0052	0.1501	0.7942
	48	88.59 0.141	0.984	-1.41	4.999 0.0022	0.9443 0.0023	-0.0057	0.1028	0.6325
	72	89.63 0.121	0.996	-0.37	4.998 0.0021	0.9505 0.0022	0.0005	0.2168	0.0204
240 -0.0017	18	238.90 0.221	0.995	-1.10	4.999 0.0013	0.9497 0.0022	-0.0003	0.7174	0.5245
	48	238.60 0.219	0.994	-1.40	4.999 0.0013	0.9478 0.0022	-0.0022	0.3151	0.8202
	72	238.73 0.221	0.995	-1.27	4.999 0.0013	0.9498 0.0022	-0.0002	0.6838	0.4954
600 -0.0007	18	598.58 0.350	0.998	-1.42	5.000 0.0008	0.9472 0.0022	-0.0028	0.4929	0.3398
	48	599.15 0.344	0.999	-0.85	5.000 0.0008	0.9474 0.0023	-0.0026	0.3374	0.4088
	72	599.10 0.346	0.999	-0.90	4.999 0.0008	0.9497 0.0022	-0.0003	0.4184	0.8557

Table 2.3. Simulations for FWCI problem under $\mu = 5, \sigma = 2$ with 10000 runs implementing $N_{\mathcal{P}_{1,2},d}$ from (2.4.4): $k = 2$, $\alpha = 0.05$, $\hat{\eta} = -1.18$, and $P_{\boldsymbol{\theta}}\{\mu \in J_{N_{P_{1,2},d}}\} = 1 - \alpha + \gamma C_2^{-1} + o(C_2^{-1})$ from Theorem 2.4.2

$2C_2$ γC_2^{-1}	$2m_0$	$\overline{n}$ $s_{\overline{n}}$	$\overline{n}/C_2$	$\bar{\eta} =$ $\overline{n} - C_2$	$\overline{x}$ $s_{\overline{x}}$	$\overline{p}$ $s_{\overline{p}}$	z	p-value$_1$	p-value$_2$
90	18	43.21	0.960	−1.79	5.000	0.9366	−0.0134	0.1191	0.0801
−0.0092		0.107			0.0022	0.0024			
	48	43.58	0.968	−1.42	4.999	0.9389	−0.0111	0.1433	0.4286
		0.098			0.0022	0.0024			
	72	45.14	1.003	0.14	4.999	0.9487	−0.0013	0.0000	0.0003
		0.079			0.0021	0.0022			
240	18	118.82	0.990	−1.18	4.999	0.9500	0.0000	1.0000	0.1222
−0.0034		0.157			0.0013	0.0022			
	48	118.66	0.989	−1.34	4.999	0.9508	0.0008	0.3112	0.0563
		0.158			0.0013	0.0022			
	72	118.74	0.990	−1.26	4.999	0.9478	−0.0022	0.6149	0.5854
		0.159			0.0013	0.0022			
600	18	298.73	0.996	−1.27	5.000	0.9478	−0.0022	0.7178	0.7161
−0.0014		0.249			0.0008	0.0022			
	48	299.15	0.997	−0.85	5.000	0.9497	−0.0003	0.1851	0.6171
		0.249			0.0008	0.0022			
	72	299.09	0.997	−0.91	4.999	0.9458	−0.0042	0.2724	0.2031
		0.246			0.0008	0.0022			

Table 2.4. Simulations for FWCI problem under $\mu = 5, \sigma = 2$ with 10000 runs implementing $N_{\mathcal{P}_{2,2},d}$ from (2.4.4): $k = 2$, $\alpha = 0.05$, $\hat{\eta} = -1.28$, and $P_{\boldsymbol{\theta}}\{\mu \in J_{N_{P_{2,2},d}}\} = 1 - \alpha + \gamma C_2^{-1} + o(C_2^{-1})$ from Theorem 2.4.2

$2C_2$ γC_2^{-1}	$2m_0$	$\overline{n}$ $s_{\overline{n}}$	$\overline{n}/C_2$	$\bar{\eta} =$ $\overline{n} - C_2$	$\overline{x}$ $s_{\overline{x}}$	$\overline{p}$ $s_{\overline{p}}$	z	p-value$_1$	p-value$_2$
90	18	43.15	0.959	−1.85	4.999	0.9318	−0.0182	0.0000	0.0074
−0.0115		0.114			0.0023	0.0025			
	48	43.59	0.969	−1.41	5.000	0.9384	−0.0116	0.2069.	0.9681
		0.103			0.0022	0.0025			
	72	45.24	1.005	0.24	4.999	0.9480	−0.0020	0.0000	0.0000
		0.083			0.0021	0.0022			
240	18	118.86	0.990	−1.14	4.999	0.9465	−0.0035	0.4074	0.7161
−0.0043		0.169			0.0013	0.0022			
	48	118.51	0.988	−1.49	4.999	0.9462	−0.0038	0.2167	0.8279
		0.170			0.0013	0.0023			
	72	118.68	0.989	−1.32	4.999	0.9455	−0.0045	0.8151	0.9307
		0.171			0.0013	0.0023			
600	18	298.85	0.996	−1.15	5.000	0.9504	0.0004	0.6263	0.3607
−0.0017		0.267			0.0008	0.0022			
	48	299.07	0.997	−0.93	5.000	0.9483	−0.0017	0.1899	1.0000
		0.267			0.0008	0.0022			
	72	298.95	0.997	−1.05	4.999	0.9511	0.0011	0.3818	0.2031
		0.263			0.0008	0.0022			

Table 2.5. Simulations for FWCI problem under $\mu = 5, \sigma = 2$ with 10000 runs implementing $N_{\mathcal{P}_{1,3},d}$ from (2.4.4): $k = 3$, $\alpha = 0.05$, $\hat{\eta} = -1.18$, and $P_{\boldsymbol{\theta}}\{\mu \in J_{N_{P_{1,3},d}}\} = 1 - \alpha + \gamma C_3^{-1} + o(C_3^{-1})$ from Theorem 2.4.2

$3C_3$ γC_3^{-1}	$3m_0$	$\overline{n}$ $s_{\overline{n}}$	$\overline{n}/C_3$	$\bar{\eta} =$ $\overline{n} - C_3$	$\overline{x}$ $s_{\overline{x}}$	$\overline{p}$ $s_{\overline{p}}$	z	p-value$_1$	p-value$_2$
90 -0.0137	18	27.10 0.099	0.903	-2.90	5.003 0.0026	0.9028 0.0030	-0.0472	0.0000	0.0000
	48	28.89 0.077	0.963	-1.11	4.999 0.0022	0.9356 0.0025	-0.0144	0.3633	0.7795
	72	30.29 0.060	1.010	0.29	4.998 0.0021	0.9496 0.0022	-0.0004	0.0000	0.0000
240 -0.0052	18	77.82 0.149	0.973	-2.18	4.999 0.0013	0.9407 0.0024	-0.0093	0.0000	0.0876
	48	78.47 0.134	0.980	-1.53	4.999 0.0013	0.9455 0.0023	-0.0045	0.0090	0.7609
	72	78.54 0.135	0.982	-1.46	4.999 0.0013	0.9421 0.0023	-0.0079	0.0381	0.2404
600 -0.0021	18	198.58 0.202	0.993	-1.42	5.000 0.0008	0.9488 0.0022	-0.0012	0.2347	0.6825
	48	198.58 0.202	0.993	-1.42	5.000 0.0008	0.9492 0.0022	-0.0008	0.2347	0.5546
	72	198.77 0.201	0.994	-1.23	4.999 0.0008	0.9506 0.0022	0.0006	0.8035	0.2197

Table 2.6. Simulations for FWCI problem under $\mu = 5, \sigma = 2$ with 10000 runs implementing $N_{\mathcal{P}_{2,3,d}}$ from (2.4.4): $k = 3$, $\alpha = 0.05$, $\hat{\eta} = -1.28$, and $P_{\boldsymbol{\theta}}\{\mu \in J_{N_{P_{2,3,d}}}\} = 1 - \alpha + \gamma C_3^{-1} + o(C_3^{-1})$ from Theorem 2.4.2

$3C_3$ γC_3^{-1}	$3m_0$	$\overline{n}$ $s_{\overline{n}}$	$\overline{n}/C_3$	$\bar{\eta} =$ $\overline{n} - C_3$	$\overline{x}$ $s_{\overline{x}}$	$\overline{p}$ $s_{\overline{p}}$	z	p-value$_1$	p-value$_2$
90	18	27.15	0.905	-2.85	5.001	0.9057	-0.0443	0.0000	0.0000
-0.0172		0.102			0.0026	0.0029			
	48	28.89	0.963	-1.11	4.998	0.9337	-0.0161	0.0358	0.6599
		0.081			0.0022	0.0025			
	72	30.43	1.014	0.43	4.998	0.9451	-0.0049	0.0000	0.0000
		0.063			0.0021	0.0023			
240	18	77.70	0.971	-2.30	4.999	0.9389	-0.0111	0.1831	0.0502
-0.0064		0.160			0.0014	0.0024			
	48	78.27	0.978	-1.73	4.999	0.9409	-0.0091	0.0018	0.2606
		0.144			0.0013	0.0024			
	72	78.50	0.981	-1.50	4.999	0.9407	-0.0093	0.1239	0.2269
		0.143			0.0013	0.0024			
600	18	198.21	0.991	-1.79	5.000	0.9461	-0.0039	0.0234	0.5719
-0.0026		0.225			0.0008	0.0023			
	48	198.37	0.992	-1.63	5.000	0.9494	-0.0006	0.1084	0.3409
		0.218			0.0008	0.0021			
	72	198.64	0.993	-1.36	4.999	0.9481	-0.0019	0.7085	0.7503
		0.214			0.0008	0.0022			

2.6.2. Simulation Study on MRPE Problem

In this section, we compare the performances of the purely sequential MRPE strategies defined respectively in (2.2.27) and (2.5.4) with $i = 1, 2$. We recall the stopping rules:

$$
\begin{aligned}
&(2.2.27)\ Q_{1,1}: \quad N_{Q_{1,1,c_1}} = \inf\left\{n \geq m_0(\geq 2): \quad n \geq (A/c_1)^{1/2} G_n\right\};\\
&(2.5.4)\ Q_{1,k}: \quad N_{Q_{1,k,c_k}} = \inf\left\{n \geq m_0(\geq 2): \quad n \geq \tfrac{1}{k}(A/c_k)^{1/2} T_{k,n}^{(1)}\right\};\\
&(2.5.4)\ Q_{2,k}: \quad N_{Q_{2,k,c_k}} = \inf\left\{n \geq m_0(\geq 2): \quad n \geq \tfrac{1}{k}(A/c_k)^{1/2} T_{k,n}^{(2)}\right\}.
\end{aligned}
\tag{2.6.3}
$$

Table 2.7. The set of notation used in Tables 2.8-2.12: $T = 10000$

n_i	: terminal sample size in i^{th} run;
$\bar{n} = T^{-1}\sum_{i=1}^{T} n_i$	: ave sample size, should compare with n_k^*;
$s_{\bar{n}} = \left\{(T^2 - T)^{-1}\sum_{i=1}^{T}(n_i - \bar{n})^2\right\}^{1/2}$	: estimated standard error (s.e.) of $\bar{n}$;
$\bar{x}_{n_i}$	: terminal sample mean in i^{th} run;
$\bar{x} = T^{-1}\sum_{i=1}^{T} \bar{x}_{n_i}$	: combined sample ave, should compare with μ;
$s_x = \left\{(T^2 - T)^{-1}\sum_{i=1}^{T}(\bar{x}_{n_i} - \bar{x})^2\right\}^{1/2}$	: estimated s.e. of $\overline{x}$;
$r_i = \frac{A\sigma^2}{kn_i} + kc_k n_i$	: estimated risk in i^{th} run;
$\overline{r} = T^{-1}\sum_{i=1}^{T} r_i$	: estimated risk, should compare with R_{k,n_k^*};
$\overline{\xi} = \overline{r}/R_{k,n_k^*}$	: should compare with 1 ;
$s_{\overline{\xi}} = \frac{1}{R_{k,n_k^*}}\left\{(T^2 - T)^{-1}\sum_{i=1}^{T}(r_i - \overline{r})^2\right\}^{1/2}$	: estimated s.e. of $\overline{\xi}$;
$\omega_i \equiv kc_k \frac{(n_i - n_k^*)^2}{n_i}$	: estimated regret in i^{th} run;
$\overline{\omega} = T^{-1}\sum_{i=1}^{T} \omega_i$	: estimated regret, should compare with $kc_k q$;
$s_{\overline{\omega}} = \{(T^2 - T)^{-1}\sum_{i=1}^{r}(\omega_i - \overline{\omega})^2\}^{1/2}$	: estimated s.e. of $\overline{\omega}$;
$\bar{n}/n_k^*$	: should compare with 1;
$\overline{\eta} \equiv \bar{n} - n_k^*$	: should compare with $\hat{\eta}$;
p-value$_1$	: p-value for testing $E[N - n_k^*] \approx \eta$;
p-value$_2$	: p-value for testing $\omega \approx kc_k q$

We intentionally pick optimal fixed sample sizes across tables so that we have: $c_1 > c_2 >$

c_3. Having fixed $A = 1$, pseudo random samples were generated from a $N(\mu = 5, \sigma^2 = 4)$ population. Recall that m_0 and n_k^* are respectively the pilot group number and the optimal fixed number of groups consisting of k observations each.

We fixed the following choices: $km_0 = 18, 36, 54$, $k = 1, 2, 3$, and kn_k^* values ranging broadly from approximately 30 through 1050. The MRPE methodologies from (2.6.3) were then implemented one by one with $T = 10000$ independent runs under each configuration. In Tables 2.8-2.12, we would use the set of notation defined precisely in Table 2.7 in addition to some that come directly from Table 2.1.

Table 2.8 summarizes the simulated performances for the MRPE problem while implementing the purely sequential procedure $Q_{1,1}$ from (2.2.27). Tables 2.9-2.10 present similar summaries for the MRPE problem while implementing our purely sequential procedures $Q_{1,2}, Q_{2,2}$ from (2.5.4) with alternative estimators $T_{2,n}^{(1)}$ and $T_{2,n}^{(2)}$ respectively as we gathered a pair of observations at-a-time (that is, we fixed $k = 2$).

In the spirits of Tables 2.9-2.10, we present similar summaries in Tables 2.11-2.12 for our purely sequential procedures $Q_{1,3}, Q_{2,3}$ from (2.5.4) as we gathered three observations at-a-time (that is, we fixed $k = 3$). Each table also provides the values of the corresponding theoretical second-order values $\hat{\eta}$ in the table's heading and the second-order term kc_kq_i in the regret expansion (Theorem 2.5.3) in Column 1.

Table 2.8. Simulations for MRPE problem under $\mu = 5, \sigma = 2, A = 1$ with 10000 runs implementing $N_{Q_{1,1,c_1}}$ from (2.2.27): $k = 1$, $\hat{\eta} = 0.13$, and $\omega = \frac{1}{2}c_1 + o(c_1)$ from Theorem 2.5.3

n_1^* $\frac{1}{2}c_1$	m_0	$\overline{n}$ $s_{\overline{n}}$	$\overline{n}/n_1^*$	$\bar{\eta} =$ $\overline{n} - n_1^*$	$\overline{x}$ $s_{\overline{x}}$	$\overline{\xi}$ $s_{\overline{\xi}}$	$\overline{\omega}$ $s_{\overline{\omega}}$	p-value$_1$ p-value$_2$
55	18	55.09	1.001	0.09	4.999	1.005	7.35×10^{-4}	0.4588
6.61×10^{-4}		0.054			0.0027	9.14×10^{-5}	1.33×10^{-5}	0.0000
	36	55.06	1.001	0.06	5.000	1.005	7.64×10^{-4}	0.2031
		0.055			0.0027	8.55×10^{-5}	1.24×10^{-5}	0.0000
	54	56.84	1.033	1.84	5.000	1.002	3.22×10^{-4}	0.0000
		0.034			0.0027	4.17×10^{-5}	6.08×10^{-6}	0.0000
500	18	500.46	1.000	0.46	5.000	1.001	8.06×10^{-6}	0.0391
8.00×10^{-6}		0.160			0.0009	7.22×10^{-6}	1.16×10^{-7}	0.6050
	36	500.02	1.000	0.02	5.000	1.001	8.20×10^{-6}	0.4918
		0.160			0.0009	7.42×10^{-6}	1.18×10^{-7}	0.0901
	54	500.03	1.000	0.03	5.001	1.001	8.03×10^{-6}	0.5268
		0.158			0.0009	7.19×10^{-6}	1.15×10^{-7}	0.7942
1000	18	1000.38	1.000	0.38	5.001	1.000	2.00×10^{-6}	0.2623
2.00×10^{-6}		0.223			0.0006	3.63×10^{-6}	2.90×10^{-8}	1.0000
	36	1000.22	1.000	0.22	5.000	1.000	2.03×10^{-6}	0.6892
		0.225			0.0006	3.40×10^{-6}	2.88×10^{-8}	0.2976
	54	999.67	1.000	−0.33	5.000	1.000	2.02×10^{-6}	0.0409
		0.225			0.0006	3.64×10^{-6}	2.91×10^{-8}	0.4919

Table 2.9. Simulations for MRPE problem under $\mu = 5, \sigma = 2, A = 1$ with 10000 runs implementing $N_{Q_{1,2,c_2}}$ from (2.5.4): $k = 2$, $\hat{\eta} = 0.13$, and $\omega = c_2 + o(c_2)$ from Theorem 2.5.3

$2n_2^*$ c_2	$2m_0$	$\overline{n}$ $s_{\overline{n}}$	$\overline{n}/n_2^*$	$\bar{\eta} =$ $\overline{n} - n_2^*$	$\overline{x}$ $s_{\overline{x}}$	$\overline{\xi}$ $s_{\overline{\xi}}$	$\overline{\omega}$ $s_{\overline{\omega}}$	p-value$_1$ p-value$_2$
60	18	30.06	1.002	0.06	4.999	1.011	1.43×10^{-3}	0.0878
1.11×10^{-3}		0.041			0.0026	2.82×10^{-4}	3.77×10^{-5}	0.0000
	36	30.05	1.002	0.05	5.001	1.010	1.35×10^{-5}	0.0510
		0.041			0.0026	1.76×10^{-4}	2.34×10^{-5}	0.0000
	54	30.72	1.024	0.72	5.000	1.005	7.12×10^{-4}	0.0000
		0.032			0.0026	6.46×10^{-5}	8.86×10^{-6}	0.0000
520	18	260.39	1.001	0.39	5.000	1.001	1.50×10^{-5}	0.0238
1.48×10^{-5}		0.115			0.0009	1.41×10^{-5}	2.16×10^{-7}	0.3545
	36	260.07	1.000	0.07	5.000	1.001	1.52×10^{-5}	0.6019
		0.115			0.0009	1.43×10^{-5}	2.20×10^{-7}	0.0690
	54	260.05	1.000	0.05	5.001	1.001	1.54×10^{-5}	0.4904
		0.116			0.0009	1.48×10^{-5}	2.28×10^{-7}	0.0085
1020	18	510.33	1.001	0.33	5.001	1.000	3.83×10^{-6}	0.2084
3.84×10^{-6}		0.159			0.0006	6.96×10^{-6}	5.46×10^{-8}	0.8547
	36	510.17	1.000	0.17	5.000	1.001	4.04×10^{-6}	0.8061
		0.163			0.0006	7.42×10^{-6}	5.82×10^{-8}	0.0006
	54	509.91	1.000	−0.09	5.000	1.000	3.88×10^{-6}	0.1691
		0.160			0.0006	6.97×10^{-6}	5.46×10^{-8}	0.4638

Table 2.10. Simulations for MRPE problem under $\mu = 5, \sigma = 2, A = 1$ with 10000 runs implementing $N_{Q_{2,2,c_2}}$ from (2.5.4): $k = 2$, $\hat{\eta} = 0.13$, and $\omega = (\pi - 2)c_2 + o(c_2)$ from Theorem 2.5.3

$2n_2^*$ $(\pi-2)c_2$	$2m_0$	$\overline{n}$ $s_{\overline{n}}$	$\overline{n}/n_2^*$	$\bar{\eta} =$ $\overline{n} - n_2^*$	$\overline{x}$ $s_{\overline{x}}$	$\overline{\xi}$ $s_{\overline{\xi}}$	$\overline{\omega}$ $s_{\overline{\omega}}$	p-value$_1$ p-value$_2$
60	18	30.10	1.003	0.10	4.999	1.012	1.57×10^{-3}	0.4854
1.27×10^{-3}		0.043			0.0026	2.93×10^{-4}	3.90×10^{-5}	0.0000
	36	30.06	1.002	0.06	5.002	1.011	1.51×10^{-3}	0.1035
		0.043			0.0027	1.88×10^{-4}	2.50×10^{-5}	0.0000
	54	30.76	1.025	0.76	5.000	1.006	7.68×10^{-4}	0.0000
		0.032			0.0026	7.40×10^{-5}	9.88×10^{-6}	0.0000
520	18	260.39	1.001	0.39	5.000	1.001	1.71×10^{-5}	0.0331
1.69×10^{-5}		0.122			0.0009	1.63×10^{-5}	2.51×10^{-7}	0.4255
	36	260.02	1.000	0.02	5.000	1.001	1.75×10^{-5}	0.3712
		0.123			0.0009	1.68×10^{-5}	2.59×10^{-7}	0.0205
	54	260.06	1.000	0.06	5.001	1.001	1.72×10^{-5}	0.5661
		0.122			0.0009	1.65×10^{-5}	2.55×10^{-7}	0.2394
1020	18	510.39	1.001	0.39	5.001	1.001	4.38×10^{-6}	0.1261
4.39×10^{-6}		0.170			0.0006	8.07×10^{-6}	6.33×10^{-8}	0.8745
	36	510.21	1.000	0.21	5.000	1.001	4.53×10^{-6}	0.6438
		0.173			0.0006	8.53×10^{-6}	6.69×10^{-8}	0.0250
	54	510.00	1.000	0.00	5.000	1.001	4.49×10^{-6}	0.4498
		0.172			0.0006	8.02×10^{-6}	6.28×10^{-8}	0.1113

Table 2.11. Simulations for MRPE problem under $\mu = 5, \sigma = 2, A = 1$ with 10000 runs implementing $N_{Q_{1,3,c_3}}$ from (2.5.4): $k = 3$, $\hat{\eta} = 0.13$, and $\omega = \frac{3}{2}c_3 + o(c_3)$ from Theorem 2.5.3

$3n_3^*$ $\frac{3}{2}c_3$	$3m_0$	$\overline{n}$ $s_{\overline{n}}$	$\overline{n}/n_3^*$	$\bar{\eta}$	$\overline{x}$ $s_{\overline{x}}$	$\overline{\xi}$ $s_{\overline{\xi}}$	$\overline{\omega}$ $s_{\overline{\omega}}$	p-value$_1$ p-value$_2$
75	18	25.06	1.002	0.06	4.999	1.014	1.48×10^{-3}	0.0655
1.07×10^{-3}		0.038			0.0023	4.02×10^{-4}	4.29×10^{-5}	0.0000
	36	25.00	1.000	0.00	5.000	1.013	1.40×10^{-3}	0.0006
		0.038			0.0024	2.68×10^{-4}	2.86×10^{-5}	0.0000
	54	25.17	1.001	0.17	5.001	1.010	1.11×10^{-3}	0.2531
		0.035			0.0023	1.36×10^{-4}	1.45×10^{-5}	0.0058
540	18	180.28	1.001	0.28	5.000	1.001	2.06×10^{-5}	0.1143
2.06×10^{-5}		0.095			0.0009	2.03×10^{-5}	3.00×10^{-7}	1.0000
	36	180.20	1.001	0.20	4.999	1.001	2.14×10^{-5}	0.4659
		0.096			0.0009	2.19×10^{-5}	3.24×10^{-7}	0.0135
	54	180.01	1.000	0.01	5.000	1.001	2.08×10^{-5}	0.2065
		0.095			0.0009	2.06×10^{-5}	3.05×10^{-7}	0.5119
1050	18	350.35	1.001	0.35	5.001	1.001	5.49×10^{-6}	0.0956
5.44×10^{-6}		0.132			0.0006	1.04×10^{-5}	7.93×10^{-8}	0.5284
	36	350.38	1.001	0.38	4.999	1.001	5.44×10^{-6}	0.0563
		0.131			0.0006	1.06×10^{-5}	8.07×10^{-8}	1.0000
	54	349.98	1.000	-0.02	5.000	1.001	5.40×10^{-6}	0.2522
		0.131			0.0006	1.01×10^{-5}	7.68×10^{-8}	0.6025

Table 2.12. Simulations for MRPE problem under $\mu = 5, \sigma = 2, A = 1$ with 10000 runs implementing $N_{Q_{2,3,c_3}}$ from (2.5.4): $k = 3$, $\hat{\eta} = 0.13$, and $\omega = 3(\frac{1}{2}\pi - 1)c_3 + o(c_3)$

$3n_3^*$ $3(\frac{1}{2}\pi - 1)c_3$	$3m_0$	$\overline{n}$ $s_{\overline{n}}$	$\overline{n}/n_3^*$	$\bar{\eta}$	$\overline{x}$ $s_{\overline{x}}$	$\overline{\xi}$ $s_{\overline{\xi}}$	$\overline{\omega}$ $s_{\overline{\omega}}$	p-value$_1$ p-value$_2$
75	18	25.07	1.003	0.07	4.999	1.016	1.71×10^{-3}	0.1434
1.21×10^{-3}		0.041			0.0023	4.78×10^{-4}	5.10×10^{-5}	0.0000
	36	24.97	0.999	-0.03	5.001	1.014	1.57×10^{-3}	0.0001
		0.040			0.0024	2.98×10^{-4}	3.18×10^{-5}	0.0000
	54	25.17	1.007	0.17	5.001	1.012	1.26×10^{-3}	0.2925
		0.038			0.0023	1.48×10^{-4}	1.58×10^{-5}	0.0016
540	18	180.31	1.002	0.31	5.000	1.002	2.38×10^{-5}	0.0747
2.35×10^{-5}		0.101			0.0009	2.40×10^{-5}	3.56×10^{-7}	0.3994
	36	180.21	1.001	0.21	4.999	1.002	2.36×10^{-5}	0.4283
		0.101				2.37×10^{-5}	3.51×10^{-7}	0.7757
	54	179.99	1.000	-0.01	5.000	1.001	2.42×10^{-5}	0.1699
		0.102				1.20×10^{-5}	3.51×10^{-7}	0.0461
1050	18	350.31	1.001	0.31	5.001	1.001	6.32×10^{-6}	0.2049
6.21×10^{-6}		0.142			0.0006	1.20×10^{-5}	9.11×10^{-8}	0.2273
	36	350.35	1.001	0.35	5.000	1.001	6.24×10^{-6}	0.1187
		0.141			0.0006	1.17×10^{-5}	8.95×10^{-8}	0.7375
	54	350.00	1.000	0.00	5.000	1.001	6.13×10^{-6}	0.3497
		0.139			0.0006	1.18×10^{-5}	9.02×10^{-8}	0.3751

In view of our detailed comments presented in the contexts of Tables 2.2-2.6, we refrain from adding elaborate commentaries on Tables 2.8-2.12. We may simply add that all first-order and second-order theoretical results are clearly seemed to be validated by our simulated exercises as we had implemented the MRPE methodologies $Q_{1,1}$ from (2.2.27), $Q_{1,k}$ from (2.5.4) with $k = 2, 3$, and $Q_{2,k}$ from (2.5.4) with $k = 2, 3$.

2.7. Illustration With Breast Cancer Data

Now, we move to illustrate our purely sequential sampling strategies for both FWCI (Section 2.4) and MRPE (Section 2.5) problems with the help of real breast cancer dataset. We provide following two links for the repository as well as the real dataset:

Dataset Link: https://archive.ics.uci.edu/ml/datasets/Breast+Cancer
+Wisconsin+%28Diagnostic%29

UCI Machine Learning Repository: http://archive.ics.uci.edu/ml

Additionally, one may refer to Dua and Graff (2019). No special permission is required to access and/or utilize the dataset.

Briefly, features are computed from a digitized image of a *fine needle aspirate* (FNA) from a breast mass. There were 357 patients who were diagnosed with a "benign" mass. The average size of the core tumor from 357 patients were collected and analyzed. Shapiro-Wilk test of normality on original was not violated with p-value 0.7795. The mean and standard deviation of the full dataset came out to be 78.08 and 11.8074, respectively.

We do not incorporate these two estimates in order to run our estimation strategies drawing observations randomly and independently from this dataset pretending to have this full body of data available to us as our intended population under purely sequential sampling. The population mean and variance are treated as unknown. Tables 2.13-2.14 respectively correspond to the FWCI and MRPE problems.

2.7.1. Illustration for the FWCI Problem

In Table 2.13, we use the following labels for identifying a specific FWCI methodology under implementation:Table

$$\begin{aligned} \delta_1 = 0 \quad &: S_n^2\text{-based purely sequential methodology from (2.2.1);} \\ \delta_1 = 1 \quad &: U_{2,n}^{(1)}\text{-based group sampling methodology from (2.4.4);} \\ \delta_1 = 2 \quad &: U_{2,n}^{(2)}\text{-based group sampling methodology from (2.4.4).} \end{aligned} \tag{2.7.1}$$

While the width of a confidence interval should be customized by a team of health professionals, we pretend fixing three choices of the d-values successively going down, namely $d = 3.27, 2.31$ and 1.89. Then, we implemented the FWCI methodology ($\delta_1 = 0$) incorporating one observation at-a-time, followed by group sampling methodology ($\delta_1 = 1, 2$) by recording 2 observations at-a-time. We fixed $km_0 = 12, 20, 30$ and then checked incoming sequential data under the requisite stopping rules. Final sample sizes along with the lower bound and upper bound are reported.

Table 2.13. FWCI estimation of population mean μ

km_0	d	Index δ_1 (2.7.1)	kn	$\widehat{\mu} \equiv \overline{x}_n$	95% Conf Interval lower CL $\widehat{\mu} - d$	upper CL $\widehat{\mu} + d$
12	3.27	0	55	78.16	74.89	81.43
		1	68	78.65	75.38	81.92
		2	48	77.04	73.77	80.31
	2.31	0	99	78.40	76.09	80.71
		1	102	78.14	75.83	80.45
		2	118	76.61	74.30	78.92
	1.89	0	145	78.76	76.87	80.65
		1	136	78.19	76.30	80.08
		2	142	78.12	76.23	80.01
20	3.27	0	49	78.65	75.38	81.92
		1	46	78.16	74.89	81.43
		2	48	78.65	75.38	81.92
	2.31	0	87	78.17	75.86	80.48
		1	82	78.08	75.77	80.39
		2	98	78.73	76.42	81.04
	1.89	0	146	77.48	75.59	79.37
		1	150	77.69	75.80	79.58
		2	156	77.57	75.68	79.46
30	3.27	0	43	78.52	75.25	81.79
		1	42	78.26	74.99	81.53
		2	44	78.04	74.77	81.31
	2.31	0	94	77.21	74.90	79.52
		1	116	77.55	75.24	79.86
		2	102	76.22	73.91	78.53
	1.89	0	141	77.59	75.70	79.46
		1	164	77.97	76.08	79.86
		2	138	77.16	75.27	79.05

Table 2.13 shows clearly that to get a more precise interval (smaller d), we need more observations. Also, the confidence intervals generated by group sampling procedures ($\delta_1 = 1, 2$) overlap with those generated by one observation at-a-time ($\delta_1 = 0$) purely sequential methodology. In few cases, the differences among observed values of kn (that is, when $N = n$) may look relatively large, but one should keep in mind that each row corresponds to a **single**

run only. Columns 6-7 respectively show the lower CL and the upper CL associated with the 95% confidence interval $[\overline{x}_n \pm d]$ obtained upon termination reported in each row. We find that these such confidence intervals are reliable in the sense that each interval $[\overline{x}_n \pm d]$ so obtained happened to include the mean (= 78.08) from the full dataset.

2.7.2. Illustration for the MRPE Problem

In Table 14, we use the following labels for identifying a specific MRPE methodology under implementation:

$$\begin{aligned} \delta_2 = 0 &\quad : G_n \text{ based purely sequential methodology from (2.27);} \\ \delta_2 = 1 &\quad : T_{2,n}^{(1)} \text{ based group sampling methodology from (5.4);} \\ \delta_2 = 2 &\quad : T_{2,n}^{(2)} \text{ based group sampling methodology from (5.4).} \end{aligned} \tag{2.7.2}$$

We implemented the MRPE methodologies with $A = 1$ by varying the value of c_k, the cost of sampling per unit observation when recording k-tuples at-a-time.

Table 2.14. MRPE strategy for population mean μ

Index		$km_0 = 12$			$km_0 = 20$			$km_0 = 30$		
δ_2			$\widehat{\mu} \equiv$			$\widehat{\mu} \equiv$			$\widehat{\mu} \equiv$	
(2.7.2)	c_k	kn	$\overline{x}_n$	s_x	kn	$\overline{x}_n$	s_x	kn	$\overline{x}_n$	s_x
0	0.0558	53	78.45	1.70	50	78.59	1.63	46	78.26	1.59
1	0.0516	44	78.11	1.85	52	74.48	1.58	56	74.28	1.87
2	0.0516	50	77.48	1.48	50	78.45	1.69	50	78.28	1.50
0	0.0139	100	78.35	1.17	94	78.66	1.14	95	78.85	1.13
1	0.0134	92	77.45	1.19	102	75.22	1.13	102	76.22	1.24
2	0.0134	102	77.80	1.17	100	78.67	1.18	92	77.57	1.05
0	0.0062	149	78.84	0.95	138	78.98	0.92	143	79.16	0.94
1	0.0060	126	78.21	0.99	146	76.38	0.94	146	77.28	1.00
2	0.0060	136	78.82	0.98	152	78.68	0.94	152	78.56	0.88

Having fixed choices for $km_0 = 12, 20$, or 30, we incorporated the three methodologies ($\delta_2 = 0, 1, 2$) as indexed via (2.7.2). The final terminal sample size(n) and $\widehat{\mu}(= \overline{x}_n)$ along with its standard standard deviation(s_x) are reported in Table 14. In this table, we observe no sizable differences between estimates of μ across the board whereas those estimates looked rather accurate in the sense that they happened to be very close to the mean ($= 78.08$) from the full dataset.

2.8. Brief Conclusion

We revisited the FWCI problem and MRPE problem both on the context of estimating unknown population mean μ when the population variance σ^2 is also unknown. In stead of sampling one individual sequentially, we proposed a newly group sampling methodology with observations gathered in groups. It provided a innovative idea when observations come in batches, or when sampling in groups is much more efficient and economic than sampling in individuals. We do so because in real life we know that packaged items purchased in bulk often cost less per unit sample than the cost of an individual item. This paper builds the whole array of estimation methodologies in order to address both FWCI and MRPE problems with appropriate first-order and second-order asymptotic analyses. These are followed by extensive sets of carefully laid out data analyses assisted via large-scale computer simulations. These are wrapped up with illustrations using breast cancer data.

Chapter 3

Purely Sequential Estimation Problems for the Mean of a Normal Population by Sampling in Groups Under Permutations within Each Group and Illustrations

3.1. Introduction and Layout

In this chapter, we briefly revisit two fundamental problems on sequential estimation: (i) the *fixed-width confidence interval* (FWCI) estimation problem, and (ii) the *minimum risk point estimation* (MRPE) problem. We do so in the context of estimating the unknown mean μ in a $N(\mu, \sigma^2)$ population where σ is also assumed unknown. We will soon explain how we go beyond the recent paper of Mukhopadhyay and Wang (2020) to further discuss these problems under the framework of sampling in groups of $k(\geq 1)$ observations gathered at-a-time.

Sampling a group of observations may be understood differently from sampling vectors of observations. The latter customarily corresponds to observations, measured as vectors, but inside a vector they may or may not be *independent and identically distributed* (i.i.d.). For example, individual's height and weight may be measured at the same time. Height and weight may be treated as a two-dimensional vector. In this case, height and weight are correlated variables, having different distributions, and one may observe such vectors of observations.

In such contexts, Mukhopadhyay and Al-Mousawi (1986) proposed two-stage, modified two-stage, purely sequential, and three-stage methodologies to construct "fixed-size" elliptic confidence regions for estimating the mean vector of a p-dimensional normal distribution. One will find a rich literature from Mukhopadhyay and de Silva (2009, chapters 11-12) and other sources.

However, we view sampling a group of independent observations at-a-time as a different matter altogether and that is exactly what we will be incorporating where the observations are assumed to be i.i.d. We consider sampling $k(\geq 1)$ observations at-a-time, recorded in groups:

$$(X_{i1}, ..., X_{ik}), \ i = 1, 2, ..., n, ...$$

from a $N(\mu, \sigma^2)$ population. We denote $\boldsymbol{\theta} = (\mu, \sigma^2)$ and assume that the joint parameter space is given by $\mathfrak{R} \times \mathfrak{R}^+$.

Our motivation for considering sampling k observations in groups at-a-time is simple and practical. We assume that the observations may keep arriving in groups of size k at-a-time. So, there may not be a possibility or practicality of gathering one observation at-a-time. More often than not when we buy goods in packets or boxes, such a packet/box will customarily cost less per unit sample. For example, a single AA battery may cost \$0.95, but a 12-pack may cost \$9.95.

Sampling in batches of equal size has been successfully explored in a recent paper of Malinovsky and Zacks (2018) while handling "proportional closeness" estimation of probability of contamination under a group testing methodology. The recent work of Mukhopadhyay and Wang (2020), as well as this present investigation provide other practical flavors different from what Malinovsky and Zacks (2018) had considered.

3.1.1. A Layout of this Chapter

Section 3.2 outlines the basic purely sequential FWCI (Section 3.2.1) and MRPE (Section 3.2.2) strategies (3.2.5) and (3.2.16) respectively defined by incorporating sampling one observation at-a-time. Then, we briefly summarize purely sequential estimation strategies (Mukhopadhyay and Wang 2020) for both FWCI and MRPE problems via implementation of (3.2.6) and (3.2.17) respectively by sampling k-tuples at-a-time. Theorems 3.2.1-3.2.2 add foundations for important asymptotic second-order properties. Those FWCI (or MRPE) strategies incorporated the unbiased and consistent estimators $U_{k,n}^{(1)}, U_{k,n}^{(2)}$ (or $T_{k,n}^{(1)}, T_{k,n}^{(2)}$) of σ^2 (or σ) from (3.2.7) (or from (3.2.19)) in defining boundary crossing.

We introduce new unbiased and consistent estimators $U_{k,n}^{\#(1)}, U_{k,n}^{\#(2)}$ (or $T_{k,n}^{\#(1)}, T_{k,n}^{\#(2)}$) of σ^2 (or σ) in (3.3.1) and (3.3.6) (or in (3.3.3) and (3.3.5)) by permuting observations within k-tuples with $(1, k-1)$-splits in order to come up with newer estimation methodologies, namely (3.3.13) and (3.3.14) respectively. These stopping rules hang in tighter around the optimal fixed-sample-sizes than those that came from Mukhopadhyay and Wang (2020).

Section 3.4 develops both asymptotic first-order and second-order properties of the estimation strategies from (3.3.13) in the context of the FWCI problem. All ensuing second-order terms and the rates of convergences have been treated rigorously with detailed data analysis and data-validation which follow with summaries obtained from large-scale simulations. Section 5 handles the MRPE problem and it is built much in the same spirits of Section 3.4.

The paper ends with Section 3.6 which includes illustrations of our purely sequential estimation strategies (3.3.13) and (3.3.14) with the help of wind energy data that are publicly available from the website of the National Renewable Energy Laboratory.

3.2. Existing Purely Sequential FWCI and MRPE Methodologies

In this section, we begin by outlining some of the basic purely sequential FWCI (Section 3.2.1) and MRPE (Section 3.2.2) methodologies, and their associated key asymptotic second-order characteristics.

3.2.1. FWCI Estimation Problem

Having fixed two pre-assigned values, $d(>0)$ and $0<\alpha<1$, the problem is one of estimating μ with a confidence interval J such that its length is $2d$, and the confidence coefficient is at least (or approximately) $1-\alpha$. Having recorded $(X_{i1},...,X_{ik})$, $i=1,2,...,n,\ k\geq 2$, we denote the sample mean and the associated confidence interval for μ as follows:

$$\overline{X}_{k,n} \equiv (nk)^{-1}\Sigma_{i=1}^{n}\Sigma_{j=1}^{k}X_{ij} \text{ and } J_{k,n}=[\overline{X}_{k,n}-d,\overline{X}_{k,n}+d]. \quad (3.2.1)$$

We hold k fixed, but otherwise leave it rather arbitrary. We begin with

$$(X_{11},...,X_{1k}),...,(X_{m_0 1},...,X_{m_0 k}),$$

a pilot data of k-tuples of size m_0 (>1), and then record additional k-tuples of the X's at-a-time as needed.

The confidence coefficient associated with $J_{k,n}$ from (3.2.1) is expressed as:

$$P_{\boldsymbol{\theta}}\{\mu\in J_{k,n}\}=P_{\boldsymbol{\theta}}\left\{|\overline{X}_{k,n}-\mu|\leq d\right\}=2\Phi\left(\frac{\sqrt{knd}}{\sigma}\right)-1. \quad (3.2.2)$$

Observe that $J_{k,n}$ already has the required fixed-width, $2d$. But, we also require that the associated confidence coefficient must be nearly (at least) $1-\alpha$.

Thus, we must have:

$$2\Phi\left(\sqrt{kn}d/\sigma\right)-1\geq 1-\alpha=2\Phi\left(a\right)-1, \tag{3.2.3}$$

where $a\equiv a_{\alpha}$ is the upper $50\alpha\%$ point of a $N(0,1)$ distribution. From (3.2.3), we then claim that the required number of k-tuples

$$\text{must be the smallest } n\geq a^{2}\sigma^{2}/(d^{2}k)\equiv C_{k}, \text{ say.} \tag{3.2.4}$$

This C_k is referred to as the required optimal fixed number of k-tuples had σ^2 been known. The magnitude of C_k, however, remains unknown. Indeed no fixed-sample-size methodology would come up with a solution for this problem regardless of whether or not the confidence interval is centered at $\overline{X}_{k,n}$. Dantzig (1940) proved this fundamental non-existential result. One may also refer to Ghosh et al. (1997, Chapter 3), Mukhopadhyay (1988,2000, Chapter 13), and Mukhopadhyay and de Silva (2009, Chapter 2) and other sources such as Sen (1981), Woodroofe (1982), Siegmund (1985), Mukhopadhyay and Solanky (1994) and Zacks (2009,2017).

Anscombe (1952,1953), Ray (1957) and Chow and Robbins (1965) originally defined the ground-breaking purely sequential methodology when $k=1$:

$$\text{Methodology } \mathcal{P}_1\text{:}\quad N_{\mathcal{P}_1,d}\equiv N_{\mathcal{P}_1}=\inf\left\{n\geq m_0(\geq 2):\quad n\geq a^2S_n^2/d^2\right\}. \tag{3.2.5}$$

where $S_n^2=(n-1)^{-1}\Sigma_{i=1}^{n}\left(X_i-\overline{X}_n\right)^2$, the sample variance.

When $k \geq 2$, Mukhopadhyay and Wang (2020) provided the following purely sequential methodology with alternative estimators in the spirit of (3.2.5):

$$\text{Methodology } \mathcal{P}_{i,k} : N_{\mathcal{P}_{i,k},d} \equiv N_{\mathcal{P}_{i,k}} = \inf \left\{ n \geq m_0(\geq 2) : \quad n \geq a^2 U_{k,n}^{(i)}/(kd^2) \right\}, \; i = 1, 2, \tag{3.2.6}$$

where they defined:

$$\begin{aligned} U_{k,n}^{(1)} &\equiv \tfrac{1}{n}\Sigma_{i=1}^{n}\tfrac{(k-1)}{k}\left\{X_{i1} - \tfrac{1}{k-1}\Sigma_{j=2}^{k}X_{ij}\right\}^2, \text{ and} \\ U_{k,n}^{(2)} &\equiv \tfrac{(k-1)\pi}{nk(2n-2+\pi)}\left\{\Sigma_{i=1}^{n}|X_{i1} - \tfrac{1}{k-1}\Sigma_{j=2}^{k}X_{ij}|\right\}^2, \end{aligned} \tag{3.2.7}$$

the unbiased and consistent estimators of σ^2.

The sampling strategies $\mathcal{P}_1$ from (3.2.5) and $\mathcal{P}_{i,k}$ from (3.2.6) terminate w.p.1, that is, $P_{\boldsymbol{\theta}}\{N_{\mathcal{P}_1} < \infty\} = 1$ and $P_{\boldsymbol{\theta}}\{N_{\mathcal{P}_{i,k}} < \infty\} = 1$. As $d \to 0$, a number of asymptotic first-order properties associated with the estimation strategies $\left(N_{\mathcal{P}_{i,k}}, J_{N_{\mathcal{P}_{i,k}}}\right)$ can be summarized:

$$\begin{gathered} N_{\mathcal{P}_{i,k}}/C_k \overset{P_{\boldsymbol{\theta}}}{\to} 1;\; E_{\boldsymbol{\theta}}[N_{\mathcal{P}_{i,k}}/C_k] \to 1 \;[\textit{Asymptotic first-order efficiency}]; \\ \text{and } P_{\boldsymbol{\theta}}\{\mu \in J_{N_{\mathcal{P}_{i,k}}}\} \to 1 - \alpha \;[\textit{Asymptotic consistency}]; \end{gathered} \tag{3.2.8}$$

$i = 1, 2$.

These properties will traditionally follow from Chow and Robbins (1965). One may also supply a direct proof of the first part from what is called the basic inequality. The second part may combine the first part, Fatou's Lemma, along with the result: $E_{\boldsymbol{\theta}}[N_{\mathcal{P}_{i,k}}] \leq C_k + O(1)$. The third part will follow from a combination of the first part and the dominated convergence theorem. One may refer to Ghosh and Mukhopadhyay (1976,1981) and Ghosh et al. (1997, Theorems 7.2.1 and 8.2.1).

Furthermore, we may let $H_{i,k}^{(1)} = C_k^{-1/2}(N_{\mathcal{P}_{i,k},d} - C_k)$, and apply the results from nonlinear renewal theory to claim:

$$
\begin{array}{ll}
(i) & \left.\begin{array}{l}(a)\ P_{\boldsymbol{\theta}}\{N_{\mathcal{P}_{1,k},d} \leq \varepsilon C_k\} = O(C_k^{-m_0/2}) \\ (b)\ P_{\boldsymbol{\theta}}\{N_{\mathcal{P}_{2,k},d} \leq \varepsilon C_k\} = O(C_k^{-m_0})\end{array}\right\} \textit{with fixed } 0 < \varepsilon < 1 \textit{ and } m_0 \geq 2; \\
(ii) & |H_{i,k}^{(1)}|^{\kappa} \textit{ is uniformly integrable if } m_0 > \max\{1, \kappa\} \textit{ with fixed } \kappa(>0),\ i = 1, 2; \\
(iii) & \left.\begin{array}{l}(a)\ H_{1,k}^{(1)} \xrightarrow{\pounds} N(0, 2) \\ (b)\ H_{2,k}^{(1)} \xrightarrow{\pounds} N(0, 2\pi - 4)\end{array}\right\} \textit{if } m_0 \geq 2; \\
(iv) & \left.\begin{array}{l}(a)\ E_{\boldsymbol{\theta}}\left[N_{\mathcal{P}_{1,k},d} - C_k\right] = \eta_{1,k}^{(1)} + o(1) \\ (b)\ E_{\boldsymbol{\theta}}\left[N_{\mathcal{P}_{2,k},d} - C_k\right] = \eta_{2,k}^{(1)} + o(1)\end{array}\right\} \textit{if } m_0 \geq 3;
\end{array} \tag{3.2.9}
$$

with $\eta_{1,k}^{(1)} \approx -1.1839$, and $\eta_{2,k}^{(1)} \approx -1.2788$ respectively. The superscript (1) used here indicates the FWCI estimation strategies.

3.2.1.1 A General Asymptotic Second-Order Expansion of the Confidence Coefficient

For completeness, we summarize this material from Mukhopadhyay and Wang (2020) and begin with a general multi-stage sampling strategy leading to a suitably defined arbitrary stopping time, $M \equiv M_{k,d}$. We are, however, assured that $M_{k,d}$ is "close" to C_k. Let us now list a number of *assumptions* that are needed in order to draw a range of specific conclusion(s):

(A1): For every fixed $n \geq m_0(> 1)$, the pilot size, the event $[M_{k,d} = n]$ depends only on the statistic $\mathbf{Z}_{k,n} \equiv \{X_{i1} - X_{ik}, ..., X_{ik-1} - X_{ik}\}, i = 1, ..., n$;

(A2): $M_{k,d}/C_k \overset{P_{\boldsymbol{\theta}}}{\to} 1$ as $d \to 0$, and also $P_{\boldsymbol{\theta}}\{M_{k,d} \leq \varepsilon C_k\} = O(C_k^{-a_1 m_0 + a_2})$ if $0 < \varepsilon < 1$, as $d \to 0$, with some $a_1 > 0$, if $m_0 > a_2/a_1$;

(A3): $H \equiv C_k^{-1/2}(M_{k,d} - C_k) \overset{\pounds}{\to} N(0, a_3)$ as $d \to 0$, with some $a_3 > 0$;

(A4): $|H|^2$ is uniformly integrable if $m_0 > a_4(> 0)$; and

(A5): $E_{\boldsymbol{\theta}}\left[M_{k,d} - C_k\right] = a_5 + o(1)$ as $d \to 0$, if $m_0 > a_6(> 0)$;

(3.2.10)

for every fixed μ, σ, and α.

With the finally accrued data $\{M_{k,d}, (X_{i1}, ..., X_{ik}), i = 1, 2, ..., M_{k,d}\}$, we propose the associated FWCI:

$$J_{M_{k,d}} = [\overline{X}_{M_{k,d}} - d, \overline{X}_{M_{k,d}} + d] \text{ for } \mu. \tag{3.2.11}$$

Theorem 3.2.1 (Mukhopadhyay and Wang, 2020)**.** *Under a general asymptotically second-order efficient multi-stage fixed-width confidence interval methodology* $(M_{k,d}, J_{M_{k,d}})$, *under the standing assumptions (A1)-(A5) from (3.2.10), for every fixed* k, μ, σ, *and* α, *we have the following asymptotic second-order expansion of the confidence coefficient associated with* $J_{M_{k,d}}$ *from (3.2.11) as* $d \to 0$:

$$P_{\boldsymbol{\theta}}\{\mu \in J_{M_{k,d}}\} = (1 - \alpha) + a\left\{a_5 - \tfrac{1}{4}a_3(1 + a^2)\right\}\phi(a)C_k^{-1} + o(d^{-2}), \text{ if the pilot size, } m_0 > \max\{1, a_6, a_4, \tfrac{1}{2a_1}(5 + 2a_2)\},$$

with C_k coming from (3.2.4) and $a \equiv a_\alpha$ is the upper $50\alpha\%$ point of $N(0,1)$.

Proof: Under the assumption (A1), we invoke Basu's (1955) theorem in the spirit of Mukhopadhyay (2000, Example 6.6.15) to claim that the random variable $I(M_{k,d} = n)$ and the sample mean, $\overline{X}_{k,n}$, are independent for every fixed $n \geq m_0$. Thus, we can express:

$$\begin{aligned} P_{\boldsymbol{\theta}}\{\mu \in J_{M_{k,d}}\} &= E_{\boldsymbol{\theta}}\left[2\Phi\left(\sqrt{kM_{k,d}}d/\sigma\right) - 1\right] = E_{\boldsymbol{\theta}}\left[2\Phi\left(aM_{k,d}^{1/2}C_k^{-1/2}\right) - 1\right] \\ &= E_{\boldsymbol{\theta}}\left[\psi\left(M_{k,d}/C_k\right)\right] \text{ where } \psi(x) \equiv 2\Phi\left(ax^{1/2}\right) - 1,\ x > 0. \end{aligned} \tag{3.2.12}$$

Refer to Mukhopadhyay and Wang (2020, Theorem 2.1) for other crucial details. ■

For the stopping rule defined in (3.2.6), we have following second order property:

$$P_{\boldsymbol{\theta}}\{\mu \in J_{M_{k,d}}\} = (1-\alpha) + \gamma_{i,k}C_k^{-1} + o(d^2) \begin{cases} \text{if } m_0 \geq 7 & \text{when } k = 1 \\ \text{if } m_0 \geq 6 & \text{when } k \geq 2 \end{cases}, \tag{3.2.13}$$

with $\gamma_{1,k} \approx -0.4125$ and $\gamma_{2,k} \approx -0.4630$ respectively.

3.2.2. MRPE Problem

We recall sampling $k(\geq 2)$ observations at-a-time, recorded in groups of independent $N(\mu, \sigma^2)$ observations. Having recorded

$$(X_{i1}, ..., X_{ik}), i = 1, 2, ..., n,$$

we propose to estimate μ with $\overline{X}_{k,n}$ under a loss function which is composed of squared error loss plus linear cost due to sampling as follows:

$$L_{k,n}(\mu, \overline{X}_{k,n}) = A(\overline{X}_{k,n} - \mu)^2 + c_k kn, \tag{3.2.14}$$

where $c_k(>0)$ is the known cost per unit observation and $A(>0)$ is also assumed known.

We assume that $0 < c_1 < ... < c_{k-1} < c_k$ which makes sampling in groups of size k cheaper and often more practical. The associated fixed-sample-size risk function can be expressed as:

$$R_{k,n} = E_{\mu,\sigma}[L_{k,n}(\mu, \overline{X}_{k,n})] = A\sigma^2(kn)^{-1} + kc_k n, \tag{3.2.15}$$

Our goal is to construct the MRPE of μ which leads to the following optimal fixed number of groups, n_k^*, had σ been known:

$$\begin{gathered} n \text{ is the minimum number of groups} \geq \frac{1}{k}\sqrt{A/c_k}\sigma = n_k^*, \text{ say, giving rise to the} \\ \text{fixed-sample-size minimum risk } R_{k,n_k^*} \equiv R_{n_k^*}(c_k) = 2kc_k n_k^*. \end{gathered} \tag{3.2.16}$$

However, the magnitude of n_k^* remains unknown since σ^2 is unknown. Indeed no fixed-sample-size methodology would come up with a solution for this problem regardless of whether or not the point estimator involves the sample mean or another estimator of choice. Again, Dantzig (1940) proved this fundamental result.

One may also refer to Ghosh et al. (1997, Chapter 3), Mukhopadhyay (2000, Chapter 13), and Mukhopadhyay and de Silva (2009, Chapter 2) and other sources such as Sen (1981), Woodroofe (1982), and Mukhopadhyay and Solanky (1994). Robbins (1959) defined the following purely sequential procedure when $k=1$:

$$\text{Methodology } Q_1\text{:}\ \ N_{Q_1,c_1} \equiv N_{Q_1} = \inf\left\{n \geq m_0(\geq 2):\ \ n \geq (A/c_1)^{1/2} S_n\right\}, \tag{3.2.17}$$

where S_n is the sample standard deviation.

Much in the spirits of (3.2.6), Mukhopadhyay and Wang (2020) updated the estimation strategy (3.2.17) and gave the following estimation procedure when $k \geq 2$:

$$\text{Methodology } Q_{i,k}:\ N_{Q_{i,k},c_k} \equiv N_{Q_{i,k}} = \inf\left\{n \geq m_0(\geq 2):\ \ n \geq \tfrac{1}{k}\sqrt{A/c_k}\, T_{k,n}^{(i)}, i=1,2\right\}, \tag{3.2.18}$$

where

$$\begin{aligned} T_{k,n}^{(1)} &\equiv \Gamma(\tfrac{n}{2})\left\{\Gamma(\tfrac{n+1}{2})\right\}^{-1}\left\{\tfrac{k-1}{2k}\Sigma_{i=1}^{n}\left(X_{i1}-\tfrac{1}{k-1}\Sigma_{j=2}^{k}X_{ij}\right)^2\right\}^{1/2}, \text{ and} \\ T_{k,n}^{(2)} &\equiv \tfrac{1}{n}\Sigma_{i=1}^{n}\sqrt{\tfrac{(k-1)\pi}{2k}}\left\{\left|X_{i1}-\tfrac{1}{k-1}\Sigma_{j=2}^{k}X_{ij}\right|\right\}, \end{aligned} \tag{3.2.19}$$

the consistent and unbiased estimators of σ.

The sampling strategies Q_1 from (3.2.17) and $Q_{i,k}$ from (3.2.18) terminate w.p.1, that is, $P_{\boldsymbol{\theta}}\{N_{Q_1} < \infty\} = 1$ and $P_{\boldsymbol{\theta}}\{N_{Q_{i,k}} < \infty\} = 1$. Upon termination of sampling, we propose the point estimator $\overline{X}_{k,Q_{i,k}}$ for μ based on the finally accrued data. The associated *sequential risk* is given by:

$$R_{k,N_{Q_{i,k}}}(c_k) = E_{\boldsymbol{\theta}}\left[L_{k,N_{Q_{i,k}}}(\mu, \overline{X}_{k,N_{Q_{i,k}}})\right] = A\sigma^2 k^{-1} E_{\boldsymbol{\theta}}[N_{Q_{i,k}}^{-1}] + kc_k E_{\boldsymbol{\theta}}[N_{Q_{i,k}}]. \tag{3.2.20}$$

In order to compare the achieved sequential risk $R_{k,N_{Q_{i,k}}}(c_k)$ from (3.2.20) with the optimal fixed-sample-size risk R_{k,n_k^*} from (3.2.16), we consider the standard requisite entities given by:

$$\begin{aligned} \text{Risk efficiency:} \quad \xi_{Q_{i,k}}(c_k) &\equiv \frac{R_{k,N_{Q_{i,k}}}(c_k)}{R_{k,n_k^*}} = \tfrac{1}{2}E_{\boldsymbol{\theta}}[N_{Q_{i,k}}n_k^{*-1}] + \tfrac{1}{2}E_{\boldsymbol{\theta}}[n_k^* N_{Q_{i,k}}^{-1}], \\ \text{Regret:} \quad \omega_{Q_{i,k}}(c_k) &\equiv R_{k,N_{Q_{i,k}}}(c_k) - R_{k,n_k^*} = kc_k E_{\boldsymbol{\theta}}\left[\frac{(N_{Q_{i,k}} - n_k^*)^2}{N_{Q_{i,k}}}\right], \end{aligned} \tag{3.2.21}$$

These metrics were originally defined by Robbins (1959). We will explore asymptotic behaviors as $c_k \to 0$. Some asymptotic first-order properties may be summarized as follows (as $c_k \to 0$):

$$\begin{aligned} &N_{Q_{i,k}}/n_k^* \xrightarrow{P_{\boldsymbol{\theta}}} 1;\ E_{\boldsymbol{\theta}}[N_{Q_{i,k}}/n_k^*] \to 1 \ [\textit{Asymptotic first-order efficiency}]; \\ &\xi_{Q_{i,k}}(c_k) \to 1 \text{ if } m_0 \geq 3 \ [\textit{Asymptotic risk efficiency}]; \end{aligned} \tag{3.2.22}$$

Further, let $H_i^{(2)} = n_k^{*-1/2}(N_{Q_{i,k}} - n_k^*)$ and borrow the results from nonlinear renewal theory to claim:

$$\begin{aligned} &(i) \quad P_{\boldsymbol{\theta}}\{N_{Q_{i,k},c_k} \leq \varepsilon n_k^*\} = O(n_k^{*-m_0}) \textit{ if } 0 < \epsilon < 1 \textit{ and } m_0 \geq 2,\ i = 1, 2; \\ &(ii) \quad \left|H_{i,k}^{(2)}\right|^\kappa \textit{ is uniformly integrable if } m_0 > \max\{1, \tfrac{1}{2}\kappa\},\ i = 1, 2; \\ &(iii) \quad \left.\begin{array}{l} (a)\ H_{1,k}^{(2)} \xrightarrow{\pounds} N(0, \tfrac{1}{2}) \\ (b)\ H_{2,k}^{(2)} \xrightarrow{\pounds} N(0, \tfrac{1}{2}\pi - 1) \end{array}\right\} \textit{ if } m_0 \geq 2; \\ &(iv) \quad \left.\begin{array}{l} (a)\ E_\theta[N_{Q_{1,k},c_k} - n_k^*] = \eta_{1,k}^{(2)} + o(1) \\ (b)\ E_\theta[N_{Q_{2,k},c_k} - n_k^*] = \eta_{2,k}^{(2)} + o(1) \end{array}\right\} \text{ if } m_0 \geq 3. \end{aligned} \tag{3.2.23}$$

with $\eta_{1,k}^{(2)} \approx 0.1330$ and $\eta_{2,k}^{(2)} \approx 0.1331$ respectively. The superscript (2) used here indicates the MRPE strategies.

3.2.2.1. Asymptotic Second-Order Regret Under Generality

In the spirit of Section 3.2.1.1, we consider a general multi-stage sampling strategy leading to a terminal stopping time $M \equiv M_{c_k}$, defined appropriately, so that we are assured "closeness" between M_{c_k} and n_k^*. Let us now list a number of *assumptions* that may be needed in order to draw certain specific conclusion(s):

(B1): For every fixed $n \geq m_0 (> 1)$, the pilot size, the event $[M_{c_k} = n]$ depends only on the statistic $\mathbf{Z}_{k,n} \equiv \{X_{i1} - X_{ik}, ..., X_{ik-1} - X_{ik}\}, i = 1, ..., n$;

(B2): $M_{c_k}/n_k^* \stackrel{P_{\boldsymbol{\theta}}}{\to} 1$ as $c_k \to 0$, and also $P_{\boldsymbol{\theta}}\{M_{c_k} \leq \varepsilon n_k^*\} = O(n_k^{*-a_1 m_0 + a_2})$ if $0 < \varepsilon < 1$, with some $a_1 > 0$, $a_2 \geq 0$ if $m_0 > a_2/a_1$;

(B3): $H \equiv n_k^{*-1/2}(M_{c_k} - n_k^*) \stackrel{\pounds}{\to} N(0, a_3)$ as $c_1 \to 0,$ with some $a_3 > 0$;

(B4): $|H|^2$ is uniformly integrable if $m_0 > a_4 (> 0)$; and

(B5): $E_{\boldsymbol{\theta}}\left[M_{c_k} - n_k^*\right] = a_5 + o(1)$ as $c_k \to 0$ if $m_0 > a_6 (> 0)$.

(3.2.24)

for every fixed μ, σ.

Theorem 3.2.2 (Mukhopadhyay and Wang, 2020). *Under a general asymptotically second-order efficient multi-stage minimum risk point estimation methodology $(M_{c_k}, \overline{X}_{k,M_{c_k}})$, under the standing assumptions (B1)-(B5) from (3.2.24), we have the following asymptotic second-order expansion of the regret as $c_k \to 0$:*

$$\omega_{Q_k}(c_k) \equiv k c_k E_{\boldsymbol{\theta}}\left[(M_{c_k} - n_k^*)^2 M_{c_k}^{-1}\right] = k a_3 c_k + o(c_k) \text{ if } m_0 > \max\{a_6, a_4, a_1^{-1}(3 + a_2)\}, \tag{3.2.25}$$

with n_k^* coming from (3.2.16).

Proof: Assumption (B5) states that our general multi-stage methodology is asymptotically second-order efficient in the Ghosh-Mukhopadhyay (1981) sense.

Now, we can express:

$$E_{\boldsymbol{\theta}}\left[\frac{(M_{c_k}-n_k^*)^2}{M_{ck}}\right] = E_{\boldsymbol{\theta}}\left[\frac{(M_{c_k}-n_k^*)^2}{M_{c_k}} I(M_{c_k} \le \tfrac{1}{2}n_k^*)\right] + E_{\boldsymbol{\theta}}\left[\frac{(M_{c_k}-n_k^*)^2}{M_{c_k}} I(M_{c_k} > \tfrac{1}{2}n_k^*)\right],$$

and handle each term. Refer to Mukhopadhyay and Wang (2020, Theorem 2.2.) for other desired results. ■

For the stopping rule defined in (3.2.18), *as* $c_k \to 0$, we have following second order property:

$$\omega_{Q_{i,k}}(c_k) \equiv R_{k,N_{Q_{i,k}}}(c_k) - R_{k,n_k^*} = kc_k p_{i,k} + o(c_k),\ i = 1, 2, \tag{3.2.26}$$

with $p_{1,k} = \frac{1}{2}$ and $p_{2,k} = \frac{1}{2}\pi - 1$ if $m_0 \ge 4.n_k^*$

3.3. New Estimators of The Scale By Permuting Within Groups And Updated Purely Sequential Estimation Strategies

In the context of the FWCI (or MRPE) problems, we defined boundary crossing criteria in (3.2.6) (or in (3.2.18)) with help of estimators $U_{k,n}^{(1)}, U_{k,n}^{(2)}$ (or $T_{k,n}^{(1)}, T_{k,n}^{(2)}$) defined via (3.2.7) (or via (3.2.19)). But, one will notice quickly that we arrived at those consistent and unbiased estimators, namely the U's and the T's, for the scale parameter σ^2 or σ respectively by comparing only the first observation X_{i1} with the average $\frac{1}{k-1}\Sigma_{j=2}^{k} X_{ij}$ of the remaining $(k-1)$ observations within each k-tuple, $i = 1, 2, \ldots$.

But, within each k-tuple, we could instead compare the l^{th} observation X_{il} with the average $\frac{1}{k-1}\Sigma_{j\neq l}^{k} X_{ij}$ of the remaining $(k-1)$ observations within each group with $l = 1, ..., k$, $i = 1, 2, ...$. That is, within each k-tuple, we can permute the observations in order to come up with new classes of U's and T's, to estimate the scale parameter σ^2 or σ. In this section, we will first introduce such new estimators, namely the updated versions of U's and the T's, for the scale parameter σ^2 or σ respectively followed by appropriately defined new stopping times in the spirits of (3.2.6) and (3.2.18).

Let us incorporate a customary abbreviation, MAD, for the *mean absolute deviation.* While we consider the i^{th} observation, a k-tuple vector $(X_{i1}, X_{i2}, ..., X_{ik})$ consisting of independent random variables, we may express the following statistics:

$$\begin{aligned} \text{within sample mean:} \quad & \overline{X}_i \equiv k^{-1}\Sigma_{j=1}^{k} X_{ij}, \\ \text{within sample variance:} \quad & S_i^2 \equiv (k-1)^{-1}\Sigma_{j=1}^{k}(X_{ij} - \overline{X}_i)^2, \\ \text{within sample MAD:} \quad & M_i \equiv k^{-1}\Sigma_{j=1}^{k} \left| X_{ij} - \overline{X}_i \right|, \end{aligned}$$

for fixed but arbitrary i and $k(\geq 2)$, $i = 1, 2, ..., n, ...$.

Next, we begin with the following constructs from the i^{th} observation vector:

$$\begin{array}{lll} X_{i1} - \frac{1}{k-1}\Sigma_{j\neq 1}X_{ij} & = X_{i1} + \frac{1}{k-1}X_{i1} - \frac{1}{k-1}\left(\Sigma_{j\neq 1}X_{ij} + X_{i1}\right) & = \frac{k}{k-1}\left(X_{i1} - \overline{X}_i\right) \\ X_{i2} - \frac{1}{k-1}\Sigma_{j\neq 2}X_{ij} & = X_{i2} + \frac{1}{k-1}X_{i2} - \frac{1}{k-1}\left(\Sigma_{j\neq 2}X_{ij} + X_{i2}\right) & = \frac{k}{k-1}\left(X_{i2} - \overline{X}_i\right) \\ \quad\cdot & \quad\cdot & \quad\cdot \\ \quad\cdot & \quad\cdot & \quad\cdot \\ \quad\cdot & \quad\cdot & \quad\cdot \\ X_{ik} - \frac{1}{k-1}\Sigma_{j\neq k}X_{ij} & = X_{ik} + \frac{1}{k-1}X_{ik} - \frac{1}{k-1}\left(\Sigma_{j\neq k}X_{ij} + X_{ik}\right) & = \frac{k}{k-1}\left(X_{ik} - \overline{X}_i\right), \end{array}$$

and then obviously, we also have:

$$\begin{array}{ll}
\left|X_{i1}-\frac{1}{k-1}\Sigma_{j\neq 1}X_{ij}\right| & =\frac{k}{k-1}\left|X_{i1}-\overline{X}_i\right| \\
\left|X_{i2}-\frac{1}{k-1}\Sigma_{j\neq 2}X_{ij}\right| & =\frac{k}{k-1}\left|X_{i2}-\overline{X}_i\right| \\
. & . \\
. & . \\
. & . \\
\left|X_{ik}-\frac{1}{k-1}\Sigma_{j\neq k}X_{ij}\right| & =\frac{k}{k-1}\left|X_{ik}-\overline{X}_i\right|
\end{array}$$

We can clearly rewrite:

$$\Sigma_{l=1}^{k}\left\{X_{il}-\tfrac{1}{k-1}\Sigma_{j\neq l}^{k}X_{ij}\right\}^2=\tfrac{k^2}{(k-1)^2}\Sigma_{l=1}^{k}\left(X_{il}-\overline{X}_i\right)^2=\tfrac{k^2}{(k-1)}S_i^2,\text{ and also}$$
$$\Sigma_{l=1}^{k}\left|X_{il}-\tfrac{1}{k-1}\Sigma_{j\neq l}^{k}X_{ij}\right|=\tfrac{k}{(k-1)}\Sigma_{l=1}^{k}\left|X_{il}-\overline{X}_i\right|=\tfrac{k^2}{(k-1)}M_i.$$

Hence, we have:

$$\begin{aligned}
&\Sigma_{i=1}^{n}\Sigma_{l=1}^{k}\left\{X_{il}-\tfrac{1}{k-1}\Sigma_{j\neq l}^{k}X_{ij}\right\}^2=\tfrac{k^2}{(k-1)}\Sigma_{i=1}^{n}S_i^2;\\
&\Rightarrow E_{\boldsymbol{\theta}}\left[\Sigma_{i=1}^{n}\Sigma_{l=1}^{k}\left\{X_{il}-\tfrac{1}{k-1}\Sigma_{j\neq l}^{k}X_{ij}\right\}^2\right]=\tfrac{k^2}{(k-1)}n\sigma^2;\\
&\Rightarrow U_{k,n}^{\#(1)}\equiv\tfrac{(k-1)}{nk^2}\Sigma_{i=1}^{n}\Sigma_{l=1}^{k}\left\{X_{il}-\tfrac{1}{k-1}\Sigma_{j\neq l}^{k}X_{ij}\right\}^2=n^{-1}\Sigma_{i=1}^{n}S_i^2;
\end{aligned} \tag{3.3.1}$$

so that $U_{k,n}^{\#(1)}$ is an unbiased and consistent estimator for σ^2 constructed in a way that is parallel to $U_{k,n}^{(1)}$ from (3.2.7). We emphasize the symbol # in the superscript which ought to remind us that we are now permuting within each i^{th} vector observation, $i=1,2,...$. Observe that S_i^2's are i.i.d. for $i=1,2,...,n,...$.

Parallel to $T_{k,n}^{(1)}$ in (3.2.19), let us construct an unbiased estimator of σ starting from

$U_{k,n}^{\#(1)}$. Suppose that $Y \sim \chi_\nu^2$ and then we may rewrite:

$$E_{\boldsymbol{\theta}}\left[U_{k,n}^{\#(1)1/2}\right] = \sigma\{(k-1)n\}^{-1/2}E[Y^{1/2}] = \sigma\{(k-1)n\}^{-1/2}r_{k,n} \text{ where}$$
$$r_{k,n} = 2^{1/2}\Gamma(\tfrac{(k-1)n}{2} + \tfrac{1}{2})\{\Gamma(\tfrac{(k-1)n}{2})\}^{-1}. \tag{3.3.2}$$

So, we may define an unbiased and consistent estimator for σ as follows:

$$T_{k,n}^{\#(1)} \equiv \{(k-1)n\}^{1/2}r_{k,n}^{-1}U_{k,n}^{\#(1)1/2} = \{(k-1)n\}^{1/2}r_{k,n}^{-1}\{n^{-1}\Sigma_{i=1}^{n}S_i^2\}^{1/2} \text{ where}$$
$$\{(k-1)n\}^{1/2}r_{k,n}^{-1} = 1 + \tfrac{1}{4(k-1)}n^{-1} + O(n^{-2}). \tag{3.3.3}$$

We also obtain:

$$\Sigma_{l=1}^{k}\left|X_{il} - \tfrac{1}{k-1}\Sigma_{j\neq l}^{k}X_{ij}\right| = \tfrac{k}{(k-1)}\Sigma_{l=1}^{k}\left|X_{il} - \overline{X}_i\right| = \tfrac{k^2}{(k-1)}M_i,$$

and we can easily verify:

$$E_{\boldsymbol{\theta}}[M_1] = \sigma\sqrt{\tfrac{2(k-1)}{\pi k}},$$

From Herrey (1965), we can express:

$$E_{\boldsymbol{\theta}}[M_1^2] = g_{1k}\sigma^2 \Rightarrow V_{\boldsymbol{\theta}}[M_1] = E_{\boldsymbol{\theta}}[M_1^2] - E_{\boldsymbol{\theta}}^2[M_1] = g_{2k}\sigma^2 \text{ where we denote:}$$
$$g_{1k} \equiv \left\{\left(\tfrac{\pi}{2} + \sin^{-1}\left(\tfrac{1}{k-1}\right) - k + \sqrt{k(k-2)}\right)\tfrac{2(k-1)}{k^2\pi} + \tfrac{2(k-1)}{\pi k}\right\}, \text{ and}$$
$$g_{2k} \equiv \left(\tfrac{\pi}{2} + \sin^{-1}\left(\tfrac{1}{k-1}\right) - k + \sqrt{k(k-2)}\right)\tfrac{2(k-1)}{k^2\pi}. \tag{3.3.4}$$

Now, having recorded the i^{th} observation, a k-tuple vector of independent random variables,

$$(X_{i1}, ..., X_{ik}), i = 1, ..., n,$$

we have:

$$\Sigma_{i=1}^{n}\Sigma_{l=1}^{k}\left|X_{il}-\tfrac{1}{k-1}\Sigma_{j\neq l}^{k}X_{ij}\right|=\tfrac{k^2}{(k-1)}\Sigma_{i=1}^{n}M_i;$$
$$\Rightarrow E_{\boldsymbol{\theta}}\left[\Sigma_{i=1}^{n}\Sigma_{l=1}^{k}\left|X_{il}-\tfrac{1}{k-1}\Sigma_{j\neq l}^{k}X_{ij}\right|\right]=\sqrt{\tfrac{2k}{\pi(k-1)}}nk\sigma; \tag{3.3.5}$$
$$\Rightarrow T_{k,n}^{\#(2)}\equiv\tfrac{1}{nk}\sqrt{\tfrac{\pi(k-1)}{2k}}\Sigma_{i=1}^{n}\Sigma_{l=1}^{k}\left|X_{il}-\tfrac{1}{k-1}\Sigma_{j\neq l}^{k}X_{ij}\right|=\sqrt{\tfrac{\pi k}{2(k-1)}}n^{-1}\Sigma_{i=1}^{n}M_i.$$

This $T_{k,n}^{\#(2)}$ is unbiased and consistent for σ constructed in a way that is parallel to our older $T_{k,n}^{(2)}$ from (3.2.19). Observe that M_i's are i.i.d. for $i=1,2,...,n,...$.

Lastly, let us find an unbiased estimator of σ^2 from $T_{k,n}^{\#(2)}$:

$$E_{\boldsymbol{\theta}}\left[T_{k,n}^{\#(2)2}\right]=\tfrac{\pi k}{2(k-1)n^2}E_{\boldsymbol{\theta}}\left[\left(\Sigma_{i=1}^{n}M_i\right)^2\right]=\tfrac{\pi k\sigma^2}{2(k-1)}\left\{g_{1k}n^{-1}+(1-n^{-1})\tfrac{2(k-1)}{\pi k}\right\}$$
$$=\left\{\left(\tfrac{\pi k g_{1k}}{2(k-1)}-1\right)n^{-1}+1\right\}\sigma^2\Rightarrow E_{\boldsymbol{\theta}}\left[\left\{\left(\tfrac{\pi k g_{1k}}{2(k-1)}-1\right)n^{-1}+1\right\}^{-1}T_{k,n}^{\#(2)2}\right]=\sigma^2.$$

where g_{1k} was defined in (3.3.4). Next, we define:

$$U_{k,n}^{\#(2)}\equiv\left\{\left(\tfrac{\pi k g_{1k}}{2(k-1)}-1\right)n^{-1}+1\right\}^{-1}T_{k,n}^{\#(2)2}, \tag{3.3.6}$$

another unbiased and consistent estimator for σ^2 where $T_{k,n}^{\#(2)}$ was defined in (3.3.5) and g_{1k} comes from (3.3.4).

3.3.1. Comparing Variances

We briefly embark upon comparing $V_{\boldsymbol{\theta}}\left[U_{k,n}^{\#(i)}\right]$ with $V_{\boldsymbol{\theta}}\left[U_{k,n}^{(i)}\right]$ as well as $V_{\boldsymbol{\theta}}\left[T_{k,n}^{\#(i)}\right]$ with $V_{\boldsymbol{\theta}}\left[T_{k,n}^{(i)}\right]$, $i=1,2$. In defining the new estimators $U_{k,n}^{\#(i)}, T_{k,n}^{\#(i)}$, since we have incorporated within group permutation, we may assume without any loss of generality that $k\geq 3$. The U's and $U^{\#}$'s (or T's and $T^{\#}$'s) respectively correspond to the FWCI (or MRPE) strategies.

Case 1: Comparing $V_{\boldsymbol{\theta}}\left[U_{k,n}^{\#(1)}\right]$ with $V_{\boldsymbol{\theta}}\left[U_{k,n}^{(1)}\right]$: Recall (3.3.1). Since S_i's are independent, we have:

$$V_{\boldsymbol{\theta}}\left[U_{k,n}^{\#(1)}\right] = n^{-1}V_{\boldsymbol{\theta}}\left[S_i^2\right] = \tfrac{2\sigma^4}{(k-1)n} \tag{3.3.7}$$

From Mukhopadhyay and Wang (2020), we express:

$$U_{k,n}^{(1)} = \tfrac{1}{n}\sum_{i=1}^{n}\sigma^2 Z_i \text{ where } Z_i = \tfrac{k-1}{k\sigma^2}\left[X_{i1} - \tfrac{1}{k-1}\sum_{j=2}^{k}X_{ij}\right]^2 \text{ are i.i.d. } \chi_1^2 \Rightarrow V_{\boldsymbol{\theta}}\left[U_{k,n}^{(1)}\right] = \tfrac{2\sigma^4}{n}. \tag{3.3.8}$$

From (3.3.7)-(3.3.8), we immediately conclude that $V_{\boldsymbol{\theta}}\left[U_{k,n}^{\#(1)}\right] \geq V_{\boldsymbol{\theta}}\left[U_{k,n}^{(1)}\right]$, equality holds if and only if $k = 2$.

Table 3.1. Simulated Variances of $U_{k,n}^{(2)}$ and $U_{k,n}^{\#(2)}$ from $N(0,1)$: 10000 replications

	$n=3$		$n=4$		$n=5$		$n=10$	
k	$\widehat{V}[U_{k,n}^{(2)}]$	$\widehat{V}[U_{k,n}^{\#(2)}]$	$\widehat{V}[U_{k,n}^{(2)}]$	$\widehat{V}[U_{k,n}^{\#(2)}]$	$\widehat{V}[U_{k,n}^{(2)}]$	$\widehat{V}[U_{k,n}^{\#(2)}]$	$\widehat{V}[U_{k,n}^{(2)}]$	$\widehat{V}[U_{k,n}^{\#(2)}]$
3	0.730	0.350	0.548	0.271	0.441	0.214	0.222	0.109
5	0.732	0.183	0.554	0.140	0.423	0.108	0.229	0.056
10	0.715	0.083	0.557	0.063	0.457	0.050	0.230	0.025
15	0.727	0.053	0.559	0.041	0.453	0.032	0.217	0.016
20	0.687	0.039	0.547	0.030	0.457	0.023	0.226	0.012
25	0.719	0.031	0.556	0.023	0.460	0.019	0.227	0.009
30	0.715	0.026	0.545	0.020	0.458	0.016	0.220	0.008

Case 2: Comparing $V_{\boldsymbol{\theta}}\left[U_{k,n}^{\#(2)}\right]$ with $V_{\boldsymbol{\theta}}\left[U_{k,n}^{(2)}\right]$: Since the analytical expressions of these variances are complicated, for selected pairs (k,n), we replicated the process of gathering observations 10000 times from a standard normal distribution and obtained simulated esti-

mates of the respected variances. We choose $k = 3, 5, 10, 15, 20, 25, 30$ and $n = 3, 4, 5, 10$. Table 1 shows summaries and it appears that $U_{k,n}^{\#(2)}$ tends to have a smaller variance than $U_{k,n}^{(2)}$.

Case 3: Comparing $V_{\boldsymbol{\theta}}\left[T_{k,n}^{\#(1)}\right]$ with $V_{\boldsymbol{\theta}}\left[T_{k,n}^{(1)}\right]$: First, we observe that

$$V_{\boldsymbol{\theta}}\left[T_{k,n}^{\#(1)}\right] = n(k-1)r_{k,n}^{-2}V_{\boldsymbol{\theta}}\left[U_{k,n}^{\#(1)1/2}\right] \text{ with } r_{k,n} \text{ defined in (3.3.2),}$$

and then,

$$V_{\boldsymbol{\theta}}\left[U_{k,n}^{\#(1)1/2}\right] = E_{\boldsymbol{\theta}}\left[U_{k,n}^{\#(1)}\right] - \left\{\sigma\left[(k-1)n\right]^{-1/2} r_{k,n}\right\}^2,$$

Then, we obtain:

$$V_{\boldsymbol{\theta}}\left[T_{k,n}^{\#(1)}\right] = \sigma^2\left[n(k-1)r_{k,n}^{-2} - 1\right] = \sigma^2\left[\tfrac{1}{2(k-1)n} + O(n^{-2})\right]. \tag{3.3.9}$$

Next, from (3.3.8), we recall:

$$T_{k,n}^{(1)} = \Gamma\left(\tfrac{n}{2}\right)\Gamma\left(\tfrac{n+1}{2}\right)^{-1}\left[\tfrac{\sigma^2}{2}\textstyle\sum_{i=1}^{n} Z_i\right]^{1/2} \text{ where } Z_i\text{'s are i.i.d. } \chi_1^2.$$

Thus,

$$V\left[T_{k,n}^{(1)}\right] = \sigma^2\left[\tfrac{n}{2}\left\{\Gamma\left(\tfrac{n}{2}\right)\right\}^2\left\{\Gamma\left(\tfrac{n+1}{2}\right)\right\}^{-2} - 1\right] = \sigma^2\left\{\tfrac{1}{2n} + O(n^{-2})\right\}. \tag{3.3.10}$$

This expressions does not depend on k. Clearly, when $k = 2$, the two variances from (3.3.9)-(3.3.10) are identical. But, again, for selected pairs (k, n), we replicated the process of gathering observations 10000 times from a standard normal distribution and obtained simulated estimates of the respected variances. We choose $k = 3, 5, 10, 15, 20, 25, 30$ and $n = 3, 4, 5, 10$. Table 2 shows summaries and it appears that $T_{k,n}^{\#(1)}$ tends to have a smaller variance than $T_{k,n}^{(1)}$.

Table 3.2. Simulated variances of $T_{k,n}^{(1)}$ and $T_{k,n}^{\#(1)}$ from $N(0,1)$: 10000 replications

	$n=3$		$n=4$		$n=5$		$n=10$	
k	$V[T_{k,n}^{(1)}]$	$V[T_{k,n}^{\#(1)}]$	$V[T_{k,n}^{(1)}]$	$V[T_{k,n}^{\#(1)}]$	$V[T_{k,n}^{(1)}]$	$V[T_{k,n}^{\#(1)}]$	$V[T_{k,n}^{(1)}]$	$V[T_{k,n}^{\#(1)}]$
3	0.178	0.086	0.132	0.064	0.104	0.051	0.051	0.025
5	0.178	0.042	0.132	0.032	0.104	0.025	0.051	0.013
10	0.178	0.019	0.132	0.014	0.104	0.011	0.051	0.006
15	0.178	0.012	0.132	0.009	0.104	0.007	0.051	0.004
20	0.178	0.009	0.132	0.007	0.104	0.005	0.051	0.003
25	0.178	0.007	0.132	0.005	0.104	0.004	0.051	0.002
30	0.178	0.006	0.132	0.004	0.104	0.003	0.051	0.002

Case 4: Comparing $V_{\boldsymbol{\theta}}\left[T_{k,n}^{\#(2)}\right]$ with $V_{\boldsymbol{\theta}}\left[T_{k,n}^{(2)}\right]$: First, from (3.3.5), we observe that

$$V\left[T_{k,n}^{\#(2)}\right] = \left\{\tfrac{\pi}{2nk} + \tfrac{1}{nk}\left[\sqrt{k(k-2)} + \sin^{-1}\left(\tfrac{1}{k-1}\right)\right] - \tfrac{1}{n}\right\}\sigma^2. \tag{3.3.11}$$

We also recall from Mukhopadhyay and Wang (2020):

$$V\left[T_{k,n}^{(2)}\right] = \left\{\tfrac{\pi}{2n} + \tfrac{n-1}{n(k-1)}\left[\sqrt{k(k-2)} + \sin^{-1}\left(\tfrac{1}{k-1}\right)\right] - 1\right\}\sigma^2. \tag{3.3.12}$$

In the spirits of Tables 3.1-3.2, we provide simulated estimates of $V\left[T_{k,n}^{(2)}\right]$ and $V\left[T_{k,n}^{\#(2)}\right]$

using 10000 replications from a standard normal distribution for selected pairs (k, n). We choose $k = 3, 5, 10, 15, 20, 25, 30$ and $n = 3, 4, 5, 10$. Table 3.3 shows summaries and it appears that $T^{\#(2)}_{k,n}$ tends to have a smaller variance than $T^{(2)}_{k,n}$.

Table 3.3. Simulated variances of $T^{(2)}_{k,n}$ and $T^{\#(2)}_{k,n}$ from $N(0,1)$: 10000 replications

	$n=3$		$n=4$		$n=5$		$n=10$	
k	$V[T^{(2)}_{k,n}]$	$V[T^{\#(2)}_{k,n}]$	$V[T^{(2)}_{k,n}]$	$V[T^{\#(2)}_{k,n}]$	$V[T^{(2)}_{k,n}]$	$V[T^{\#(2)}_{k,n}]$	$V[T^{(2)}_{k,n}]$	$V[T^{\#(2)}_{k,n}]$
3	0.275	0.092	0.239	0.069	0.216	0.055	0.172	0.028
5	0.211	0.046	0.166	0.035	0.139	0.028	0.085	0.014
10	0.194	0.021	0.147	0.016	0.119	0.013	0.063	0.006
15	0.192	0.013	0.145	0.010	0.116	0.008	0.059	0.004
20	0.191	0.010	0.144	0.007	0.115	0.006	0.058	0.003
25	0.191	0.008	0.143	0.006	0.115	0.005	0.058	0.002
30	0.190	0.007	0.143	0.005	0.115	0.004	0.058	0.002

3.3.2. Purely Sequential Stopping Times Incorporating Permuted Estimators

We move ahead by incorporating new unbiased and consistent estimators of (i) σ^2, namely, $U^{\#}_{k,n}$'s, and (ii) σ, namely, $T^{\#}_{k,n}$'s, in order to define updated purely sequential (i) FWCI methodologies and (ii) MRPE methodologies respectively. We formally propose the following strategies:

$$\text{Methodology } \mathcal{P}^{\#}_{i,k}: \quad N_{\mathcal{P}^{\#}_{i,k},d} \equiv N_{\mathcal{P}^{\#}_{i,k}} = \inf\left\{n \geq m_0(\geq 2); \quad n \geq a^2 U^{\#(i)}_{k,n}/(kd^2)\right\} \\ \text{where } U^{\#(i)}_{k,n} \text{ are defined in (3.3.1) and (3.3.6);} \tag{3.3.13}$$

as well as

$$\text{Methodology } Q^{\#}_{i,k}: \quad N_{Q^{\#}_{i,k},c_k} \equiv N_{Q^{\#}_{i,k}} = \inf\left\{n \geq m_0(\geq 2): \quad n \geq \tfrac{1}{k}\sqrt{A/c_k}T^{\#(i)}_{k,n}\right\}$$
$$\text{where } T^{\#(i)}_{k,n} \text{ are defined in (3.3.3) and (3.3.5);} \tag{3.3.14}$$

$i = 1, 2$, much in the spirits of (3.2.6) and (3.2.17) respectively.

We may reiterate that $U^{\#(i)}_{k,n}$'s are unbiased for σ^2 and $T^{\#(i)}_{k,n}$'s are unbiased for σ, $i = 1, 2$. We also expect $U^{\#(i)}_{k,n}, T^{\#(i)}_{k,n}$ to have smaller variances than those of $U^{(i)}_{k,n}, T^{(i)}_{k,n}$ respectively for each $i = 1, 2$. Thus, it makes sense to propose (3.3.13) and (3.3.14) in order to update (3.2.6) and (3.2.17) respectively. Our second-order theories as well as comparisons of the ensuing FWCI and MRPE strategies via simulations will reveal the benefits derived from such substitutions. More discussions will follow.

3.4. Purely Sequential FWCI Strategies Under Sampling In Groups Incorporating Within Group Permutation

Now, we are going to update the asymptotic first-order and second-order properties of the purely sequential FWCI estimation strategies $\mathcal{P}^{\#}_{i,k}$ from (3.3.13), $i = 1, 2$. We emphasize that $\mathcal{P}_{i,k}$ (or its updated version, $\mathcal{P}^{\#}_{i,k}$) from (3.2.6) (or from (3.3.13)) used $U^{(i)}_{k,n}$ (or its updated version, $U^{\#(i)}_{k,n}$) in the definition of the boundary crossing, $i = 1, 2$.

We expect $\mathcal{P}^{\#}_{i,k}$'s to perform better than $\mathcal{P}_{i,k}$'s in view of (i) $U^{(i)}_{k,n}, U^{\#(i)}_{k,n}$ are both unbiased and consistent estimators of σ^2, and (ii) $V_{\boldsymbol{\theta}}[U^{\#(i)}_{k,n}]$ is expected not to exceed $V_{\boldsymbol{\theta}}[U^{(i)}_{k,n}]$, $i = 1, 2$. We will soon explore where and how such intuitively expected additional benefit of implementing $\mathcal{P}^{\#}_{i,k}$ over $\mathcal{P}_{i,k}$ may manifest itself.

3.4.1. Preliminaries and Selected First-Order Properties

We begin by stating a number of interesting results (Lemmas 3.4.1-3.4.4) without giving too many specific details towards their full justifications.

Lemma 3.4.1. *Under the stopping rules* $\mathcal{P}^{\#}_{i,k}$ *defined in (3.3.13), for all fixed* $\boldsymbol{\theta}, d, \alpha, k,$ *and* $m_0(\geq 2)$, *the statistics* $U^{\#(i)}_{k,n}$ *and* $\overline{X}_{k,n}$ *are distributed independently for each fixed* $n \geq m_0$.

Proof: We only need to show (i) S_i^2 and $\overline{X}_i$ as well as (ii) M_i and $\overline{X}_i$ are independent for all fixed $n(\geq 2)$ and $i = 1, ..., n$. Assume $\sigma = \sigma_0(> 0)$ is a fixed but arbitrary. In this situation, $\overline{X}_i$ is complete and sufficient for μ, while S_i^2 and M_i are individually ancillary for μ. Thus having fixed $\sigma = \sigma_0$, using Basu's Theorem we have (i) S_i^2 and $\overline{X}_i$ are independent, (ii) M_i and $\overline{X}_i$ are independent. This independence doesn't depend on the choice of σ_0. Thus, we claim $U^{\#(i)}_{k,n}, i = 1, 2$ and $\overline{X}_{k,n}$ are distributed independently. ■

Lemma 3.4.2. *Denote* $G_{\mathcal{P}^{\#}_{i,k}} \equiv N^{1/2}_{\mathcal{P}^{\#}_{i,k}} \left(\overline{X}_{k,N_{\mathcal{P}^{\#}_{i,k}}} - \mu \right) \sigma^{-1}$ *under the stopping rule* $N_{P^{\#}_{i,k}}$ *defined in (3.3.13). Then, for all fixed* $\boldsymbol{\theta}, d, \alpha, k,$ *and* $m_0(\geq 2)$, *we have*:

$$G_{\mathcal{P}^{\#}_{i,k}} \sim N(0,1), \; E_{\boldsymbol{\theta}}\left[\overline{X}_{k,N_{\mathcal{P}^{\#}_{i,k}}}\right] = \mu, \; \textit{and} \; V_{\boldsymbol{\theta}}\left[\overline{X}_{k,N_{\mathcal{P}^{\#}_{i,k}}}\right] = k^{-1}\sigma^2 E_{\boldsymbol{\theta}}\left[N^{-1}_{\mathcal{P}^{\#}_{i,k}}\right],$$

$i = 1, 2.$

Proof: By (3.3.13), $N_{\mathcal{P}^{\#}_{i,k}}$ depends on $U^{\#(i)}_{k,n}$ only. Combine Lemma 3.4.1 and (3.3.1) to claim that $\overline{X}_{k,N_{\mathcal{P}^{\#}_{i,k}}}$ and $N_{\mathcal{P}^{\#}_{i,k}}$ are independent. The rest follows. ■

Lemma 3.4.3. *For the stopping time* $N_{\mathcal{P}^{\#}_{i,k}}$ *defined in (3.3.13), for all fixed* $\boldsymbol{\theta}, \alpha, k,$ *and*

$m_0(\geq 2)$, *we have as* $d \to 0$:

$$\begin{aligned} &(i) \quad N_{\mathcal{P}^{\#}_{i,k}}/C_k \overset{P_{\boldsymbol{\theta}}}{\to} 1; \text{ and} \\ &(ii) \quad E_{\boldsymbol{\theta}}\left[N_{\mathcal{P}^{\#}_{i,k}}/C_k\right] \to 1 \text{ [Asymptotic first-order efficiency];} \end{aligned} \tag{3.4.1}$$

parallel to (3.2.8), where $C_k(= a^2\sigma^2/(d^2k))$ *defined in (3.2.4) is the optimal fixed number of groups.*

Lemma 3.4.4. *For the stopping time* $N_{\mathcal{P}^{\#}_{i,k}}$ *from (3.2.13) and the proposed FWCI estimation strategy*

$$\left(N_{\mathcal{P}^{\#}_{i,k}}, J_{k,N_{\mathcal{P}^{\#}_{i,k}}} = \left[\overline{X}_{k,N_{\mathcal{P}^{\#}_{i,k}}} \pm d\right]\right),$$

for all fixed $\boldsymbol{\theta}, \alpha, k$, *and* $m_0(\geq 2)$, *we have:*

$$\lim_{d\to 0} P_{\boldsymbol{\theta}}\left\{\mu \in J_{k,N_{\mathcal{P}^{\#}_{i,k}}}\right\} = 1 - \alpha \quad \text{[Asymptotic consistency].} \tag{3.4.2}$$

Proof: We express the coverage probability as follows:

$$P_{\boldsymbol{\theta}}\left\{\mu \in J_{k,N_{\mathcal{P}^{\#}_{i,k}}}\right\} = E_{\boldsymbol{\theta}}\left\{2\Phi\left(\sqrt{kN_{\mathcal{P}^{\#}_{i,k}}}\, d/\sigma\right) - 1\right\}.$$

Now, Lemma 3.4.3, part (i) and the dominated convergence theorem together complete the proof. ■

3.4.2. Selected Asymptotic Second-Order Properties

We lay down Theorems 3.4.1-3.4.2 shortly summarizing a number of major asymptotic second-order properties associated with the stopping rules $\mathcal{P}^{\#}_{i,k}$ from (3.3.13) by relying upon nonlinear renewal theory from Woodroofe (1977,1982), Lai and Siegmund (1977,1979), Mukhopadhyay and Solanky (1994), Ghosh et al. (1997), and Mukhopadhyay and de Silva (2009). We begin by enumerating a number of essential technical details.

Case 1: $i = 1$

First, we rewrite the stopping time corresponding to $\mathcal{P}^{\#}_{1,k}$ from (3.3.13) as follows:

$$N_{\mathcal{P}^{\#}_{1,k}} = \inf\left\{n \geq m_0 : \quad n^2 C_k^{-1} \geq \Sigma_{i=1}^{n} W_i\right\} \text{ where } W_i = S_i^2/\sigma^2, \tag{3.4.3}$$

with $(k-1)W_i \overset{\text{i.i.d.}}{\sim} \chi^2_{k-1}$. Then, obviously, $E_{\boldsymbol{\theta}}\left[W_i\right] = 1$ and $V_{\boldsymbol{\theta}}\left[W_i\right] = \frac{2}{k-1}$. Next, for arbitrary $u(>0)$, we obtain (with $\widetilde{u} \propto u$):

$$P_{\boldsymbol{\theta}}\left\{W_1 \leq u\right\} = P_{\boldsymbol{\theta}}\left\{S_i^2 \leq \widetilde{u}\right\} \propto \int_0^{\widetilde{u}} y^{\frac{k-3}{2}} e^{-y/2} dy \leq B u^{(k-1)/2},$$

with $B(>0)$, not involving u.

Now, we compare this representation (3.4.3) with that used in nonlinear renewal theory (Mukhopadhyay and Solanky 1994, p. 49) to immediately express:

$$\lambda = 1, \tau^2 = \tfrac{2}{k-1}, h^* = C_k^{-1}, n^* = C_k, \delta = 2, \beta^* = 1, l_0 = 0, p = \tfrac{2}{k-1}, \text{ and } b = \tfrac{1}{2}(k-1). \tag{3.4.4}$$

We also write:

$$\begin{aligned} D^{\#(1)}_{1,k} &= \Sigma_{n=1}^{\infty} n^{-1} E\left\{\max(0, \Sigma_{i=1}^{n} W_i - 2n)\right\}; \\ \nu^{\#(1)}_{1,k} &= \tfrac{k+1}{2(k-1)} - D^{\#(1)}_{1,k}; \\ \eta^{\#(1)}_{1,k} &= \tfrac{k-3}{2(k-1)} - D^{\#(1)}_{1,k}. \end{aligned} \tag{3.4.5}$$

These expressions correspond to $U_{k,n}^{\#(1)}$. The superscript (1) used in (3.4.5) identifies with the FWCI problem.

Observe that indeed all three entities $D_{1,k}^{\#(1)}, v_{1,k}^{\#(1)}$ and $\eta_{1,k}^{\#(1)}$ depend on the choice of k which is the group size. We will incorporate simulations to estimate these entities for some specific choices of k in Section 3.4.3.1. Here and elsewhere, for any pre-fixed group size k, the entities $D_{1,k}^{\#(1)}, v_{1,k}^{\#(1)}$ and $\eta_{1,k}^{\#(1)}$ are treated as known constants.

Case 2: $i = 2$

Next, we rewrite the stopping time corresponding to $\mathcal{P}_{2,k}^{\#}$ from (3.3.13) as follows:

$$N_{\mathcal{P}_{2,k}^{\#}} = \inf\left\{n \geq m_0 : \quad n^{3/2} C_k^{-1/2}\left[1 + l_0 n^{-1} + o(n^{-1})\right] \geq \Sigma_{i=1}^{n} W_i'\right\}, \tag{3.4.6}$$

where $l_0 = \frac{\pi k g_{1k}}{4(k-1)} - \frac{1}{2}$ and $W_i' = \sqrt{\frac{\pi k}{2(k-1)}}\frac{M_i}{\sigma}$. Then, we have $E_{\boldsymbol{\theta}}\left[W_i'\right] = 1$, $V_{\boldsymbol{\theta}}\left[W_i'\right] = \frac{k\pi g_{1k}}{2(k-1)} - 1$. Here, g_{1k} comes from (3.3.4).

Again, for arbitrary $u(>0)$, we obtain (with $\widetilde{u} \propto u$):

$$P_{\boldsymbol{\theta}}\left\{W_1' \leq u\right\} = P_{\boldsymbol{\theta}}\left\{\tfrac{1}{k}\Sigma_{j=1}^{k}|X_j - \overline{X}_j| \leq \widetilde{u}\right\} \leq kP\left\{|X_{11} - \overline{X}_1| \leq \widetilde{u}\right\} \leq Bu,$$

with $B(>0)$, not involving u.

Now, by comparing this representation with that used in nonlinear renewal theory (Mukhopadhyay and Solanky 1994, p. 49), we immediately get:

$$\lambda = 1, \tau^2 = \tfrac{k\pi g_{1k}}{2(k-1)} - 1, h^* = C_k^{-1/2}, n^* = C_k, \delta = \tfrac{3}{2}, \beta^* = 2, l_0 = \tfrac{1}{2}\tau^2, p = 4\tau^2, \text{ and } b = 1. \tag{3.4.7}$$

We also obtain:

$$
\begin{aligned}
D_{2,k}^{\#(1)} &= \Sigma_{n=1}^{\infty} n^{-1} E\left\{\max(0, \Sigma_{i=1}^{n} W_i' - \tfrac{3}{2}n)\right\}; \\
\nu_{2,k}^{\#(1)} &= \tfrac{1}{4} + \tau^2 - D_{2,k}^{\#(1)}; \qquad (3.4.8) \\
\eta_{2,k}^{\#(1)} &= \tfrac{1}{2} - 2\tau^2 - 2D_{2,k}^{\#(1)}.
\end{aligned}
$$

Again, the superscript (1) used in (3.4.8) identifies with the FWCI problem.

Theorem 3.4.1. *For the stopping time $N_{\mathcal{P}_{i,k}^{\#}}$ defined by (3.3.13), we denote $H_{i,k}^{\#(1)} = C_k^{-1/2}(N_{\mathcal{P}_{i,k}^{\#}} - C_k)$. Then, for every fixed $\boldsymbol{\theta}, k$ and α, we have the following results as $d \to 0$:*

$$
\begin{array}{ll}
(i) & \left.\begin{array}{l} (a)\ P_{\boldsymbol{\theta}}\{N_{\mathcal{P}_{1,k}^{\#}} \le \varepsilon n_k^*\} = O(C_k^{-(k-1)m_0/2}) \\ (b)\ P_{\boldsymbol{\theta}}\{N_{\mathcal{P}_{2,k}^{\#}} \le \varepsilon n_k^*\} = O(C_k^{-m_0}) \end{array}\right\} \textit{if } 0 < \varepsilon < 1 \textit{ and } m_0 \ge 2; \\
(ii) & \left|H_{i,k}^{\#(1)}\right|^2 \textit{ is uniformly integrable if } m_0 \ge 2; \\
(iii) & \left.\begin{array}{l} (a)\ H_{1,k}^{\#(1)} \xrightarrow{\pounds} N(0, \frac{2}{k-1}) \\ (b)\ H_{2,k}^{\#(1)} \xrightarrow{\pounds} N(0, \frac{2k\pi g_{1k}}{(k-1)} - 4) \end{array}\right\} \textit{if } m_0 \ge 2; \\
(iv) & \left.\begin{array}{l} (a)\ E_{\boldsymbol{\theta}}[N_{\mathcal{P}_{1,k}^{\#}} - C_k] = \eta_{1,k}^{\#(1)} + o(1) \\ (b)\ E_{\boldsymbol{\theta}}[N_{\mathcal{P}_{2,k}^{\#}} - C_k] = \eta_{2,k}^{\#(1)} + o(1) \end{array}\right\} \text{if } m_0 \ge 3;
\end{array}
\qquad (3.4.9)
$$

with $\eta_{1,k}^{\#(1)}$, $\eta_{2,k}^{\#(1)}$ and g_{1k} coming from (3.4.5), (3.4.9) and (3.3.4) respectively.

Theorem 3.4.2. *For the stopping time $N_{\mathcal{P}_{i,k}^{\#}}$ defined by (3.2.13) with the proposed FWCI estimation strategy*

$$
\left(N_{\mathcal{P}_{i,k}^{\#}}, J_{k,\mathcal{P}_{i,k}^{\#}} = \left[\overline{X}_{k,\mathcal{P}_{i,k}^{\#}} \pm d\right]\right),
$$

for every fixed $\boldsymbol{\theta}, k$ *and* α, *we have the following result as* $d \to 0$:

$$P_{\boldsymbol{\theta}}\left\{\mu \in J_{k,\mathcal{P}^{\#}_{i,k}}\right\} = (1-\alpha) + \gamma^{\#}_{i,k} C_k^{-1} + o(d^2) \text{ [Asymptotic second-order consistency]}, \tag{3.4.10}$$

if $m_0 \geq 6$. *Here,* $\gamma^{\#}_{i,k} = \left\{\eta^{\#(1)}_{i,k} - \frac{1}{4}p^{\#}_{i,k}(a^2+1)\right\} a\phi(a)$, $i = 1, 2$, *with* $p^{\#}_{1,k} = \frac{2}{k-1}, p^{\#}_{2,k} = \frac{k\pi g_{1k}}{2(k-1)} - 1$, *but* $\eta^{\#(1)}_{1,k}$, $\eta^{\#(1)}_{2,k}$, g_{1k} *come from (3.4.5), (3.4.9) and (3.3.4) respectively.*

Proof: We revisit the assumptions (A1)-(A5) from (3.2.10) in Section 3.2.1.1. They hold for the sampling strategy defined by (3.3.13) along with the following details:

$$i = 1: \quad a_1 = \tfrac{k-1}{2}, a_2 = 0, a_3 = \tfrac{2}{k-1}, a_4 = 1, a_5 = \eta^{\#(1)}_{1,k} \text{ and } a_6 = 2;$$
$$i = 2: \quad a_1 = 1, a_2 = 0, a_3 = \tfrac{2k\pi g_{1k}}{(k-1)} - 4, a_4 = 1, a_5 = \eta^{\#(1)}_{2,k} \text{ and } a_6 = 2.$$

This result follows immediately from the more general statement given in Theorem 3.2.1. ■

Once we combine (3.2.9) parts (ii) and (iii), we conclude the following results if $m_0 \geq 3$:

$$\begin{aligned} V_{\boldsymbol{\theta}}[N_{\mathcal{P}_{1,k}}] &= 2C_k + o(C_k) \Rightarrow C_k^{-1} V_{\boldsymbol{\theta}}[N_{\mathcal{P}_{1,k}}] \approx 2, \text{ and} \\ V_{\boldsymbol{\theta}}[N_{\mathcal{P}_{2,k}}] &= \{2\pi - 4\} C_k + o(C_k) \Rightarrow C_k^{-1} V_{\boldsymbol{\theta}}[N_{\mathcal{P}_{2,k}}] \approx 2\pi - 4. \end{aligned} \tag{3.4.11}$$

Similarly, Theorem 3.4.1 parts (ii) and (iii) combined provide the following results if $m_0 \geq 2$:

$$\begin{aligned} V_{\boldsymbol{\theta}}[N_{\mathcal{P}^{\#}_{1,k}}] &= \tfrac{2}{k-1} C_k + o(C_k) \Rightarrow C_k^{-1} V_{\boldsymbol{\theta}}[N_{\mathcal{P}^{\#}_{1,k}}] \approx \tfrac{2}{k-1}, \text{ and} \\ V_{\boldsymbol{\theta}}[N_{\mathcal{P}^{\#}_{2,k}}] &= \left\{\tfrac{2k\pi g_{1k}}{(k-1)} - 4\right\} C_k + o(C_k) \Rightarrow C_k^{-1} V_{\boldsymbol{\theta}}[N_{\mathcal{P}^{\#}_{2,k}}] \approx \left\{\tfrac{2k\pi g_{1k}}{(k-1)} - 4\right\}, \end{aligned} \tag{3.4.12}$$

with g_{1k} coming from (3.3.4).

Comparing these approximate (or asymptotic) expression of $C_k^{-1} V_{\boldsymbol{\theta}}[N_{\mathcal{P}^{\#}_{1,k}}]$ with that of $C_k^{-1}V_{\boldsymbol{\theta}}[N_{\mathcal{P}_{1,k}}]$, we observe that the stopping variable $N_{\mathcal{P}_{1,k}^{\#}}$ is tighter around C_k than $N_{\mathcal{P}_{1,k}}$. Next, $C_k^{-1}V_{\boldsymbol{\theta}}[N_{\mathcal{P}_{2,k}}] = 2\pi - 4 \approx 2.2832$, however, we exhibit some selected values of $C_k^{-1}V_{\boldsymbol{\theta}}[N_{\mathcal{P}_{2,k}^{\#}}]$:

k:	3	5	10	15	20	25	30
$\frac{2k\pi g_{1k}}{(k-1)} - 4$:	1.1019	0.55717	0.25056	0.16174	0.11942	9.4661×10^{-2}	7.8405×10^{-2}

We observe that the stopping variable $N_{\mathcal{P}_{2,k}^{\#}}$ is tighter around C_k than $N_{\mathcal{P}_{2,k}}$ (when $k > 2$). This is one concrete way how we rip the benefit of implementing $\mathcal{P}_{i,k}^{\#}$ over $\mathcal{P}_{i,k}$. More discussions will be forthcoming.

3.4.3. Simulation Studies on FWCI Problem

In this section, we investigate the performances of FWCI problems via computer simulations when the sample sizes are varied from small (50) to medium (150) to large (300).

3.4.3.1. Estimation of $D_{i,k}^{\#(1)}$ and $\eta_{i,k}^{\#(1)}$

We revisit the definitions and expressions of $D_{i,k}^{\#(1)}$ and $\eta_{i,k}^{\#(1)}$ for $i = 1, 2$ from (3.4.5) and (3.4.8). Again, the superscript (1) identifies with the FWCI problem. We recall that

$$D_{1,k}^{\#(1)} = \Sigma_{n=1}^{\infty} n^{-1} E\left\{\max(0, \Sigma_{i=1}^{n} W_i - 2n)\right\} \text{ and } D_{2,k}^{\#(1)} = \Sigma_{n=1}^{\infty} n^{-1} E\left\{\max(0, \Sigma_{i=1}^{n} W_i' - \tfrac{3}{2}n)\right\},$$

where W_i and W_i' come from (3.4.3) and (3.4.6).

Clearly the values of $D_{i,k}^{\#(1)}$ depend on the choice of group size, k. We used Matlab R2019b to estimate $D_{i,k}^{\#(1)}$. More specifically, consider $D_{1,k}^{\#(1)}$: With fixed k, we defined $Y_n = \max\left\{0, \Sigma_{i=1}^{n} W_i - 2n\right\}$ and consequently $D_{1,k}^{\#(1)} = \Sigma_{n=1}^{\infty} n^{-1} E\left[Y_n\right]$. To estimate $E\left[Y_n\right]$, having fixed n, we generated n independent random variables $\{W_1, ..., W_n\}$ from χ^2_{k-1} distribution, and computed $\max\{0, \Sigma_{i=1}^{n} W_i - 2n\}$. The associated sample mean, $\overline{Y}_n$, was calculated from 10000 replications. Then, we repeated this approach for every fixed $n = 1, 2, ..., 1000$ and estimated $D_{1,k}^{\#(1)}$ by $\Sigma_{n=1}^{\infty} n^{-1}\overline{Y}_n$.

Obviously, $D_{2,k}^{\#(1)}$ was estimated in the same fashion. Table 3.4 shows the estimated values for $D_{i,k}^{\#(1)}$, $\eta_{i,k}^{\#(1)}$ and the estimated standard error ($s_{\widehat{\eta}_{i,k}^{\#(1)}}$) of $\widehat{\eta}_{i,k}^{\#(1)}$ when $k = 3$ and 5.

Table 3.4. Estimated values of $D_{i,k}^{\#(1)}$, $\eta_{i,k}^{\#(1)}$, $s_{\widehat{\eta}_{i,k}^{\#(1)}}$

i	k	Strategy $\mathcal{P}_{i,k}^{\#}$	$\widehat{D}_{i,k}^{\#(1)}$	$\widehat{\eta}_{i,k}^{\#(1)}$	$s_{\widehat{\eta}_{i,k}^{\#(1)}}$
1	3	$\mathcal{P}_{1,3}^{\#}$	0.2573	−0.2573	1.91×10^{-3}
1	5	$\mathcal{P}_{1,5}^{\#}$	0.0790	0.1710	8.55×10^{-4}
2	3	$\mathcal{P}_{2,3}^{\#}$	0.1015	−0.2540	1.33×10^{-3}
2	5	$\mathcal{P}_{2,5}^{\#}$	0.0289	0.1637	6.22×10^{-4}

3.4.3.2. Simulating the Estimation Strategies from (3.3.13)

Next, we set out to compare the performances of the purely sequential FWCI estimation strategies defined from (3.3.13). Having fixed $\alpha = 0.05$, pseudo random samples were generated from a $N(\mu = 5, \sigma^2 = 4)$ population. Recall that m_0 and C_k were respectively the pilot number of groups and the optimal fixed number of groups consisting of k-tuples each. For example, if $m_0 = 3$ and $k = 5$, we begin sampling with 3 groups each having in it 5 observations. We fixed m_0 and C_k with selected k, and d was determined accordingly from the expression of C_k in (3.2.4). In Tables 3.6-3.9, we would use a the set of notation defined precisely in Table 3.5.

Table 3.5. The set of notations used in Tables 3.6-3.10 with $T = 10000$

n_i	: terminal sample size in i^{th} run;
$\bar{n} = T^{-1}\Sigma_{i=1}^{T} n_i$	: ave sample size, should compare with C_k;
$s_{\bar{n}} = \left\{(T^2 - T)^{-1}\Sigma_{i=1}^{T}(n_i - \bar{n})^2\right\}^{1/2}$	: estimated standard error (s.e.) of $\bar{n}$;
$\bar{x}_{n_i}$	: terminal sample mean in i^{th} run;
$\bar{x} = T^{-1}\Sigma_{i=1}^{T}\bar{x}_{n_i}$	: combined sample ave, should compare with μ;
$s_{\overline{x}} = \left\{(T^2 - T)^{-1}\Sigma_{i=1}^{T}(\bar{x}_{n_i} - \bar{x})^2\right\}^{1/2}$	: estimated s.e. of $\overline{x}$;
$p_i = I\left(\lvert\bar{x}_{n_i} - \mu\rvert \le d\right)$	: 1(or 0) if J_{n_i} covers (or does not cover) μ in i^{th} run;
$\bar{p} = T^{-1}\Sigma_{i=1}^{T} p_i$	: estimated cov probability, should compare with $1 - \alpha$;
$z \equiv \bar{p} - (1 - \alpha)$	: should compare with $\gamma_k^{\#} C_k^{-1}$;
$s_{\bar{p}} \equiv s_z = \left\{T^{-1}\bar{p}(1 - \bar{p})\right\}^{1/2}$	: estimated s.e. of $\bar{p}$;
$\bar{n}/C_k$	: should compare with 1;
$\overline{\eta}_k^{(1)} \equiv \bar{n} - C_k$	: should compare with $\hat{\eta}_k^{\#(1)}$;
p-value$_1$	: p-value for testing $E_\theta[N_{P_k} - C_k] \approx \eta_k^{\#(1)}$;
p-value$_2$	: p-value for testing $E_\theta[\overline{p} - (1 - \alpha)] \approx \gamma_k^{\#} C_k^{-1}$.

To be able to compare the performance of $P_{i,k}$ from (3.2.6) with $P_{i,k}^{\#}$ from (3.3.13), we implement simulations under the same distribution, the m_0's and C_k's and the number of replications ($T = 10000$), but replace the estimators used in $P_{i,k}$ with our newly proposed permuted estimators. Mukhopadhyay and Wang (2020) provided the simulation results for $k = 2$ and 3. Notice that when $k = 2$, our methodologies $P_{i,k}^{\#}$ defined from (3.3.13) will coincide with the methodologies $P_{i,k}$. Hence, we fixed $k = 3, 5$ in producing the sets of Tables 3.6-3.7 and Tables 3.8-3.9 respectively.

Table 3.6. Simulations for the FWCI problem under $\mu = 5, \sigma = 2$ with 10000 runs implementing $N_{\mathcal{P}^{\#}_{1,3}}$ from (3.3.13): $k = 3$, $\alpha = 0.05$, $\hat{\eta}^{\#(1)}_{1,3} = -0.2573$, and $P_{\boldsymbol{\theta}}\{\mu \in J_{N_{\mathcal{P}^{\#}_{1,3}}}\} = 1 - \alpha + \gamma^{\#}_{1,3} C_3^{-1} + o(C_3^{-1})$ from (3.4.10)

C_3 / $\gamma^{\#}_{1,3} C_3^{-1}$	m_0	$\overline{n}$ / $s_{\overline{n}}$	$\overline{n}/C_3$	$\bar{\eta}^{(1)}_{1,3} =$ $\overline{n} - C_3$	$\overline{x}$ / $s_{\overline{x}}$	$\overline{p}$ / $s_{\overline{p}}$	$z =$ $\overline{p} - 0.95$	p-value$_1$ / p-value$_2$
30	6	29.58	0.986	-0.42	4.999	0.9418	-0.0082	0.00
-0.0056		0.059			2.18×10^{-3}	2.34×10^{-3}		0.27
	16	29.74	0.991	-0.26	5.000	0.9423	-0.0077	1.00
		0.057			2.15×10^{-3}	2.33×10^{-3}		0.37
	24	30.21	1.007	0.21	5.001	0.9461	-0.0039	0.00
		0.048			2.12×10^{-3}	2.26×10^{-3}		0.45
80	6	79.78	0.997	-0.22	4.999	0.9514	0.0014	0.65
-0.0021		0.091			1.30×10^{-3}	2.15×10^{-3}		0.10
	16	79.63	0.995	-0.37	4.999	0.9499	-0.0001	0.22
		0.091			1.30×10^{-3}	2.18×10^{-3}		0.36
	24	79.90	0.996	-0.30	4.999	0.9493	-0.0007	0.67
		0.091			1.29×10^{-3}	2.19×10^{-3}		0.52
200	6	199.82	0.999	-0.18	5.000	0.9488	-0.0012	0.56
-0.0008		0.141			8.21×10^{-4}	2.20×10^{-3}		0.87
	16	199.86	0.999	-0.14	4.999	0.9473	-0.0027	0.39
		0.142			8.26×10^{-4}	2.23×10^{-3}		0.41
	24	199.70	0.999	-0.30	5.000	0.9451	-0.0049	0.78
		0.144			8.30×10^{-4}	2.28×10^{-3}		0.07

Table 3.7. Simulations for the FWCI problem under $\mu = 5, \sigma = 2$ with 10000 runs implementing $N_{\mathcal{P}^{\#}_{2,3}}$ from (3.3.13): $k = 3$, $\alpha = 0.05$, $\hat{\eta}^{\#(1)}_{2,3} = -0.2540$, and $P_{\boldsymbol{\theta}}\{\mu \in J_{N_{\mathcal{P}^{\#}_{2,3}}}\} = 1 - \alpha + \gamma^{\#}_{2,3}C_3^{-1} + o(C_3^{-1})$ from (3.4.10)

C_3 $\gamma^{\#}_{2,3}C_3^{-1}$	m_0	$\overline{n}$ $s_{\overline{n}}$	$\overline{n}/C_3$	$\overline{\eta}^{(1)}_{2,3} =$ $\overline{n} - C_3$	$\overline{x}$ $s_{\overline{x}}$	$\overline{p}$ $s_{\overline{p}}$	$z =$ $\overline{p} - 0.95$	p-value$_1$ p-value$_2$
30	6	29.55	0.984	-0.45	4.999	0.9427	-0.0073	0.00
-0.0022		0.059			2.18×10^{-3}	2.32×10^{-3}		0.03
	16	29.71	0.990	-0.29	4.999	0.9426	-0.0074	0.50
		0.057			2.17×10^{-3}	2.33×10^{-3}		0.03
	24	30.23	1.008	0.23	5.001	0.9465	-0.0035	0.00
		0.048			2.13×10^{-3}	2.25×10^{-3}		0.58
80	6	79.80	0.998	-0.20	4.999	0.9456	-0.0044	0.54
-0.0008		0.092			1.30×10^{-3}	2.27×10^{-3}		0.12
	16	79.59	0.995	-0.41	4.999	0.9502	0.0002	0.09
		0.091			1.29×10^{-3}	2.18×10^{-3}		0.63
	24	79.76	0.997	-0.24	4.999	0.9520	0.0020	0.85
		0.092			1.29×10^{-3}	2.14×10^{-3}		0.18
200	6	199.82	0.999	-0.18	5.000	0.9486	-0.0014	0.62
-0.0003		0.141			8.16×10^{-4}	2.21×10^{-3}		0.63
	16	199.71	0.999	-0.29	5.000	0.9491	-0.0009	0.81
		0.144			8.21×10^{-4}	2.20×10^{-3}		0.80
	24	199.67	0.998	-0.33	5.000	0.9467	-0.0033	0.58
		0.144			8.29×10^{-4}	2.25×10^{-3}		0.19

Each table shows the values of the estimated theoretical second-order values $\hat{\eta}^{\#(1)}_{i,k}$ in its heading. In column 1, we show the optimal fixed sample size C_k along with the theoretical second-order term $\gamma^{\#}_{i,k}C_k^{-1}$ from (3.4.10). Choices of pilot group number m_0 are shown in column 2. Columns 3-5 suggest that the average number of groups ($\overline{n}$) are close to the pre-assigned optimal fixed number of groups, C_k, with small standard errors, $s_{\overline{n}}$.

The first-order efficiency term, $\overline{n}/C_k$, shows values that are very close to 1 with minor oversampling at times. The second-order efficiency term, $\bar{\eta}_{i,k}^{(1)}(=\overline{n}-C_k)$, stay reasonably close to the estimated theoretical value ($\hat{\eta}_{i,k}^{\#(1)}$). To validate this result, recall that $\hat{\eta}_{i,k}^{\#(1)}$ values were provided along with their estimated s.e. values $s_{\hat{\eta}}$ in Table 3.4. We empirically tested:

$$H_0 : E_\theta[N_{\mathcal{P}_{i,k}^{\#}} - C_k] = \eta_{i,k}^{\#(1)} \text{ vs. } H_1 : E_\theta[N_{\mathcal{P}_{i,k}^{\#}} - C_k] \neq \eta_{i,k}^{\#(1)}, \tag{3.4.13}$$

Column 9 shows p-values (p-value$_1$) of such customary two-sample tests performed based upon the datasets on ($\overline{n}-C_k, s_{\overline{n}}$) and ($\hat{\eta}_{i,k}^{\#(1)}, s_{\hat{\eta}}$). Most of these p-values turned out larger than 0.05 which seemed to indicate that our simulations validated the second-order efficiency property reasonably well. However, a number of these p-values fell under 0.05 especially when $C_k = 25$. Such discrepancies may have occurred since $C_k = 25$ is deemed rather small. Column 7 shows that the overall terminal sample mean $\overline{x}$ came close to μ with small $s_{\overline{x}}$.

Columns 7-9 summarize the performance of estimated coverage probability: Column 7 shows that the estimated coverage probabilities $\overline{p}$ which are remarkably close to $1-\alpha$ along with small $s_{\overline{p}}$ values across all tables. In the spirit of a two-sample test highlighted by (3.4.13), column 8 explores practical validity of the estimated second-order term, $\overline{p}-(1-\alpha)$, the difference between the estimated coverage probability $\overline{p}$ and the set target $1-\alpha$. We also empirically tested:

$$H_0 : P_{\boldsymbol{\theta}} \left\{ \mu \in J_{N_{\mathcal{P}^{\#}_{i,k}}} \right\} - (1-\alpha) = \gamma^{\#}_{i,k} C_k^{-1} \text{ vs. } H_1 : P_{\boldsymbol{\theta}} \left\{ \mu \in J_{N_{\mathcal{P}^{\#}_{i,k}}} \right\} - (1-\alpha) \neq \gamma^{\#}_{i,k} C_k^{-1}, \tag{3.4.14}$$

The associated p-values (p-value$_2$) are shown in column 9. In Tables 3.6-3.7, we note that most of these p-values are larger than 0.05 which seem to indicate that our simulations validated the second-order expansion of our coverage probability reasonably well. However, a number of these p-values fell under 0.05 especially when $C_k = 25$. Such discrepancies may have occurred since $C_k = 25$ is deemed rather small.

Table 3.8. Simulations for the FWCI problem under $\mu = 5, \sigma = 2$ with 10000 runs implementing $N_{\mathcal{P}^{\#}_{1,5}}$ from (3.3.13): $k = 5$, $\alpha = 0.05$, $\hat{\eta}^{\#(1)}_{1,5} = 0.1710$, and $P_{\boldsymbol{\theta}}\{\mu \in J_{N_{\mathcal{P}^{\#}_{1,5}}}\} = 1 - \alpha + \gamma^{\#}_{1,5} C_5^{-1} + o(C_5^{-1})$ from (3.4.10)

C_5 $\gamma^{\#}_{1,5} C_5^{-1}$	m_0	$\overline{n}$ $s_{\overline{n}}$	$\overline{n}/C_5$	$\bar{\eta}^{(1)}_{1,5} =$ $\overline{n} - C_5$	$\overline{x}$ $s_{\overline{x}}$	$\overline{p}$ $s_{\overline{p}}$	z	p-value$_1$ p-value$_2$
30	6	30.13	1.004	0.13	4.997	0.9496	-0.0004	0.36
-1.66×10^{-3}		0.040			1.64×10^{-3}	2.19×10^{-3}		0.56
	16	30.15	1.005	0.15	4.999	0.9534	0.0034	0.53
		0.039			1.63×10^{-3}	2.11×10^{-3}		0.02
	24	30.28	1.009	0.23	4.999	0.9476	-0.0024	0.00
		0.037			1.63×10^{-3}	2.23×10^{-3}		0.74
80	6	80.20	1.003	0.20	4.999	0.9521	0.0021	0.64
-6.22×10^{-4}		0.063			9.97×10^{-4}	2.14×10^{-3}		0.20
	16	80.26	1.003	0.26	5.000	0.9504	0.0004	0.15
		0.064			1.00×10^{-3}	2.17×10^{-3}		0.64
	24	80.14	1.002	0.14	5.000	0.9524	0.0024	0.64
		0.063			9.95×10^{-4}	2.13×10^{-3}		0.16
200	6	200.32	1.002	0.32	5.000	0.9480	-0.0020	0.12
-2.49×10^{-4}		0.099			6.29×10^{-4}	2.22×10^{-3}		0.43
	16	200.20	1.001	0.20	4.999	0.9511	0.0011	0.74
		0.100			6.32×10^{-4}	2.16×10^{-3}		0.53
	24	200.07	1.000	0.007	5.000	0.9514	0.0014	0.30
		0.101			6.32×10^{-4}	2.15×10^{-3}		0.44

Table 3.9. Simulations for the FWCI problem under $\mu = 5, \sigma = 2$ with 10000 runs implementing $N_{\mathcal{P}^{\#}_{2,5}}$ from (3.3.13): $k = 5$, $\alpha = 0.05$, $\hat{\eta}^{\#(1)}_{2,5} = 0.1637$, and $P_{\boldsymbol{\theta}}\{\mu \in J_{N_{\mathcal{P}^{\#}_{2,5}}}\} = 1 - \alpha + \gamma^{\#}_{2,5} C_5^{-1} + o(C_5^{-1})$ from (3.4.10)

C_5 $\gamma^{\#}_{2,5} C_5^{-1}$	m_0	$\overline{n}$ $s_{\overline{n}}$	$\overline{n}/C_5$	$\bar{\eta}^{(1)}_{2,5} =$ $\overline{n} - C_5$	$\overline{x}$ $s_{\overline{x}}$	$\overline{p}$ $s_{\overline{p}}$	$z =$ $\overline{p} - 0.95$	p-value$_1$ p-value$_2$
30	6	30.08	1.003	0.08	4.997	0.9481	-0.0019	0.04
-1.87×10^{-5}		0.041			1.65×10^{-3}	2.22×10^{-3}		0.40
	16	30.12	1.004	0.12	4.999	0.9534	0.0034	0.33
		0.041			1.63×10^{-3}	2.11×10^{-3}		0.10
	24	30.27	1.009	0.27	4.999	0.9504	0.0004	0.00
		0.038			1.63×10^{-3}	2.17×10^{-3}		0.85
80	6	80.18	1.002	0.18	4.999	0.9518	0.0018	0.86
-7.00×10^{-6}		0.065			9.98×10^{-4}	2.14×10^{-3}		0.40
	16	80.21	1.003	0.21	5.000	0.9484	-0.0016	0.50
		0.065			1.10×10^{-3}	2.21×10^{-3}		0.47
	24	80.12	1.001	0.12	5.000	0.9523	0.0023	0.48
		0.066			9.96×10^{-4}	2.13×10^{-3}		0.28
200	6	200.31	1.002	0.31	5.000	0.9509	0.0009	0.17
-2.80×10^{-6}		0.104			6.27×10^{-4}	2.16×10^{-3}		0.68
	16	200.12	1.001	0.12	4.999	0.9486	-0.0014	0.70
		0.104			6.35×10^{-4}	2.21×10^{-3}		0.53
	24	200.00	1.000	0.00	5.000	0.9533	0.0033	0.13
		0.105			6.27×10^{-4}	2.11×10^{-3}		0.12

Tables 3.8-3.9 similarly show summaries from simulations when with $k = 5$ with $\alpha = 0.05$ fixed. Comparing Tables 3.6-3.7 with Tables 3.8-3.9, we observe that the standard error values ($s_{\overline{n}}$) of the average stopping variable ($\overline{n}$) decrease as the group size (k) increases from

3 to 5.

From Section 3.3.1, we reiterate that we expect the stopping variable $N_{P_{i,k}^{\#}}$ from (3.3.13) to be more tightly around C_k than the stopping variable $N_{P_{i,k}}$ from (3.2.6). In order to investigate this issue further, we provide Table 3.10 which shows side-by-side the $s_{\overline{n}}$ values associated with $N_{P_{i,k}^{\#}}, N_{P_{i,k}}$ when $k = 3, 5$ and $i = 1, 2$. Having fixed (i, k, m_0, C_k), these columns were obtained from 10000 independent replications based on pseudo random observation from a $N(\mu = 5, \sigma^2 = 4)$ population based upon freshly run simulations.

Table 3.10. Comparing $s_{\overline{n}}$ from $P_{i,k}$ and $P_{i,k}^{\#}$ for the FWCI problem under $\mu = 5, \sigma = 2$ when $k = 3, 5$ and $\alpha = 0.05$ with 10000 runs

C_k	m_0	$P_{1,3}$	$P_{1,3}^{\#}$	$P_{2,3}$	$P_{2,3}^{\#}$	$P_{1,5}$	$P_{1,5}^{\#}$	$P_{2,5}$	$P_{2,5}^{\#}$
30	6	0.099	0.059	0.102	0.059	0.091	0.040	0.096	0.041
	16	0.077	0.057	0.081	0.057	0.077	0.039	0.081	0.041
	24	0.060	0.048	0.063	0.048	0.060	0.037	0.063	0.038
80	6	0.149	0.091	0.160	0.092	0.137	0.063	0.149	0.065
	16	0.134	0.091	0.144	0.091	0.134	0.064	0.144	0.065
	24	0.135	0.091	0.143	0.092	0.131	0.063	0.142	0.066
120	6	0.202	0.141	0.225	0.141	0.205	0.099	0.219	0.104
	16	0.202	0.142	0.218	0.144	0.205	0.100	0.217	0.104
	24	0.201	0.144	0.214	0.144	0.204	0.101	0.217	0.105

Comparing column 3 (or 5) with column 7 (or 9), we get a distinct feeling that the $s_{\overline{n}}$ values are not changing a whole lot when we keep m_0, C_k fixed but k varies. This is consistent with (3.2.9), parts (ii)-(iii). We certainly see that the stopping variable $N_{P_{i,k}^{\#}}$ is tighter around C_k than the stopping variable $N_{P_{i,k}}$ whereas the closeness between $N_{P_{i,k}^{\#}}$ and C_k appears to become stronger as k goes up.

3.5. Purely Sequential MRPE Strategies Under Sampling In Groups Incorporating Within Group Permutation

In this section, we are going to update the asymptotic first-order and second-order properties of the purely sequential MRPE strategies $Q_{i,k}^{\#}$ from (3.3.14), $i = 1, 2$. We emphasize that $Q_{i,k}$ (or its updated version, $Q_{i,k}^{\#}$) from (3.2.18) (or from (3.3.14)) used $T_{k,n}^{(i)}$ (or its updated version, $T_{k,n}^{\#(i)}$) in the definition of the boundary crossing, $i = 1, 2$.

We expect $Q_{i,k}^{\#}$'s to perform better than $Q_{i,k}$'s in view of (i) $T_{k,n}^{(i)}, T_{k,n}^{\#(i)}$ are both unbiased and consistent estimators of σ^2, and (ii) $V_{\boldsymbol{\theta}}[T_{k,n}^{\#(i)}]$ is expected not to exceed $V_{\boldsymbol{\theta}}[T_{k,n}^{(i)}]$, $i = 1, 2$. We will soon explore where and how such intuitively expected additional benefit of implementing $Q_{i,k}^{\#}$ over $Q_{i,k}$ may manifest itself.

3.5.1. Preliminaries and Selected First-Order Properties

We begin by stating a number of interesting results (Lemmas 3.5.1-3.5.3) without giving details towards their justifications.

Lemma 3.5.1. *Under the stopping time $N_{Q_{i,k}^{\#}}$ defined in (3.3.14), for all fixed $\boldsymbol{\theta}, c_k, A, k$, and $m_0 (\geq 2)$, the statistic $T_{k,n}^{\#(i)}$ and $\overline{X}_{k,n}$ are distributed independently for each fixed $n \geq m_0$.*

Lemma 3.5.2. *Denote $G_{Q_{i,k}^{\#}} \equiv N_{Q_{i,k}^{\#}}^{1/2} \left(\overline{X}_{k,N_{Q_{i,k}^{\#}}} - \mu \right) \sigma^{-1}$ under the stopping time $Q_{i,k}^{\#}$ defined in (3.3.14), for all fixed $\boldsymbol{\theta}, c_k, A, k$, and $m_0 (\geq 2)$, we have:*

$$G_{Q_{i,k}^{\#}} \sim N(0, 1), \;\; E_{\boldsymbol{\theta}} \left[\overline{X}_{k,Q_{i,k}^{\#}} \right] = \mu \;\; \text{and} \;\; V_{\boldsymbol{\theta}} \left[\overline{X}_{k,Q_{i,k}^{\#}} \right] = k^{-1} \sigma^2 E_{\boldsymbol{\theta}} \left[N_{Q_{i,k}^{\#}}^{-1} \right],$$

$i = 1, 2$.

Lemma 3.5.3. *For the stopping time $N_{Q_{i,k}^{\#}}$ defined in (3.3.14), for all fixed $\boldsymbol{\theta}, A, k$, and*

$m_0(\geq 2)$, we have as $c_k \to 0$:

$$\begin{aligned} &(i) \quad N_{Q^{\#}_{i,k}}/n^*_k \overset{P_{\boldsymbol{\theta}}}{\to} 1; \text{ and} \\ &(ii) \quad E_{\boldsymbol{\theta}}\left[N_{Q^{\#}_{i,k}}/n^*_k\right] \to 1 \text{ [Asymptotic first-order efficiency]}; \end{aligned} \tag{3.5.1}$$

parallel to (3.2.22) where $n^*_k(=\frac{1}{k}\sqrt{A/c_k}\sigma)$ defined in (3.2.16) is the optimal fixed number of groups.

Upon termination, we propose to estimate μ by $\overline{X}_{k,N_{Q^{\#}_{i,k}}}$ obtained from the fully accrued dataset, namely,

$$\left\{N_{Q^{\#}_{i,k}}, (X_{i1}, X_{i2}, ..., X_{ik}), i = 1, 2, ..., N_{Q^{\#}_{i,k}}\right\}.$$

Theorem 3.5.1. *For the stopping rule $Q^{\#}_{i,k}$ defined in (3.2.14), for all fixed $\boldsymbol{\theta}, A$, and k, we have as $c_k \to 0$:*

$$\xi_{Q^{\#}_{i,k}}(c_k) \to 1 \quad \text{[Asymptotic risk efficiency]}, \tag{3.5.2}$$

if $m_0 \geq 3$ with n^*_k and $\xi_{Q^{\#}_{i,k}}(c_k)$ coming from (3.2.16) and (3.2.21) respectively.

Proof: Clearly, in the spirit of (3.2.21), we have:

$$\xi_{Q^{\#}_{i,k}}(c_k) \equiv R_{k,N_{Q^{\#}_{i,k}}}(c_k)/R_{k,n^*_k} = \tfrac{1}{2}E_{\boldsymbol{\theta}}[N_{Q^{\#}_{i,k}}n^{*-1}_k] + \tfrac{1}{2}E_{\boldsymbol{\theta}}[n^*_k N^{-1}_{Q^{\#}_{i,k}}].$$

We can verify that $E_{\boldsymbol{\theta}}[n^*_k N^{-1}_{Q^{\#}_{i,k}}] \to 1$ as $c_k \to 0$, if $m_0 \geq 3$ and from Lemma 3.5.3, part (ii), we already know that $E_{\boldsymbol{\theta}}[N_{Q^{\#}_{i,k}}n^{*-1}_k] \to 1$ as $c_k \to 0$ if $m_0 \geq 2$. Hence, the result follows. ■

3.5.2. Selected Asymptotic Second-Order Properties

In this section, Theorems 3.5.2 - 3.5.3 summarize a number of major asymptotic second-order properties associated with the stopping rules $Q^{\#}_{i,k}$ from (3.3.14). However, before we tackle these theorems, we begin with a number of essential technical details.

Case 1: $i = 1$

First, we rewrite the stopping time corresponding to $Q^{\#}_{1,k}$ from (3.3.14) as follows:

$$N_{Q^{\#}_{1,k}} = \inf\left\{n \geq m_0 : \quad n^3 n_k^{*-2}\left(1 - \tfrac{1}{2(k-1)}n^{-1} + O(n^{-2})\right) \geq \Sigma_{i=1}^{n} W_i\right\}, \tag{3.5.3}$$

where $(k-1)W_i \overset{\text{i.i.d.}}{\sim} \chi^2_{k-1}$. Then, obviously, $E_{\boldsymbol{\theta}}\left[W_i\right] = 1$ and $V_{\boldsymbol{\theta}}\left[W_i\right] = \frac{2}{k-1}$. Next, for arbitrary $u(>0)$, we obtain (with $\widetilde{u} \propto u$):

$$P_{\boldsymbol{\theta}}\left\{W_1 \leq u\right\} = P_{\boldsymbol{\theta}}\left\{S_i^2 \leq \widetilde{u}\right\} \propto \int_0^{\widetilde{u}} y^{\frac{k-3}{2}} e^{-y/2} dy \leq B u^{(k-1)2},$$

with $B(>0)$, not involving u.

Now, we compare this representation (3.5.3) with that used in nonlinear renewal theory (Mukhopadhyay and Solanky 1994, p. 49), we immediately have:

$$\lambda = 1, \tau^2 = \tfrac{2}{k-1}, h^* = n_k^{*-2}, n^* = n_k^*, \delta = 3, \beta^* = \tfrac{1}{2}, l_0 = -\tfrac{1}{2(k-1)}, p = \tfrac{1}{2(k-1)}, \text{ and } b = \tfrac{1}{2}(k-1). \tag{3.5.4}$$

We again apply nonlinear renewal theory to express:

$$\begin{aligned} D^{\#(2)}_{1,k} &= \Sigma_{n=1}^{\infty} n^{-1} E\left\{\max(0, \Sigma_{i=1}^{n} W_i - 3n)\right\}; \\ \nu^{\#(2)}_{1,k} &= 1 + \tfrac{1}{2(k-1)} - D^{\#(2)}_{1,k}; \\ \eta^{\#(2)}_{1,k} &= \tfrac{1}{2} - \tfrac{1}{4(k-1)} - \tfrac{1}{2} D^{\#(2)}_{1,k}. \end{aligned} \tag{3.5.5}$$

These expressions correspond to $T_{k,n}^{\#(1)}$. The superscript (2) used in (3.5.5) identifies with the MRPE problem.

Case 2: $i = 2$

We rewrite the stopping time corresponding to $Q_{2,k}^{\#}$ from (3.3.14) as follows:

$$N_{Q_{2,k}^{\#}} = \inf\left\{n \geq m_0 : \quad n^2 n_k^{*-1} \geq \Sigma_{i=1}^{n} W_i'\right\}, \tag{3.5.6}$$

where $W_i' = \sqrt{\frac{\pi k}{2(k-1)}}\frac{M_i}{\sigma}$. Then, $E_{\boldsymbol{\theta}}\left[W_i'\right] = 1$, $V_{\boldsymbol{\theta}}\left[W_i'\right] = \frac{k\pi g_{1k}}{2(k-1)} - 1$, with g_{1k} defined in (3.3.4). Next, for arbitrary $u(> 0)$, we obtain (with $\widetilde{u} \propto u$):

$$P_{\boldsymbol{\theta}}\left\{W_1' \leq u\right\} = P_{\boldsymbol{\theta}}\left\{\tfrac{1}{k}\Sigma_{j=1}^{k}|X_j - \overline{X}_i| \leq \widetilde{u}\right\} \leq kP\left\{|X_1 - \overline{X}_i| \leq \widetilde{u}\right\} \leq Bu,$$

with $B(> 0)$, not involving u.

Now, by comparing this representation with that used in nonlinear renewal theory (Mukhopadhyay and Solanky 1994, p.49), we immediately get:

$$\lambda = 1, \tau^2 = \tfrac{k\pi g_{1k}}{2(k-1)} - 1, h^* = n_k^{*-1}, n^* = n_k^*, \delta = 2, \beta^* = 1, l_0 = 0, p = \tfrac{k\pi g_{1k}}{2(k-1)} - 1, \text{ and } b = 1. \tag{3.5.7}$$

We again apply nonlinear renewal theory to express:

$$\begin{aligned} D_{2,k}^{\#(2)} &= \Sigma_{n=1}^{\infty} n^{-1} E\left\{\max(0, \Sigma_{i=1}^{n} W_i' - 2n)\right\}; \\ \nu_{2,k}^{\#(2)} &= \tfrac{1}{2}(1 + \tau^2) - D_{2,k}^{\#(2)}; \\ \eta_{2,k}^{\#(2)} &= \tfrac{1}{2}(1 - \tau^2) - D_{2,k}^{\#(2)}. \end{aligned} \tag{3.5.8}$$

These expressions correspond to $T_{k,n}^{\#(2)}$. The superscript (2) used in (3.5.8) identifies with the MRPE problem. We note that all three expressions depend on the choice of k. In Section 3.5.3.1, we will provide simulated estimates of the entities from both (3.5.5) and (3.5.8).

Theorem 3.5.2. *For the stopping time $N_{Q_{i,k}^{\#}}$ defined by (3.3.14), we denote $H_{i,k}^{\#(2)} = n_k^{*-1/2}(N_{Q_{i,k}^{\#}} - n_k^*)$. Then, for every fixed $\boldsymbol{\theta}, c_k$ and d, we have the following results as $c_k \to 0$:*

$$
\begin{array}{ll}
(i) & \left.\begin{array}{l}(a)\ P_{\boldsymbol{\theta}}\{N_{Q_{1,k}^{\#}} \le \varepsilon n_k^*\} = O(n_k^{*-(k-1)m_0/2}) \\ (b)\ P_{\boldsymbol{\theta}}\{N_{Q_{2,k}^{\#}} \le \varepsilon n_k^*\} = O(n_k^{*-m_0})\end{array}\right\} \text{ if } 0 < \varepsilon < 1 \text{ and } m_0 \ge 2; \\
(ii) & \left|H_{i,k}^{\#(2)}\right|^2 \text{ is uniformly integrable if } m_0 \ge 2; \\
(iii) & \left.\begin{array}{l}(a)\ H_{1,k}^{\#(2)} \xrightarrow{\pounds} N(0, \frac{1}{2(k-1)}) \\ (b)\ H_{2,k}^{\#(2)} \xrightarrow{\pounds} N(0, \frac{k\pi g_{1k}}{2(k-1)} - 1)\end{array}\right\} \text{ if } m_0 \ge 2; \\
(iv) & \left.\begin{array}{l}(a)\ E_\theta[N_{Q_{1,k}^{\#}} - n_k^*] = \eta_{1,k}^{\#(2)} + o(1) \\ (b)\ E_\theta[N_{Q_{2,k}^{\#}} - n_k^*] = \eta_{2,k}^{\#(2)} + o(1)\end{array}\right\} \text{if } m_0 \ge 2;
\end{array}
\quad (3.5.9)
$$

with $\eta_{1,k}^{\#(2)}, \eta_{2,k}^{\#(2)}$ and g_{1k} coming from (3.5.5), (3.5.8) and (3.3.4) respectively.

Theorem 3.5.3. *Given the regret function defined in (3.2.21), we have as $c_k \to 0$:*

$$
\omega_{Q_{i,k}^{\#}}(c_k) \equiv R_{k,N_{Q_{i,k}^{\#}}}(c_k) - R_{n_k^*} = kc_k p_{i,k}^{\#} + o(c_k), \quad (3.5.10)
$$

with $p_{1,k}^{\#} = \frac{1}{2(k-1)}$ and $p_{2,k}^{\#} = \frac{k\pi g_{1k}}{2(k-1)} - 1$ if $m_0 \ge 4$. Here, g_{1k} comes from (3.3.4).

Proof: We revisit the assumptions (B1)-(B5) from (3.2.24) in Section 3.2.2.1 and they hold for the sampling strategy defined by (3.3.14) along with the following details:

$$i=1: \quad a_1=\tfrac{k-1}{2}, a_2=0, a_3=\tfrac{1}{2(k-1)}, a_4=1, a_5=\eta_{1,k}^{\#(2)}, a_6=1;$$
$$i=1: \quad a_1=1, a_2=0, a_3=\tfrac{k\pi g_{1k}}{2(k-1)}-1, a_4=1, a_5=\eta_{2,k}^{\#(2)}, a_6=1.$$

This result follows immediately from the more general statement given in Theorem 3.2.2. ■

Once we combine (3.2.23) parts (ii) and (iii), we conclude the following results if $m_0 \geq 2$:

$$\begin{aligned} V_\theta\left[N_{Q_{1,k}}\right] &= \tfrac{1}{2}n_k^* + o(n_k^*) \Rightarrow n_k^{*-1}V_\theta\left[N_{Q_{1,k}}\right] \approx \tfrac{1}{2}, \\ V_\theta\left[N_{Q_{2,k}}\right] &= \left\{\tfrac{\pi}{2}-1\right\}n_k^* + o(n_k^*) \Rightarrow n_k^{*-1}V_\theta\left[N_{Q_{2,k}}\right] \approx \tfrac{\pi}{2}-1. \end{aligned} \tag{3.5.11}$$

Theorem 3.5.2 part (ii) and (iii) provide the following results for $m_0 \geq 2$:

$$\begin{aligned} V_\theta\left[N_{Q^{\#}_{1,k}}\right] &= \tfrac{1}{2(k-1)}n_k^* + o(n_k^*) \Rightarrow n_k^{*-1}V_\theta\left[N_{Q^{\#}_{1,k}}\right] \approx \tfrac{1}{2(k-1)}, \text{ and} \\ V_\theta\left[N_{Q^{\#}_{2,k}}\right] &= \left\{\tfrac{k\pi g_{1k}}{2(k-1)}-1\right\}n_k^* + o(n_k^*) \Rightarrow n_k^{*-1}V_\theta\left[N_{Q^{\#}_{2,k}}\right] \approx \tfrac{k\pi g_{1k}}{2(k-1)}-1, \end{aligned} \tag{3.5.12}$$

with g_{1k} coming from (3.3.4).

Comparing these approximate (or asymptotic) expression of $n_k^{*-1}V_\theta\left[N_{Q^{\#}_{1,k}}\right]$ with that of $n_k^{*-1}V_\theta\left[N_{Q_{1,k}}\right]$, we observe that the stopping time $N_{Q^{\#}_{1,k}}$ is tighter around n_k^* than $N_{Q_{1,k}}$. Next, $n_k^{*-1}V_\theta\left[N_{Q_{2,k}}\right] = \frac{\pi}{2}-1 \approx 0.5708$, we exhibit some selected values of $n_k^{*-1}V_\theta\left[N_{Q^{\#}_{2,k}}\right]$:

k:	3	5	10	15	20	25	30
$\frac{k\pi g_{1k}}{2(k-1)}-1$:	0.2754	0.1393	0.0626	0.0404	0.0299	0.0237	0.0196

In other words, we observe that the stopping time $N_{Q^{\#}_{2,k}}$ is tighter around n_k^* than $N_{Q_{2,k}}$

(when $k > 2$). This is one concrete way how we rip the benefit of implementing $Q^{\#}_{i,k}$ over $Q_{i,k}$. More discussions will be forthcoming.

3.5.3. Simulation Studies on MRPE Problem

In this section, we investigate the performances of MRPE problems via computer simulations when the sample sizes are varied from small (50) to medium (150) to large (300).

3.5.3.1. Estimation of $D^{\#(2)}_{i,k}$ and $\eta^{\#(2)}_{i,k}$

We revisit the definitions and expressions of $D^{\#(2)}_{i,k}$ and $\eta^{\#(2)}_{i,k}$ for $i = 1, 2$ from (3.5.5) and (3.5.8). Again, the superscript (2) corresponds to MRPE problem. We recall that

$$D^{\#(2)}_{1,k} = \Sigma^{\infty}_{n=1} n^{-1} E\left\{\max(0, \Sigma^{n}_{i=1} W_i - 3n)\right\} \text{ and } D^{\#(2)}_{2,k} = \Sigma^{\infty}_{n=1} n^{-1} E\left\{\max(0, \Sigma^{n}_{i=1} W'_i - 2n)\right\},$$

where W_i and W'_i are defined after (3.5.3) and (3.5.6).

Clearly the values of $D^{\#(2)}_{i,k}$ depend on the choice of group size, k. We used Matlab R2019b to estimate $D^{\#(2)}_{i,k}$ in the same fashion that we had described in Section 3.4.3.1. Table 3.11 shows the estimated values for $D^{\#(2)}_{i,k}$, $\eta^{\#(2)}_{i,k}$ and the estimated standard error ($s_{\widehat{\eta}^{\#(2)}_{i,k}}$) of $\widehat{\eta}^{\#(2)}_{i,k}$ when $k = 3$ and 5.

Table 3.11. Estimated values of $D^{\#(2)}_{i,k}$, $\eta^{\#(2)}_{i,k}$, $s_{\widehat{\eta}^{\#(2)}_{i,k}}$

i	k	Strategy $Q^{\#}_{i,k}$	$D^{\#(2)}_{i,k}$	$\eta^{\#(2)}_{i,k}$	$s_{\widehat{\eta}^{\#(2)}_{i,k}}$
1	3	$Q^{\#}_{1,3}$	0.0634	0.3433	5.38×10^{-4}
1	5	$Q^{\#}_{1,5}$	0.0110	0.4320	1.72×10^{-4}
2	3	$Q^{\#}_{2,3}$	0.0130	0.3493	2.57×10^{-4}
2	5	$Q^{\#}_{2,5}$	0.0017	0.4289	7.47×10^{-4}

3.5.3.2. Simulating the Estimation Strategies from (3.3.14)

In this section, we set out to compare the performances of the purely sequential MRPE strategies defined (3.3.14). We fixed the choices of A, m_0 and k, but then took $n_k^* = 25, 180, 350$ so that (3.2.16) led to associated values of c_k. The MRPE strategies from (3.3.14) were then implemented one by one with $T = 10000$ independent runs under each configuration. In Tables 3.13-3.17, we would use the set of notation defined precisely in Table 3.12 in addition to some that may be borrowed from Table 3.5.

Table 3.12. Additional set of notations used in Tables 3.13-3.17 with $T = 10000$ beyond those in Table 3.5

$r_i = \frac{A\sigma^2}{kn_i} + kc_kn_i$	: estimated risk in i^{th} run;
$\overline{r} = T^{-1}\Sigma_{i=1}^{T} r_i$	: estimated risk, should compare with R_{k,n_k^*};
$\overline{\xi} = \overline{r}/R_{k,n_k^*}$	: should compare with 1;
$s_{\overline{\xi}} = \frac{1}{R_{k,n_k^*}}\left\{(T^2-T)^{-1}\Sigma_{i=1}^{T}(r_i-\overline{r})^2\right\}^{1/2}$	: estimated s.e. of $\overline{\xi}$;
$\omega_i \equiv kc_k\frac{(n_i-n_k^*)^2}{n_i}$	: estimated regret in i^{th} run;
$\overline{\omega} = T^{-1}\Sigma_{i=1}^{T}\omega_i$	: estimated regret, should compare with kc_kq;
$s_{\overline{\omega}} = \left\{(T^2-T)^{-1}\Sigma_{i=1}^{T}(\omega_i-\overline{\omega})^2\right\}^{1/2}$	: estimated s.e. of $\overline{\omega}$;
p-value$_1$	: p-value for testing $E_\theta[N_{Q_k} - n_k^*] \approx \hat{\eta}_k^{\#(2)}$;
p-value$_2$	: p-value for testing $\omega \approx kc_kq_k^{\#}$.

To be able to compare the performance of $Q_{i,k}$ from (3.2.18) with $Q_{i,k}^{\#}$ from (3.3.14), we implement simulations under the same distribution, the m_0's and n_k^*'s, and the number of replications ($T = 10000$). Notice when $k = 2$, our methodologies $Q_{i,k}^{\#}$ defined by (3.3.14) are identical with $Q_{i,k}$ and hence in Tables 3.13-3.14 (or Tables 3.15-3.16) implemented $Q_{i,k}^{\#}$ with $k = 3$ (or $k = 5$). Each table also provides the values of the corresponding theoretical second-order values $\hat{\eta}_{i,k}^{\#(2)}$ in the table's heading and the second-order term $kc_kp_{i,k}^{\#}$ for the regret expansion from Theorem 3.5.3 in column 1.

Table 3.13. Simulations for the MRPE problem under $\mu = 5, \sigma = 2, A = 1$ with 10000 runs implementing $N_{Q_{1,3}^{\#}}$ from (3.3.14): $k = 3$, $\hat{\eta}_{1,3}^{\#(2)} = 0.3433$, and $\omega = kc_kp_{i,k}^{\#} + o(c_k)$ from (3.5.10)

n_3^* $3c_3p_{1,3}^{\#}$	m_0	$\overline{n}$ $s_{\overline{n}}$	$\overline{n}/n_3^*$	$\overline{\eta}_{1,3}^{(2)} =$ $\overline{n} - n_3^*$	$\overline{x}$ $s_{\overline{x}}$	$\overline{\xi}$ $s_{\overline{\xi}}$	$\overline{\omega}$ $s_{\overline{\omega}}$	p-value$_1$ p-value$_2$
25	6	25.34	1.013	0.34	5.000	1.006	5.97×10^{-4}	0.76
5.33×10^{-4}		0.026			2.31×10^{-3}	1.01×10^{-4}	1.08×10^{-5}	0.00
	12	25.34	1.013	0.34	5.000	1.006	5.94×10^{-4}	0.90
		0.026			2.29×10^{-3}	9.24×10^{-5}	9.86×10^{-6}	0.00
	18	25.31	1.013	0.31	4.997	1.005	5.45×10^{-4}	0.24
		0.026			2.32×10^{-3}	7.54×10^{-5}	8.05×10^{-6}	0.14
180	6	180.35	1.002	0.35	5.000	1.001	1.05×10^{-5}	0.88
1.03×10^{-5}		0.068			8.48×10^{-4}	1.00×10^{-5}	1.48×10^{-7}	0.19
	12	180.41	1.002	0.41	5.000	1.001	1.01×10^{-5}	0.28
		0.066			8.53×10^{-4}	9.69×10^{-6}	1.43×10^{-7}	0.18
	18	180.36	1.002	0.36	5.000	1.001	1.03×10^{-5}	0.74
		0.067			8.51×10^{-4}	9.92×10^{-6}	1.47×10^{-7}	0.95
350	6	350.33	1.001	0.33	5.000	1.001	2.69×10^{-6}	0.89
2.72×10^{-6}		0.093			6.15×10^{-4}	5.08×10^{-6}	3.87×10^{-8}	0.46
	12	350.48	1.001	0.48	5.000	1.000	2.74×10^{-6}	0.16
		0.094			6.15×10^{-4}	5.07×10^{-6}	3.86×10^{-8}	0.59
	18	350.40	1.001	0.40	5.000	1.000	2.75×10^{-6}	0.56
		0.094			6.19×10^{-4}	5.14×10^{-6}	3.91×10^{-8}	0.40

Table 3.14. Simulations for the MRPE problem under $\mu = 5, \sigma = 2, A = 1$ with 10000 runs implementing $N_{Q^{\#}_{2,3}}$ from (3.3.14): $k = 3$, $\hat{\eta}^{\#(2)}_{2,3} = 0.3493$, and $\omega = kc_k p^{\#}_{i,k} + o(c_k)$ from (3.5.10)

n_3^*	m_0	$\overline{n}$	$\overline{n}/n_3^*$	$\bar{\eta}_3^{(2)} =$	$\overline{x}$	$\overline{\xi}$	$\overline{\omega}$	p-value$_1$
$3c_3p^{\#}_{2,3}$		$s_{\overline{n}}$		$\overline{n} - n_3^*$	$s_{\overline{x}}$	$s_{\overline{\xi}}$	$s_{\overline{\omega}}$	p-value$_2$
25	6	25.34	1.013	0.34	5.000	1.006	6.54×10^{-4}	0.79
5.88×10^{-4}		0.027			2.29×10^{-3}	1.04×10^{-4}	1.11×10^{-5}	0.00
	12	25.35	1.014	0.35	5.000	1.006	6.45×10^{-4}	0.95
		0.027			2.29×10^{-3}	1.04×10^{-4}	1.10×10^{-5}	0.00
	18	25.34	1.014	0.34	4.997	1.006	6.02×10^{-4}	0.94
		0.026			2.32×10^{-3}	8.15×10^{-5}	8.69×10^{-6}	0.10
180	6	180.32	1.002	0.32	5.000	1.001	1.13×10^{-5}	0.72
1.13×10^{-5}		0.070			8.56×10^{-4}	1.11×10^{-5}	1.64×10^{-7}	1.00
	12	180.41	1.002	0.41	5.000	1.001	1.11×10^{-5}	0.36
		0.070			8.62×10^{-4}	1.08×10^{-5}	1.60×10^{-7}	0.12
	18	180.49	1.003	0.49	5.000	1.001	1.16×10^{-5}	0.05
		0.071			8.59×10^{-4}	1.11×10^{-5}	1.64×10^{-7}	0.13
350	6	350.30	1.001	0.30	5.000	1.000	3.00×10^{-6}	0.61
2.99×10^{-6}		0.098			6.14×10^{-4}	5.62×10^{-6}	4.28×10^{-8}	0.97
	12	350.44	1.001	0.44	5.000	1.000	2.95×10^{-6}	0.34
		0.097			6.13×10^{-4}	5.50×10^{-6}	4.19×10^{-8}	0.24
	18	350.48	1.001	0.48	5.000	1.000	2.99×10^{-6}	0.19
		0.098			6.17×10^{-4}	5.42×10^{-6}	4.13×10^{-8}	0.86

Table 3.15. Simulations for the MRPE problem under $\mu = 5, \sigma = 2, A = 1$ with 10000 runs implementing $N_{Q^{\#}_{1,5}}$ from (3.3.14): $k = 5, \hat{\eta}^{\#(2)}_{1,5} = 0.4320$, and $\omega = kc_k p^{\#}_{i,k} + o(c_k)$ from (3.5.10)

n^*_5 $5c_5p^{\#}_{1,5}$	m_0	$\overline{n}$ $s_{\overline{n}}$	$\overline{n}/n^*_5$	$\bar{\eta}^{(2)}_5 =$ $\overline{n} - n^*_4$	$\overline{x}$ $s_{\overline{x}}$	$\overline{\xi}$ $s_{\overline{\xi}}$	$\overline{\omega}$ $s_{\overline{\omega}}$	p-value$_1$ p-value$_2$
25	6	25.43	1.017	0.43	4.998	1.003	1.72×10^{-4}	0.94
1.60×10^{-4}		0.018			1.76×10^{-3}	3.83×10^{-5}	2.45×10^{-6}	0.00
	12	25.43	1.017	0.43	4.999	1.003	1.70×10^{-4}	0.97
		0.018			1.77×10^{-3}	3.78×10^{-5}	2.42×10^{-6}	0.00
	18	25.43	1.017	0.43	5.000	1.003	1.72×10^{-4}	0.96
		0.018			1.77×10^{-3}	3.88×10^{-5}	2.49×10^{-6}	0.00
180	6	180.44	1.002	0.44	5.000	1.000	3.13×10^{-6}	0.81
3.09×10^{-6}		0.048			6.64×10^{-4}	4.94×10^{-6}	4.39×10^{-8}	0.28
	12	180.52	1.003	0.52	5.000	1.000	3.15×10^{-6}	0.06
		0.047			6.66×10^{-4}	5.11×10^{-6}	4.54×10^{-8}	0.18
	18	180.42	1.002	0.42	5.000	1.000	3.13×10^{-6}	0.74
		0.048			6.68×10^{-4}	4.95×10^{-6}	4.40×10^{-8}	0.30
350	6	350.52	1.001	0.52	5.000	1.000	8.11×10^{-7}	0.19
8.16×10^{-7}		0.066			4.76×10^{-4}	2.50×10^{-6}	1.14×10^{-8}	0.61
	12	350.46	1.001	0.46	5.000	1.000	8.28×10^{-7}	0.67
		0.067			4.76×10^{-4}	2.54×10^{-6}	1.16×10^{-8}	0.30
	18	350.40	1.001	0.40	5.000	1.000	8.28×10^{-7}	0.63
		0.067			4.81×10^{-4}	2.56×10^{-6}	1.18×10^{-8}	0.34

Table 3.16. Simulations for the MRPE problem under $\mu = 5, \sigma = 2, A = 1$ with 10000 runs implementing $N_{Q^{\#}_{2,5}}$ from (3.3.14): $k = 5$, $\hat{\eta}^{\#(2)}_{2,5} = 0.4289$, and $\omega = kc_k p^{\#}_{i,k} + o(c_k)$ from (3.5.10)

n^*_5 $5c_5 p^{\#}_{2,5}$	m_0	$\overline{n}$ $s_{\overline{n}}$	$\overline{n}/n^*_5$	$\bar{\eta}^{(2)}_{2,5} =$ $\overline{n} - n^*_5$	$\overline{x}$ $s_{\overline{x}}$	$\overline{\xi}$ $s_{\overline{\xi}}$	$\overline{\omega}$ $s_{\overline{\omega}}$	p-value$_1$ p-value$_2$
25	6	25.43	1.017	0.43	4.998	1.003	1.91×10^{-4}	0.76
1.78×10^{-4}		0.019			1.77×10^{-3}	4.21×10^{-5}	2.70×10^{-6}	0.00
	12	25.42	1.017	0.42	4.999	1.003	1.91×10^{-4}	0.57
		0.019			1.17×10^{-3}	4.30×10^{-5}	2.75×10^{-6}	0.00
	18	25.44	1.017	0.44	5.000	1.003	1.91×10^{-6}	0.91
		0.019			1.78×10^{-3}	4.23×10^{-5}	2.70×10^{-6}	0.00
180	6	180.45	1.003	0.45	5.000	1.000	3.37×10^{-6}	0.71
3.44×10^{-6}		0.050			6.67×10^{-4}	5.46×10^{-6}	4.85×10^{-8}	0.18
	12	180.51	1.003	0.51	5.000	1.000	3.53×10^{-6}	0.14
		0.051			6.61×10^{-4}	5.61×10^{-6}	4.99×10^{-8}	0.06
	18	180.46	1.003	0.46	5.000	1.000	3.44×10^{-6}	0.64
		0.050			6.69×10^{-4}	5.47×10^{-6}	4.87×10^{-8}	0.93
350	6	350.54	1.002	0.54	5.000	1.001	8.87×10^{-7}	0.13
9.10×10^{-7}		0.069			4.78×10^{-4}	2.77×10^{-6}	1.27×10^{-8}	0.07
	12	350.49	1.001	0.49	5.000	1.000	9.09×10^{-7}	0.41
		0.070			4.74×10^{-4}	2.82×10^{-6}	1.29×10^{-8}	0.97
	18	350.39	1.001	0.39	5.000	1.000	9.30×10^{-7}	0.53
		0.071			4.83×10^{-4}	2.88×10^{-6}	1.31×10^{-8}	0.99

We gave extensive sets of comments on data analysis in Section 3.4.3.2, and hence we refrain from too much more explanations in this section on the MRPE problem, especially since our sentiments expressed earlier carry over in this section with obvious modifications.

However, in order to compare the variations, Table 3.17, constructed in the spirits of Table 3.10, shows $s_{\overline{n}}$ values obtained from simulations under both $Q^{\#}_{i,k}$ as well as $Q_{i,k}$, $k = 3, 5$. As expected, the stopping variable $N_{Q^{\#}_{i,k}}$ appears tighter around n^*_k than $N_{Q_{i,k}}$.

Table 3.17. Comparison of $s_{\overline{n}}$ from $Q_{i,k}$ and $Q^{\#}_{i,k}$ for the MRPE problem under $\mu = 5, \sigma = 2, A = 1$ when $k = 3, 5$ with 10000 runs

n^*_k	m_0	$Q_{1,3}$	$Q^{\#}_{1,3}$	$Q_{2,3}$	$Q^{\#}_{2,3}$	$Q_{1,5}$	$Q^{\#}_{1,5}$	$Q_{2,5}$	$Q^{\#}_{2,5}$
25	6	0.038	0.026	0.041	0.027	0.039	0.018	0.040	0.019
	12	0.038	0.026	0.040	0.027	0.037	0.018	0.040	0.019
	18	0.035	0.026	0.038	0.026	0.035	0.018	0.037	0.019
180	6	0.095	0.068	0.101	0.070	0.096	0.048	0.101	0.050
	12	0.096	0.066	0.101	0.070	0.096	0.047	0.102	0.051
	18	0.095	0.067	0.102	0.071	0.096	0.048	0.102	0.050
350	6	0.132	0.093	0.142	0.098	0.133	0.066	0.143	0.069
	12	0.131	0.094	0.141	0.097	0.132	0.067	0.143	0.070
	18	0.131	0.094	0.139	0.098	0.132	0.067	0.141	0.071

3.6. Illustrations With Wind Energy Data

Wind power delivers sustainable and renewable energy. Offshore wind energy research uses wind farms constructed in the ocean to harvest wind energy in order to generate electricity. Higher wind speed occurs offshore compared to wind speed on land. Offshore wind can be stronger in the afternoon when consumers demand more electricity. Recent interest in generating clean energy from offshore wind has increased demand for reliable statistical inference methodologies.

Block Island Wind Farm is the first commercial offshore wind farm in the United States. Wind power data collected by turbine are provided by the *National Renewable Energy Lab-*

oratory (NREL). We referred to King et al. (2014), Lieberman-Cribbin et al. (2014), Draxl et al. (2015a,b) and other sources including the following open source:.

https://www.nrel.gov/wind/data-tools.html

This site explains: "... OpenFAST is an open-source wind turbine simulation tool that was established with the FAST v8 code as its starting point. The release of OpenFAST represents a transition to better support an open-source developer community across research laboratories, industry, and academia around FAST-based aero-hydro-servo-elastic engineering models of wind-turbines and wind-plants. ... Based on a wiki platform, OpenEI is becoming a global leader in the energy data realm - specifically analyses of renewable energy and energy efficiency. Through this growing community, users can view, edit, add, and download data."

We choose the data in January to control the effect of ramp events, that is, the events when wind speed dramatically differs from the theoretical distribution. The observed variables used in this study were wind speed ($v_t, m/s$) and wind direction (θ_t, deg) measured at $100m$ height. Time series data were collected sequentially every 5 minutes during the whole month.

Assuming that wind velocity (v_t) is analyzed in two directions where each component is independently distributed with equal variance and mean. In this situation, the overall wind speed may be characterized by Rayleigh distribution. Based on such modeling assumption, one will customarily observe $v_t \cos(\theta_t)$ and $v_t \sin(\theta_t)$ at time t from normal distributions.

We focus on one dimension, say, $X_t = v_t \sin(\theta_t)$. One often assumes:

$$X_t = \mu + \phi X_{t-1} + \varepsilon_t \text{ where } \varepsilon_t \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2), \tag{3.6.1}$$

that is, the data is by a AR(1) model. Then, μ may represent the unknown seasonal information of wind speed, ϕ is the known coefficient determined by the wind turbine and σ^2 represents the constant variance of a white noise. It immediately follows that we have:

$$Y_t = X_t - \phi X_{t-1} \overset{\text{i.i.d.}}{\sim} N(\mu, \sigma^2). \tag{3.6.2}$$

In order to provide a precise and accurate estimate, a point estimate or a confidence interval estimate, of μ at the lowest cost, the data analyst will be expected to gather data sequentially and check every step of the way whether one should terminate sampling or not. However, there may not be a compelling need to update observations every 5 minutes. In other words, sampling one observation at-a-time may not be the most efficient way to gather data.

We can and should reduce the frequency by recording data more conveniently collected in groups. In this scenario, taking a pair ($k = 2$) of observations at-a-time is equivalent to looking at the data every 10 minutes. Now, let us illustrate both the FWCI and MRPE problems with the help of available wind speed data from NREL.

3.6.1. Illustration for the FWCI Problem

The width of a confidence interval plays an important role in forecasting. Considering the interquartile range of the Y_t's which spans from $-0.06(m/s)$ to $0.08(m/s)$, we pre-fixed

$\alpha = 0.05$ and three possible choices of d-values successively going down:

$$\boxed{d = 0.05(m/s), 0.01(m/s) \text{ and } 0.005(m/s).}$$

Table 3.18. 95% FWCI strategy (3.3.13) for the mean μ

d	k	i	kn	$\widehat{\mu}$	$s_{\widehat{\mu}}$	Lower CL $\widehat{\mu} - d$	Upper CL $\widehat{\mu} + d$
0.05	1		60	−0.0792	0.0164	−0.1042	−0.0542
	3	1	60	−0.0792	0.0164	−0.1042	−0.0542
	3	2	60	−0.0792	0.0164	−0.1042	−0.0542
	5	1	60	−0.0792	0.0164	−0.1042	−0.0542
	5	2	60	0.0219	0.0167	−0.0031	0.0469
0.01	1		1248	−0.0365	0.0051	−0.0415	−0.0315
	3	1	1251	−0.0361	0.0051	−0.0411	−0.0311
	3	2	1059	−0.0490	0.0056	−0.0540	−0.0440
	5	1	1275	−0.0354	0.0050	−0.0404	−0.0304
	5	2	1110	−0.0443	0.0056	−0.0493	−0.0393
0.005	1		4386	−0.0466	0.0026	−0.0491	−0.0441
	3	1	4380	−0.0466	0.0026	−0.0491	−0.0441
	3	2	3699	−0.0533	0.0029	−0.0559	−0.0509
	5	1	4245	−0.0481	0.0026	−0.0505	−0.0455
	5	2	3445	−0.0478	0.0030	−0.0503	−0.0453

Then, we implemented the FWCI strategy, first incorporating sampling observations every 5 minutes ($k = 1$), followed by updating observations every 15 minutes ($k = 3$) or then every 25 minutes ($k = 5$). The simulation studies from Sections 3.4.3.2 and 3.5.3.2 suggested that the pilot group size (m_0) had very little to nearly no impact on the terminal number of groups, N.

Thus, in these illustrations, we chose potentially large pilot sample size m_0 (such that $km_0 = 60$) then checked incoming sequential data under the requisite stopping rules (3.3.13). Final sample size (kn) from a single run produced each row in Table 3.18 along with the *lower confidence limit* (Lower CL) and *upper confidence limit* (Upper CL) associated with the terminal FWCI are reported in columns 7-8. The first column of Table 3.18 indicates the pre-fixed width of interval and column 2 shows the group size. With $k = 1$, the original purely sequential procedure (3.2.5) was used. With $k > 2$, sampling strategies from (3.3.13) were used with the corresponding identifier i (column 3).

In the first five rows of the table, the pilot size m_0 satisfied the pre-defined stopping boundary crossing. Thus, the sampling terminated right there and the final sample size was the pilot sample size 60. In this study, sampling k observations at-a-time tended to terminate earlier than that when sampling one observation at-a-time sequentially. We found that these constructed confidence intervals were more often than not reliable in the sense that most such intervals reported in Table 3.18 happened to include the mean ($= -0.0482$) from the full dataset.

3.6.2. Illustration for the MRPE Problem

Considering the loss caused by inaccurate point estimate of average wind speed, we implemented the MRPE strategies (3.3.14) with $A = 1000$ by varying the value of c_k, the cost of sampling per unit observation when recording k-tuples at-a-time. Assume it costs c_1 to download one single observation from the system. The total cost due to waiting time increases marginally when we download k observations at-a-time where we let $kc_k = \frac{1}{6}c_1(5 + k)$.

For example, if the operation time for downloading a single observation is 0.03 second, we may let $c_1 = 0.030\,\text{sec}$. Then, we should have: $2c_2 = 0.035\,\text{sec}$ to download two observations at-a-time, and $3c_3 = 0.040\,\text{sec}$ for three observations at-a-time. In other words, we have:

$$c_1 > c_2 > ... > c_k,$$

as expected.

Table 3.19 MRPE strategy (3.3.14) for the mean μ

c_1	k	i	kn	$\widehat{\mu}$	$s_{\widehat{\mu}}$	Loss
0.03	1		193	−0.1065	0.0098	4.1104
	3	1	453	−0.0460	0.0097	0.1089
	3	2	438	−0.0503	0.0099	0.1495
	5	1	495	−0.0415	0.0090	0.0851
	5	2	495	−0.0415	0.0090	0.0851
0.003	1		878	−0.0486	0.0066	0.0747
	3	1	1233	−0.0367	0.0051	0.0681
	3	2	1137	−0.0427	0.0054	0.0254
	5	1	1355	−0.0359	0.0047	0.0752
	5	2	1275	−0.0354	0.0050	0.0805
0.0003	1		2332	−0.0411	0.0034	0.0155
	3	1	3768	−0.0538	0.0028	0.1220
	3	2	3255	−0.0433	0.0031	0.0072
	5	1	4155	−0.0476	0.0026	0.0267
	5	2	3670	−0.0530	0.0029	0.1047

Suppose there are three types of operational systems with the corresponding values:

$$\boxed{c_1 = 0.03\,\text{sec},\ 0.003\,\text{sec},\ \text{and}\ 0.0003\,\text{sec},}$$

respectively. Then, we implemented the MRPE methodology (3.3.14) by gathering observations every 5 minutes ($k = 1$), followed by updating observations every 15 minutes ($k = 3$) or every 25 minutes ($k = 5$). The final terminal sample size (n), $\widehat{\mu}(= \overline{x})$ along with its standard deviation (s_x) and associated loss are reported in Table 3.19. Again, column 2 shows the group size and column 3 indicates corresponding sampling strategy refer from (3.3.14). Column 7 shows the observed loss when using the loss function defined in (3.2.14).

In Table 3.19, we observed no significant difference between estimates of μ across the board when taking the standard error into account. However, we realize that sampling a group of observations seemed to terminate relatively earlier than sampling one observation purely sequentially. As the sample size increased, however, the sample mean $\widehat{\mu}$ happened to be very close to the mean (0.0033) from full dataset. It may be a good idea to remind ourselves that each row in Table 3.19 was produced from a single run in the spirit of Table 3.18.

3.7. Brief Conclusion

We revisited the FWCI problem and MRPE problem both on the context of estimating unknown population mean μ when the population variance σ^2 is also unknown. The methodology discussed in Chapter 2 has potential drawback. We update the sampling in groups strategies with newly proposed estimators under permutation within groups which have obviously advantages in two ways: (i) the variances of the new estimators are uniformly smaller, and (ii) the variance of the stopping times are tighter around the optimal fixed sample size. These advantages are validated by the mathematical derivation as well as large scale simulation. It's not surprisingly that the new estimators can be written as functions of the within group sample variance, and the within group MAD. We have proved that such newly developed methods have associated asymptotic first-order and second-order properties. Our proposed theory are supported by the illustrations with real data.

Chapter 4

Second-Order Asymptotics for Comparing Treatment Means from Purely Sequential Estimation Strategies Under Possible Outlying Observations

4.1. Introduction and Layout

We revisit two-sample *minimum risk point estimation* (MRPE) problems which were amply developed in the books of Mukhopadhyay and de Silva (2009, Section 13.3) and Ghosh et al. (1997, Section 7.3). In that light, we begin by considering two *independent* populations, $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$, corresponding to two treatments with all four parameters assumed unknown, $(\mu_1, \mu_2) \in R^2$ and $(\sigma_1, \sigma_2) \in R^{+2}$. In this paper, our goal is to address sequential MRPE problems for the parameter $\mu_1 - \mu_2 (\equiv \delta$, say) under the *squared error loss* (SEL) plus linear cost of sampling.

The relevant literature is vast. Original formulations of such sequential MRPE problems were laid out in Mukhopadhyay (1976,1977) which incorporated a sampling *allocation scheme*, to be explained briefly in Section 4.2.2, and a sequential stopping number, much in the spirits of Robbins et al. (1967) and Srivastava (1970). Ghosh and Mukhopadhyay (1980) had dealt with asymptotic second-order purely sequential MRPE problems for δ assuming both σ_1^2, σ_2^2 were unknown and $\sigma_1^2 \neq \sigma_2^2$.

Mukhopadhyay and Moreno (1991) first pointed out that a sampling allocation scheme was not essential to carry out these kinds of MRPE strategies when σ_1^2, σ_2^2 were unknown but $\sigma_1^2 \neq \sigma_2^2$. Mukhopadhyay and Purkayastha (1994) considerably strengthened a similar line of thought. In all fairness, we should however point out that Robbins et al. (1967) and Srivastava (1970) focused exclusively on nonparametric sequential *fixed-width confidence interval* (FWCI) estimation problems for δ.

Mukhopadhyay (1976) additionally developed sequential estimation problems of comparing three treatment means when all variances were unknown and unequal. In the two-sample case where σ_1^2, σ_2^2 were unknown, but the ratio $\sigma_1^2\sigma_2^{-2}$ was known, FWCI problems for δ were developed by Mukhopadhyay et al. (2010) which were further improved upon by Aoshima et al. (2011). The literature on associated multivariate problems is also widely developed and spread out. One may review a broad set of related methodologies from Woodroofe (1977,1982), Siegmund (1985), Ghosh and Sen (1991), Ghosh et al. (1997, pp. 222-223, 256-259), Mukhopadhyay and de Silva (2009, chapter 13, pp.356-359), and Zacks (2009,2017) among other sources.

4.1.1. Layout in This Chapter

In Section 4.2, we summarize the basic MRPE problems along with some requisite preliminaries. We introduce two distinct situations: (i) $\sigma_1^2 = \sigma_2^2 = \sigma^2$ with σ^2 unknown (Section 4.2.1), and (ii) σ_1^2, σ_2^2 are both unknown but $\sigma_1^2 \neq \sigma_2^2$ (Section 4.2.2).

Under either scenario, in the spirits of the recently improved methodologies (Mukhopadhyay and Chattopadhyay 2013; Mukhopadhyay and Hu 2017; Hu and Mukhopadhyay 2019) developed in the contexts of one-sample problems, we proceed to emphasize both asymptotic *first-order* and *second-order* results in addressing our two-sample problems. In doing so, we move to replace the multiples of sample standard deviations used in defining requisite boundary crossing conditions with the GMD's, the MAD's, along with a number of their combinations with the sample standard deviations.

The scenario of common but unknown variance is fully developed in Section 4.3. Hu and Mukhopadhyay (2019) laid down the ground-work in the case of their one-sample problems. We substantially improvised upon such ground-work to be able to develop the required asymptotic first-order and second-order analyses in our present two-sample problems and then apply these powerful techniques to succeed. Major theoretical findings are summarized by Theorems 4.3.1-4.3.2 and (4.3.23)-(4.3.24). The associated theory and methodology are supplemented by the analysis of data obtained from simulations (Section 4.3.5) as well as with illustrations using real data (Section 4.3.6).

The scenario of unknown but unequal variances is fully developed in Section 4.4. We further our improvisations from Section 3 and generalize them to suit the present scenario. Again, we move to replace the multiples of sample standard deviations used in defining requisite boundary crossing conditions with the GMD's, the MAD's, along with a number of their combinations with the sample standard deviations. We pushed theory hard enough to be able to develop the required asymptotic first-order and second-order analyses in our two-sample problems and then apply these powerful techniques to succeed. Major theoretical findings are summarized by (4.4.5)-(4.4.8) and (4.4.14)-(4.4.15). The associated theories and methodologies are supplemented by the analysis of data obtained from simulations (Section 4.4.2) as well as with illustrations using real data (Section 4.3).

This discourse ends with a brief set of concluding thoughts.

4.2. Formulation of Problems and Some Preliminaries

We formulate and handle two distinct scenarios. The first situation (Section 4.2.1) deals with equal but unknown common variance. The second situation (Section 4.2.2) assumes unequal and unknown variances.

4.2.1. Equal But Unknown Common Variance

Under this scenario, we record equal number of observations from both treatments. Let $X_{i1}, X_{i2}, ..., X_{in}, ...$ be *independent and identically distributed* (i.i.d.) observations from $N(\mu_i, \sigma^2)$, $i = 1, 2$. We denote $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma) \in R^2 \times R^+$ assuming that all three parameters are unknown.

Having recorded $\{X_{i1}, X_{i2}, ..., X_{in}, i = 1, 2\}$, we let:

$$\begin{gathered}\overline{X}_{in} = n^{-1}\Sigma_{j=1}^{n} X_{ij},\ S_{in}^2 = (n-1)^{-1}\Sigma_{j=1}^{n}(X_{ij} - \overline{X}_{in})^2,\ i = 1, 2,\ T_n = \overline{X}_{1n} - \overline{X}_{2n}, \\ \text{and } S_{Pn}^2 = \tfrac{1}{2}\{S_{1n}^2 + S_{2n}^2\}, \text{ the pooled sample variance, } n \geq 2.\end{gathered} \tag{4.2.1}$$

Suppose that the loss function in estimating $\delta(= \mu_1 - \mu_2)$ by means of $T_n(= \overline{X}_{1n} - \overline{X}_{2n})$ is given by:

$$L_n(\delta, T_n) \equiv L_n = A(T_n - \delta)^2 + 2cn \text{ with } A > 0,\ c > 0, \tag{4.2.2}$$

where A is a known weight function and c is the known cost per unit observation. The associated risk is:

$$R_n \equiv E_{\boldsymbol{\theta}}[L_n] = 2A\sigma^2 n^{-1} + 2cn,$$

which is minimized approximately (for large n) when we have the following:

$$\begin{gathered}n \text{ is the smallest integer } (Ac^{-1})^{1/2}\sigma = n^*, \text{ say, which implies} \\ \text{the associated fixed-sample-size risk, } R_{n^*} = 4cn^*.\end{gathered} \tag{4.2.3}$$

This n^* is referred to as the optimal fixed sample size required from *both* populations had σ^2 been known. But, σ^2 remains unknown and hence so is n^*. In the spirit of Robbins (1959), we first propose the following purely sequential estimation strategy. One may refer to Mukhopadhyay and de Silva (2009, Section 13.3.1, pp. 356-359).

We begin with pilot data $\{X_{i1}, X_{i2}, ..., X_{im}, i = 1, 2\}$ where $m(\geq 2)$ is the pilot size. Then, we follow up with one additional pair of observations (X_1, X_2) at-a-time according to the following stopping rule:

$$\text{Methodology } \mathcal{Q}_{1,0}: \quad N_{\mathcal{Q}_{1,0}}(c) \equiv N = \inf\{n \geq m : \quad n \geq (Ac^{-1})^{1/2} S_{\mathcal{P}n}\}. \tag{4.2.4}$$

Indeed, under the methodology $\mathcal{Q}_{1,0}$, we have $P_{\boldsymbol{\theta}}\{N < \infty\} = 1$, and thus upon termination, we propose to estimate δ with the terminal estimator $T_N(= \overline{X}_{1N} - \overline{X}_{2N})$ based on fully accrued data:

$$\{N_{\mathcal{Q}_{1,0}}, X_{i1}, X_{i2}, ..., X_{iN_{\mathcal{Q}_{1,0}}}, i = 1, 2\}.$$

Next, we note that the random variable $I(N = n)$ depends only on $(S_{\mathcal{P}m}, ..., S_{\mathcal{P}n})$ for every fixed $n \geq m$. Thus, incorporating Basu's (1955) theorem, we can claim that the random variables $I(N = n)$ and $(\overline{X}_{1n}, \overline{X}_{2n})$, and hence $I(N = n)$ and T_n, are independent for all fixed $n \geq m$. One may additionally refer to Mukhopadhyay (2000, pp. 324-327) and review how Basu's (1955) theorem applies.

Thus, the associated achieved sequential risk and other requisite entities are given by:

$$\begin{aligned}
\text{Sequential risk:}\quad & R_{N,\mathcal{Q}_{1,0}} \equiv E_{\boldsymbol{\theta}}[L_{N_{\mathcal{Q}_{1,0}}}] = 2A\sigma^2 E_{\boldsymbol{\theta}}[N_{\mathcal{Q}_{1,0}}^{-1}] + 2cE_{\boldsymbol{\theta}}[N_{\mathcal{Q}_{1,0}}],\\
\text{Risk efficiency:}\quad & \kappa_{\mathcal{Q}_{1,0}}(c) \equiv \kappa = R_{N,\mathcal{Q}_{1,0}}/R_{n^*} = \tfrac{1}{2}E_{\boldsymbol{\theta}}[n^* N_{\mathcal{Q}_{1,0}}^{-1}] + \tfrac{1}{2}E_{\boldsymbol{\theta}}[N_{\mathcal{Q}_{1,0}} n^{*-1}], \qquad (4.2.5)\\
\text{Regret:}\quad & \omega_{\mathcal{Q}_{1,0}}(c) \equiv \omega = R_{N,\mathcal{Q}_{1,0}} - R_{n^*} = 2cE_{\boldsymbol{\theta}}\left[\tfrac{(N_{\mathcal{Q}_{1,0}} - n^*)^2}{N_{\mathcal{Q}_{1,0}}}\right].
\end{aligned}$$

In a one-sample problem, the *risk-efficiency* (κ) and *regret* (ω) measures were originally proposed by Robbins (1959) which were further developed by others including Starr (1966), Starr and Woodroofe (1969), Ghosh and Mukhopadhyay (1976,1980,1981), Woodroofe (1977,1982), and Mukhopadhyay (1988,1991).

4.2.2. Unequal and Unknown Variances

In this situation, we record unequal number of observations from the two treatments. Now, we may denote $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma_1, \sigma_2) \in R^2 \times R^{+2}$ assuming that all four parameters are unknown.

Let $X_{i1}, X_{i2}, ..., X_{in_i}, ...$ be i.i.d. observations from $N(\mu_i, \sigma_i^2)$, $i = 1, 2$. Having recorded $\{X_{i1}, X_{i2}, ..., X_{in_i}, \ i = 1, 2\}$, we denote:

$$\begin{gathered}
\overline{X}_{in_i} = n_i^{-1}\Sigma_{j=1}^{n_i} X_{ij},\ S_{in_i}^2 = (n_i - 1)^{-1}\Sigma_{j=1}^{n_i}(X_{ij} - \overline{X}_{in_i})^2,\ n_i \geq 2,\ i = 1, 2,\ \mathbf{n} = (n_1, n_2),\\
T_{\mathbf{n}} = \overline{X}_{1n_1} - \overline{X}_{2n_2} \text{ and the total sample size } n = n_1 + n_2.
\end{gathered} \qquad (4.2.6)$$

In the spirit of (4.2.2), suppose that the loss function in estimating $\delta(= \mu_1 - \mu_2)$ by $T_{\mathbf{n}}(= \overline{X}_{1n_1} - \overline{X}_{2n_2})$ is given by:

$$L_{\mathbf{n}}(\delta, T_{\mathbf{n}}) \equiv L_{\mathbf{n}} = A(T_{\mathbf{n}} - \delta)^2 + c(n_1 + n_2) \text{ with } A > 0,\ c > 0, \qquad (4.2.7)$$

where A is a known weight function and c is the cost per unit observation. We could easily incorporate unequal costs of sampling from the two treatments, but we avoid that in favor of a more readable presentation.

The associated risk is given by:

$$R_{\mathbf{n}} \equiv E_{\boldsymbol{\theta}}[L_{\mathbf{n}}] = A(\sigma_1^2 n_1^{-1} + \sigma_2^2 n_2^{-1}) + c(n_1 + n_2),$$

which is minimized approximately (for large n_1, n_2) when we have the following:

$$\begin{gathered} n_i \text{ is the smallest integer } \geq (Ac^{-1})^{1/2}\sigma_i = n_i^*, \text{ say, } i = 1, 2, \text{ which imply the} \\ \text{minimum fixed-sample-size risk, } R_{\mathbf{n}^*} = 2cn^*,\ n^* = n_1^* + n_2^*, \text{ the total} \\ \text{optimal fixed-sample-size, and we let } \mathbf{n}^* = (n_1^*, n_2^*). \end{gathered} \tag{4.2.8}$$

This n^* is the combined optimal fixed sample size required from *both* treatments had σ_1^2, σ_2^2 been known. But, σ_1^2, σ_2^2 are unknown and hence n_1^*, n_2^*, n^* remain unknown. In the spirit of Robbins et al. (1967), one might propose a stopping rule determining both sample sizes after incorporating an appropriate *sampling allocation* scheme. The role of an allocation scheme was to precisely decide which treatment to sample from to record the next observation every step of the way before termination. Indeed, Mukhopadhyay (1976,1977) and Ghosh and Mukhopadhyay (1980) implemented such ideas which seemed reasonable at the time.

However, Mukhopadhyay and Moreno (1991) first observed that for the problem on hand, no sampling allocation scheme was truly essential at all especially since n_i^* depended exclusively on σ_i, $i = 1, 2$. In a number of subsequent papers handling problems from multiple comparisons, the authors followed that lead. For example, one may refer to Mukhopadhyay and Chattopadhyay (1991), Mukhopadhyay and Purkayastha (1994) and other sources. Also, one may refer to Mukhopadhyay and de Silva (2009, Section 13.3.2).

We begin with pilot data $\{X_{i1}, X_{i2}, ..., X_{im}, i = 1, 2\}$, $m \geq 2$, $i = 1, 2$. Then, we follow up with one additional observation X_i at-a-time from the i^{th} treatment according to the following stopping rule:

Methodology $\mathcal{Q}_{2,0}$: The total sample size $N_{\mathcal{Q}_{2,0}}(c) \equiv N = N_{1,\mathcal{Q}_{2,0}} + N_{2,\mathcal{Q}_{2,0}}$ where

$$N_{i,\mathcal{Q}_{2,0}} \equiv N_i = \inf\{n \geq m : \quad n \geq (Ac^{-1})^{1/2} S_{in}\},\ i = 1, 2. \tag{4.2.9}$$

Observe that the purely sequential sampling strategy (4.2.9) from both treatments are carried out in parallel, and independently of each other. In other words, $N_{1,\mathcal{Q}_{2,0}}$ and $N_{2,\mathcal{Q}_{2,0}}$ are both genuine stopping variables distributed independently of each other.

Indeed, under the methodology $\mathcal{Q}_{2,0}$, we have $P_{\boldsymbol{\theta}}\{N_{i,\mathcal{Q}_{2,0}} < \infty\} = 1$, $i = 1, 2$, and thus upon termination, we propose to estimate δ with the terminal estimator $T_{\mathbf{N}}(= \overline{X}_{1N_1} - \overline{X}_{2N_2})$ based on fully accrued data:

$$\{N_{i,\mathcal{Q}_{2,0}}, X_{i1}, X_{i2}, ..., X_{iN_{i,\mathcal{Q}_{2,0}}}, i = 1, 2\}.$$

Then, the associated sequential risk and other requisite entities are given by:

$$\begin{aligned}
\text{Sequential risk:}\quad & R_{\mathbf{N},\mathcal{Q}_{2,0}} \equiv E_{\boldsymbol{\theta}}[L_{\mathbf{N}_{\mathcal{Q}_{2,0}}}] = A\Sigma_{i=1}^{2}\sigma_i^2 E_{\boldsymbol{\theta}}[N_{i,\mathcal{Q}_{2,0}}^{-1}] + c\Sigma_{i=1}^{2}E_{\boldsymbol{\theta}}[N_{i,\mathcal{Q}_{2,0}}],\\
\text{Risk efficiency:}\quad & \kappa_{\mathcal{Q}_{2,0}}(c) \equiv \kappa = R_{\mathbf{N},\mathcal{Q}_{2,0}}/R_{\mathbf{n}^*} = \tfrac{1}{2}\Sigma_{i=1}^{2}\tfrac{n_i^*}{n^*}\left\{E_{\boldsymbol{\theta}}[n_i^* N_{i,\mathcal{Q}_{2,0}}^{-1}] + E_{\boldsymbol{\theta}}[N_{i,\mathcal{Q}_{2,0}} n_i^{*-1}]\right\},\\
\text{Regret:}\quad & \omega_{\mathcal{Q}_{2,0}}(c) \equiv \omega = R_{\mathbf{N},\mathcal{Q}_{2,0}} - R_{\mathbf{n}^*} = c\Sigma_{i=1}^{2}E_{\boldsymbol{\theta}}\left[\tfrac{(N_{i,\mathcal{Q}_{2,0}}-n_i^*)^2}{N_{i,\mathcal{Q}_{2,0}}}\right].
\end{aligned} \tag{4.2.10}$$

4.3. Common Unknown Variance

We are back to the scenario where the two treatments have a common but unknown variance σ^2, that is we are back on track where we had left with the preliminaries in Section 2.1. Recall that we had denoted $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma)$.

4.3.1. Summary from Nonlinear Renewal Theory

We note that we may equivalently express $N_{\mathcal{Q}_{1,0}} = J + 1$ w.p.1 where we denote:

$$J_c \equiv J = \inf\{n \geq m-1: \quad 2n^3(1+n^{-1})^2 n^{*-2} \geq \Sigma_{j=1}^{n} Y_j\}, \quad \text{the } Y_j\text{'s being i.i.d. } \chi_2^2. \tag{4.3.1}$$

Now, we apply nonlinear renewal theory by proceeding along the lines of Woodroofe (1977,1982), Lai and Siegmund (1977,1979), Mukhopadhyay (1988), Mukhopadhyay and Solanky (1994, Section 2.4.2), Ghosh et al. (1997, pp. 64-65) and Mukhopadhyay and de Silva (2009, Section A.4).

More specifically, we first identify with a basic set of notations laid out in Mukhopadhyay and Solanky (1994, Section 2.4.2):

$$\delta = 3,\ L_0 = 2,\ \theta = 2,\ \tau^2 = 4,\ b = 1,\ \beta^* = \tfrac{1}{2},\ n_0^* = n^*,\ h^* = 2n^{*-2}, \text{ and } p = \tfrac{1}{4}, \tag{4.3.2}$$

in order to properly study J from (4.3.1). Let us also define:

$$\nu = \tfrac{5}{2} - D \text{ where } D = \Sigma_{n=1}^{\infty} n^{-1} E\left[\max\left(0, \chi^2_{2n} - 6n\right)\right] \text{ and } \eta = \tfrac{1}{4}\nu - \tfrac{11}{8}. \tag{4.3.3}$$

Next, from the established properties of J summarized in Theorem 2.4.8 from Mukhopadhyay and Solanky (1994), we conclude as $c \to 0$:

$$\begin{aligned}
&\text{a)} \quad P_{\boldsymbol{\theta}}\{N_{\mathcal{Q}_{1,0}} \le \varepsilon n^*\} = O(n^{*-2m+2}) \text{ if } 0 < \varepsilon < 1,\, m \ge 2;\\
&\text{b)} \quad N^*_{\mathcal{Q}_{1,0}} \equiv n^{*-1/2}(N_{\mathcal{Q}_{1,0}} - n^*) \xrightarrow{\pounds} N(0, \tfrac{1}{4}) \text{ if } m \ge 2;\\
&\text{c)} \quad \left|N^*_{\mathcal{Q}_{1,0}}\right|^r \text{ is uniformly integrable if } m > 1 + \tfrac{1}{4}r,\, r > 0.
\end{aligned} \tag{4.3.4}$$

Now, we highlight the following asymptotic first-order results as $c \to 0$ when $m \ge 2$:

$$\begin{aligned}
&\text{a)} \quad E_{\boldsymbol{\theta}}[N_{\mathcal{Q}_{1,0}} n^{*-1}] \to 1 \text{ [\textit{asymptotic first-order efficiency}]};\\
&\text{b)} \quad E_{\boldsymbol{\theta}}[n^* N^{-1}_{\mathcal{Q}_{1,0}}] \to 1;\\
&\text{c)} \quad \kappa_{\mathcal{Q}_{1,0}}(c) \to 1 \text{ [\textit{asymptotic risk efficiency}]};
\end{aligned} \tag{4.3.5}$$

with the risk efficiency, $\kappa_{\mathcal{Q}_{1,0}}(c)$, coming from (4.2.5).

Conclusion (4.3.5), part (a) follows directly from Chow and Robbins (1965) where this property was called as asymptotic efficiency. The same property alternatively follows by invoking Wiener's (1939) ergodic theorem in combination with the dominated convergence theorem. Ghosh and Mukhopadhyay (1981) began referring to the same property as *asymptotic first-order efficiency* after they developed the notion of *asymptotic second-order efficiency* property.

We can easily verify (4.3.5), part (b) as follows: We first note that $n^{*-1} N_{\mathcal{Q}_{1,0}} \xrightarrow{P_{\boldsymbol{\theta}}} 1$ as

$c \to 0$, and hence $n^* N_{\mathcal{Q}_{1,0}}^{-1} \overset{P_{\boldsymbol{\theta}}}{\to} 1$ as $c \to 0$. Clearly,

$$n^* N_{\mathcal{Q}_{1,0}}^{-1} I(N_{\mathcal{Q}_{1,0}} > \tfrac{1}{2} n^*)$$

is bounded and hence *uniformly integrable* (u.i.) which implies

$$E_{\boldsymbol{\theta}} \left[n^* N_{\mathcal{Q}_{1,0}}^{-1} I(N_{\mathcal{P}_{1,0}} > \tfrac{1}{2} n^*) \right] = 1 + o(1) \text{ if } m \geq 2. \tag{4.3.6}$$

Also, we have

$$E_{\boldsymbol{\theta}} \left[n^* N_{\mathcal{Q}_{1,0}}^{-1} I(N_{\mathcal{Q}_{1,0}} \leq \tfrac{1}{2} n^*) \right] \leq m^{-1} n^* P_{\boldsymbol{\theta}} \{ N_{\mathcal{Q}_{1,0}} \leq \tfrac{1}{2} n^* \} = O(n^{*-2m+3}) = o(1) \text{ if } m > \tfrac{3}{2}. \tag{4.3.7}$$

Combining (4.3.6)-(4.3.7), we conclude (4.3.5), part (b).

Next, (4.3.5), part (c) follows from the expression of $\kappa_{\mathcal{Q}_{1,0}}(c)$ given in (4.2.5) after we combine (4.3.5), parts (a) and (b).

Theorem 4.3.1. *For the MRPE strategy* $\left(N_{\mathcal{Q}_{1,0}}, T_{N_{\mathcal{Q}_{1,0}}} \right)$ *from (4.2.4), we have as* $c \to 0$:

(*i*) $E_{\boldsymbol{\theta}}[N_{\mathcal{Q}_{1,0}}] = n^* + (\frac{1}{4}\nu - \frac{3}{8}) + o(1)$ *if* $m \geq 2$ [*asymptotic second-order efficiency*];

(*ii*) $\omega_{\mathcal{Q}_{1,0}}(c) = \frac{1}{2} c + o(c)$ *if* $m \geq 3$ [*asymptotic second-order regret*];

where n^*, $\omega_{\mathcal{Q}_{1,0}}(c)$, *and* ν *come from (4.2.3), (4.2.5), and (4.3.3) respectively.*

Proof: First, we set out to evaluate the important and requisite entities from (4.3.3) numerically that are correct up to 6 decimal places:

$$D \approx 0.125823 \Rightarrow \eta = \tfrac{1}{4}\nu - \tfrac{11}{8} = -\tfrac{3}{4} - \tfrac{1}{4}D \approx -0.781456, \tag{4.3.8}$$

so that (Mukhopadhyay and Solanky 1994, Theorem 2.4.8, part (v) with $\omega = 1$), we can claim:

$$E_{\boldsymbol{\theta}}[N_{\mathcal{Q}_{1,0}}] = n^* + \eta + 1 + o(1) \approx n^* + 0.218544 + o(1) \text{ if } m \geq 2. \tag{4.3.9}$$

The equation (4.3.9) shows that the purely sequential estimation strategy (4.2.4) is asymptotically second-order efficient in the sense of Ghosh and Mukhopadhyay (1981).

Next, we recall from (4.2.5):

$$\omega_{\mathcal{Q}_{1,0}}(c) = 2cE_{\boldsymbol{\theta}}\left[\tfrac{(N_{\mathcal{Q}_{1,0}}-n^*)^2}{N_{\mathcal{Q}_{1,0}}}\right],$$

and we observe (w.p.1):

$$0 < \tfrac{(N_{\mathcal{Q}_{1,0}}-n^*)^2}{N_{\mathcal{Q}_{1,0}}} I(N_{\mathcal{Q}_{1,0}} > \tfrac{1}{2}n^*) < 2\tfrac{(N_{\mathcal{Q}_{1,0}}-n^*)^2}{n^*},$$

but $\frac{(N_{\mathcal{Q}_{1,0}}-n^*)^2}{n^*}$ is u.i. using (4.3.4), part (c) if $m \geq 2$, so that $\frac{(N_{\mathcal{Q}_{1,0}}-n^*)^2}{N_{\mathcal{Q}_{1,0}}} I(N_{\mathcal{Q}_{1,0}} > \frac{1}{2}n^*)$ is also u.i. if $m \geq 2$.

Now, since

$$\tfrac{(N_{\mathcal{Q}_{1,0}}-n^*)^2}{N_{\mathcal{Q}_{1,0}}} I(N_{\mathcal{Q}_{1,0}} > \tfrac{1}{2}n^*) \xrightarrow{\pounds} \tfrac{1}{4}\chi_1^2 \text{ as } c \to 0,$$

we combine with the uniform integrability to claim:

$$E_{\boldsymbol{\theta}}\left[\frac{(N_{\mathcal{Q}_{1,0}}-n^*)^2}{N_{\mathcal{Q}_{1,0}}} I(N_{\mathcal{Q}_{1,0}} > \tfrac{1}{2}n^*)\right] = \tfrac{1}{4} + o(1) \text{ if } m \geq 2. \tag{4.3.10}$$

On the other hand, we can express:

$$\begin{aligned} 0 < E_{\boldsymbol{\theta}}\left[\frac{(N_{\mathcal{Q}_{1,0}}-n^*)^2}{N_{\mathcal{Q}_{1,0}}} I(N_{\mathcal{Q}_{1,0}} \leq \tfrac{1}{2}n^*)\right] &\leq E_{\boldsymbol{\theta}}\left[\frac{N_{\mathcal{Q}_{1,0}}^2+n^{*2}}{m} I(N_{\mathcal{Q}_{1,0}} \leq \tfrac{1}{2}n^*)\right] \\ \leq \tfrac{5}{4m} n^{*2} P_{\boldsymbol{\theta}}\{N_{\mathcal{Q}_{1,0}} \leq \tfrac{1}{2}n^*\} &= O(n^{*2})O(n^{*2-2m}) = o(1) \text{ if } m \geq 3, \end{aligned} \tag{4.3.11}$$

by (4.3.4), part (a). Combining (4.3.10)-(4.3.11), we can claim:

$$\omega_{\mathcal{Q}_{1,0}}(c) \equiv \tfrac{1}{4}(2c) + o(c) = \tfrac{1}{2}c + o(c) \text{ if } m \geq 3.$$

The proof is now complete. ■

4.3.2. Stopping Rules with Selected Combinations from $\mathbf{S_{Pn}}$, GMD, and MAD

Let us begin by putting forward a kind of generic purely sequential estimation strategy as follows:

$$\text{Methodology } \mathcal{Q}_1 \begin{cases} N_{\mathcal{Q}_1}(c) \equiv N_{\mathcal{Q}_1} = \inf\left\{n \geq m : \quad n \geq (Ac^{-1})^{1/2}\left(W_n + n^{-\lambda}\right)\right\}, \lambda > \tfrac{1}{2}, \\ \text{estimating } \delta \text{ with a terminal estimator, } T_{N_{\mathcal{Q}_1}} (\equiv T_N = \overline{X}_{1N} - \overline{X}_{2N}), \\ \text{from fully accrued data } \{N_{\mathcal{Q}_1}, X_{i1}, X_{i2}, ..., X_{iN_{\mathcal{Q}_1}}, i = 1, 2\}. \end{cases} \tag{4.3.12}$$

Here, "W_n" stands for a consistent pooled estimator of σ. Indeed, under the methodology $\mathcal{Q}_1$, we expect $P_{\boldsymbol{\theta}}\{N_{\mathcal{Q}_1} < \infty\} = 1$ to hold, so that the terminal estimator $T_{N_{\mathcal{Q}_1}}$ will be meaningful. Recall that (4.2.4) gave the methodology $\mathcal{Q}_1 \equiv \mathcal{Q}_{1,0}$ with $W_n = S_{\mathcal{P}n}$ which was consistent, but biased, for σ. But, note that in defining $\mathcal{Q}_{1,0}$, we did not use a fudge factor such as $n^{-\lambda}$.

We can easily verify for $n \geq 2$:

$$E_{\boldsymbol{\theta}}\left[a_n S_{Pn}\right] = \sigma \text{ where } a_n = (n-1)^{1/2}\Gamma\left(n-1\right)\left\{\Gamma\left(n-\tfrac{1}{2}\right)\right\}^{-1}. \tag{4.3.13}$$

Then, the original estimation strategy (4.2.4) may be modified as follows:

$$\text{Methodology } \mathcal{Q}_{1,1}\text{:} \quad N_{\mathcal{Q}_{1,1}}(c) \equiv N = \inf\{n \geq m: \quad n \geq (Ac^{-1})^{1/2} a_n S_{Pn}\}, \tag{4.3.14}$$

with the terminal estimator $T_N(= \overline{X}_{1N} - \overline{X}_{2N})$ for δ based on fully accrued data upon termination:

$$\{N_{\mathcal{Q}_{1,1}}, X_{i1}, X_{i2}, ..., X_{iN_{\mathcal{Q}_{1,1}}}, i = 1, 2\}.$$

Next, in order to explore other tangible choices of W_n, let us first denote:

$$G_{in} = \tfrac{1}{2}\sqrt{\pi}\binom{n}{2}^{-1} \underset{1\leq j<k\leq n}{\Sigma\Sigma} \left|X_{ij} - X_{ik}\right|, \text{ and } M_{in} = \sqrt{\tfrac{\pi n}{2(n-1)}} n^{-1}\Sigma_{k=1}^{n} \left|X_{ik} - \overline{X}_{in}\right|, \tag{4.3.15}$$

with $i = 1, 2$. The G_{in}'s and M_{in}'s from (4.3.15) respectively correspond to unbiased and consistent estimators of σ derived from GMD and MAD respectively based on random samples from the i^{th} treatment, $i = 1, 2$.

We will successively substitute W_n in (4.3.12), the generic statistic that was used earlier to define the first-time boundary crossing, with the following unbiased and consistent estimators of σ. That way, we will introduce a rather large array of newer MRPE strategies:

$$\begin{aligned}
&\text{Methodology } \mathcal{Q}_1 \equiv \mathcal{Q}_{1,1}: && \text{Replace } S_{Pn} \text{ in (4.2.4) with } a_n S_{Pn} \text{ and } a_n \text{ from (4.3.13);}\\
&\text{Methodology } \mathcal{Q}_1 \equiv \mathcal{Q}_{1,2}: && \text{Replace } W_n \text{ in (4.3.12) with } \tfrac{1}{2}(G_{1n}+G_{2n});\\
&\text{Methodology } \mathcal{Q}_1 \equiv \mathcal{Q}_{1,3}: && \text{Replace } W_n \text{ in (4.3.12) with } \tfrac{1}{2}(M_{1n}+M_{2n});\\
&\text{Methodology } \mathcal{Q}_1 \equiv \mathcal{Q}_{1,4}: && \text{Replace } W_n \text{ in (4.3.12) with } \tfrac{1}{2}(G_{1n}+M_{2n});\\
&\text{Methodology } \mathcal{Q}_1 \equiv \mathcal{Q}_{1,5}: && \text{Replace } W_n \text{ in (4.3.12) with } \tfrac{1}{2}(M_{1n}+G_{2n}).
\end{aligned} \tag{4.3.16}$$

One should observe that in (4.3.16), we mixed unbiased-consistent-independent, and yet rather unconventional, estimators of σ obtained from the observations gathered on both treatments.

Next, we successively consider replacing W_n in (4.3.12) with the following unbiased, consistent, but not necessarily independent, estimators of σ one-by-one defined via certain specific pooled averages from the list consisting of $a_n S_{Pn}$, G_{1n}, G_{2n}, M_{1n}, and M_{2n}. Thereby, we will be introducing another array of newer MRPE strategies as follows:.

$$\begin{aligned}
&\text{Methodology } \mathcal{Q}_1 \equiv \mathcal{Q}_{1,6}: && \text{Replace } W_n \text{ in (4.3.12) with } \tfrac{1}{2}(a_n S_{Pn} + G_{1n});\\
&\text{Methodology } \mathcal{Q}_1 \equiv \mathcal{Q}_{1,7}: && \text{Replace } W_n \text{ in (4.3.12) with } \tfrac{1}{2}(a_n S_{Pn} + M_{1n});\\
&\text{Methodology } \mathcal{Q}_1 \equiv \mathcal{Q}_{1,8}: && \text{Replace } W_n \text{ in (4.3.12) with } \tfrac{1}{2}(a_n S_{Pn} + G_{2n});\\
&\text{Methodology } \mathcal{Q}_1 \equiv \mathcal{Q}_{1,9}: && \text{Replace } W_n \text{ in (4.3.12) with } \tfrac{1}{2}(a_n S_{Pn} + M_{2n});\\
&\text{Methodology } \mathcal{Q}_1 \equiv \mathcal{Q}_{1,10}: && \text{Replace } W_n \text{ in (4.3.12) with } \tfrac{1}{3}(a_n S_{Pn} + G_{1n} + G_{2n});\\
&\text{Methodology } \mathcal{Q}_1 \equiv \mathcal{Q}_{1,11}: && \text{Replace } W_n \text{ in (4.3.12) with } \tfrac{1}{3}(a_n S_{Pn} + M_{1n} + M_{2n});\\
&\text{Methodology } \mathcal{Q}_1 \equiv \mathcal{Q}_{1,12}: && \text{Replace } W_n \text{ in (4.3.12) with } \tfrac{1}{3}(a_n S_{Pn} + G_{1n} + M_{2n});\\
&\text{Methodology } \mathcal{Q}_1 \equiv \mathcal{Q}_{1,13}: && \text{Replace } W_n \text{ in (4.3.12) with } \tfrac{1}{3}(a_n S_{Pn} + M_{1n} + G_{2n}).
\end{aligned} \tag{4.3.17}$$

However, in (4.3.16)-(4.3.17), we do not want to go overboard by considering convex combinations other than those that we have illustrated. Such arbitrary convex combinations will cause no sense of excessive strain with regard to theory and/or methodology. We simply refrain from going along that route for brevity, and yet, hope to get our main ideas across. The estimation strategies from (4.3.12) define genuine stopping variables in the sense that the result $P_{\boldsymbol{\theta}}\{N_{\mathcal{Q}_1} < \infty\} = 1$ holds while using W_n's from (4.3.16)-(4.3.17).

Next, under such methodologies, we note that the random variable $I(N_{\mathcal{Q}_1} = n)$ depends only on $(W_m, ..., W_n)$ which is a location invariant vector-valued random variable for every fixed $n \geq m$. Hence, incorporating Basu's (1955) theorem, we can again claim that the random variables $I(N_{\mathcal{Q}_1} = n)$ and $(\overline{X}_{1n}, \overline{X}_{2n})$, and hence $I(N_{\mathcal{Q}_1} = n)$ and T_n, are independently distributed for all fixed $n \geq m$. One may additionally refer to Mukhopadhyay (2000, pp. 324-327).

4.3.3. Asymptotic Second-Order Efficiency and Regret Expansion for $\mathcal{Q}_1 \equiv \mathcal{Q}_{1,1}$

With a_n from (4.3.13), we can express:

$$a_n = 1 + \tfrac{1}{8}n^{-1} + O(n^{-2}) \Rightarrow a_{n+1}^{-2} = 1 - \tfrac{1}{4}n^{-1} + O(n^{-2}). \tag{4.3.18}$$

One may refer to Abramowitz and Stegun (1972, p. 257, 6.1.47).

We note that we may equivalently express $N_{\mathcal{Q}_{1,1}} = R + 1$ w.p.1 where we denote:

$$\begin{gathered} R_c \equiv r = \inf\{n \geq m - 1: \quad 2n^3(1 + n^{-1})^2 a_{n+1}^{-2} n^{*-2} \geq \Sigma_{j=1}^{n} Y_j\}, \\ \text{the } Y_j\text{'s being i.i.d. } \chi_2^2. \end{gathered} \tag{4.3.19}$$

Again, we can apply nonlinear renewal theory from Woodroofe (1977) by proceeding along the lines of Woodroofe (1977,1982), Lai and Siegmund (1977,1978), Mukhopadhyay (1988), Mukhopadhyay and Solanky (1994, Section 2.4.2), Ghosh et al. (1997, pp. 64-65) and Mukhopadhyay and de Silva (2009, Section A.4).

We first identify with a basic set of notation from Mukhopadhyay and Solanky (1994, Section 2.4.2) to note:

$$\begin{gathered} \delta = 3,\ L_0 = \tfrac{7}{4},\ \theta = 2,\ \tau^2 = 4,\ b = 1,\ \beta^* = \tfrac{1}{2},\ n_0^* = n^*,\ h^* = 2n^{*-2}, \\ p = \tfrac{1}{4}, \text{ so that } \eta = \tfrac{1}{4}\nu - \tfrac{5}{4} \text{ with } \nu \text{ coming from (4.3.3).} \end{gathered} \tag{4.3.20}$$

in order to handle R from (4.3.19).

Next, from the established properties of R as summarized in Theorem 2.4.8 from Mukhopadhyay and Solanky (1994), we conclude as $c \to 0$:

$$\begin{aligned}
&\text{a)} \quad P_{\boldsymbol{\theta}}\{N_{\mathcal{Q}_{1,1}} \le \varepsilon n^*\} = O(n^{*-2m+2}) \text{ if } 0 < \varepsilon < 1,\ m \ge 2; \\
&\text{b)} \quad N^*_{\mathcal{Q}_{1,1}} \equiv n^{*-1/2}(N_{\mathcal{Q}_{1,1}} - n^*) \xrightarrow{\pounds} N(0, \tfrac{1}{4}) \text{ if } m \ge 2; \\
&\text{c)} \quad \left| N^*_{\mathcal{Q}_{1,1}} \right|^r \text{ is uniformly integrable if } m > 1 + \tfrac{1}{4}r,\ r > 0.
\end{aligned} \tag{4.3.21}$$

Then, in the spirit of (4.3.5), we highlight the following asymptotic first-order results as $c \to 0$ when $m \ge 2$:

$$\begin{aligned}
&\text{a)} \quad E_{\boldsymbol{\theta}}[N_{\mathcal{Q}_{1,1}} n^{*-1}] \to 1 \ [\textit{asymptotic first-order efficiency}]; \\
&\text{b)} \quad E_{\boldsymbol{\theta}}[n^* N^{-1}_{\mathcal{Q}_{1,1}}] \to 1; \\
&\text{c)} \quad \kappa_{\mathcal{Q}_{1,1}}(c) \to 1 \ [\textit{asymptotic risk efficiency}];
\end{aligned} \tag{4.3.22}$$

The next set of second-order results follow along the lines of proof given in the case of Theorem 4.3.1. So, we show no further details.

Theorem 4.3.2. *For the MRPE strategy* $\left(N_{\mathcal{Q}_{1,1}}, T_{N_{\mathcal{Q}_{1,1}}}\right)$ *from (4.3.14), we have as* $c \to 0$:

(*i*) $E_{\boldsymbol{\theta}}[N_{\mathcal{Q}_{1,1}}] = n^* + \frac{1}{4}(\nu - 1) + o(1)$ *if* $m \ge 2$ [*asymptotic second-order efficiency*];

(*ii*) $\omega_{\mathcal{Q}_{1,1}}(c) = \frac{1}{2}c + o(c)$ *if* $m \ge 3$ [*asymptotic second-order regret expansion*];

where n^*, $\omega_{\mathcal{Q}_{1,1}}(c)$, *and* ν *come from (4.2.3), (4.2.5) corresponding to the methodology* $\mathcal{Q}_{1,1}$, *and (4.3.3) respectively.*

4.3.4. Asymptotic Second-Order Regret Expansion for $\mathcal{Q}_1 \equiv \mathcal{Q}_{1,j},\ j = 2, ..., 13$

We had previously handled $\mathcal{Q}_{1,0}$, $\mathcal{Q}_{1,1}$, and thus it remains for us to address $\mathcal{Q}_{1,j}$, $j = 2, ..., 13$. Now, all sufficient conditions (C1)-(C7) from Hu and Mukhopadhyay (2019) are easily verified in the context of (4.3.12) when the W_n's are substituted from (4.3.16)-(4.3.17). Hence, in the case of each such associated estimation strategy, we can claim that the second-order regret expansion would be given by:

$$\omega_{\mathcal{Q}_{1,j}}(c) \equiv 2cE_{\boldsymbol{\theta}}\left[\frac{(N_{\mathcal{Q}_{1,j}}-n^*)^2}{N_{\mathcal{Q}_{1,j}}}\right] = \gamma_j c + o(c),\ j = 2, ..., 13, \tag{4.3.23}$$

with appropriate expressions for "γ_j" which can be made precise as well as explicit. To quote such results in (4.3.23), we will not require any appropriate condition(s) on m since we included a term $n^{-\lambda}$ on the right-hand side of our definition of the boundary crossing criterion in (4.3.12). What that does is this: We can claim that $N \geq (A/c)^{1/2(1+\lambda)}$ holds w.p.1 whereas this lower bound also goes to ∞ as $c \to 0$.

Now, we revert back to the basic methodologies, namely $\mathcal{Q}_{1,j},\ j = 0, 1, ..., 5$, from our previous list so that we can explicitly provide respective expressions for "γ_j" under each scenario in view of the recently published work of Hu and Mukhopadhyay (2019). We simply report the following conclusions without going into substantial details from behind the scene:

$$\begin{aligned}
&\text{a)}\quad \omega_{\mathcal{Q}_{1,0}}(c) = \tfrac{1}{2}c + o(c) \text{ if } m \geq 3;\\
&\text{b)}\quad \omega_{\mathcal{Q}_{1,1}}(c) = \tfrac{1}{2}c + o(c) \text{ if } m \geq 3;\\
&\text{c)}\quad \omega_{\mathcal{Q}_{1,2}}(c) = \tfrac{1}{3}(\pi + 6\sqrt{3} - 12)c + o(c) \ \approx 0.511c + o(c);\\
&\text{d)}\quad \omega_{\mathcal{Q}_{1,3}}(c) = \tfrac{1}{2}(\pi - 2)c + o(c) \approx 0.571c + o(c);\\
&\text{e)}\quad \omega_{\mathcal{Q}_{1,4}}(c) = \tfrac{1}{12}(5\pi + 12\sqrt{3} - 30)c + o(c) \approx 0.541c + o(c);\\
&\text{f)}\quad \omega_{\mathcal{Q}_{1,5}}(c) = \tfrac{1}{12}(5\pi + 12\sqrt{3} - 30)c + o(c) \approx 0.541c + o(c).
\end{aligned} \tag{4.3.24}$$

The corresponding finite mathematical expressions of γ_j's associated with the MRPE strategies $\mathcal{Q}_1 \equiv \mathcal{Q}_{1,j},\ j = 6, ..., 13$, exist and they are theoretically easy to comprehend. But, their analytical closed-form expressions are significantly more cumbersome. Hence, instead of going that route, we proceed to carry out simulations.

Table 4.1. Explanation of the set of notations used under $T = 10000$ replications

n_t	: sample size in t^{th} run;
$\overline{n} = T^{-1}\Sigma_{t=1}^{T} n_t$	: should estimate n^*;
$s(\overline{n}) = \left\{(T^2 - T)^{-1}\Sigma_{t=1}^{T}(n_t - \overline{n})^2\right\}^{1/2}$	: estimated standard error (s.e.) of $\overline{n}$;
$s^2_{Pn_t}$	: pooled sample variance in t^{th} run;
$r_{n_t} = As^2_{n_t}/n_t + 2cn_t$	: estimated risk in t^{th} run;
$\overline{r} = T^{-1}\Sigma_{t=1}^{T} r_{n_t}$	: should estimate R_{n^*};
$\widehat{\kappa} = \overline{r}/R_{n^*}$	: should estimate risk efficiency κ;
$s(\widehat{\kappa}) = \left\{(T^2 - T)^{-1}\Sigma_{t=1}^{T}\left(r_{n_t} - \overline{r}\right)^2\right\}^{1/2}/R_{n^*}$	: estimated s.e. of $\widehat{\kappa}$;
$\omega_{n_t} = 2c(n_t - n^*)^2/n_t$	: estimated regret in t^{th} run;
$\widehat{\omega} = T^{-1}\Sigma_{t=1}^{T}\omega_{n_t}$	: should estimate ω;
$s(\widehat{\omega}) = \left\{(T^2 - T)^{-1}\Sigma_{t=1}^{T}\left(\omega_{n_t} - \widehat{\omega}\right)^2\right\}^{1/2}$	: estimated s.e. of $\widehat{\omega}$.
$\gamma_j \equiv \gamma$	: theoretical approx. of $\omega_{\mathcal{Q}_{1,j}}(c), j = 0, 1, ..., 5$.

4.3.5. Simulations

In the spirit of Hu and Mukhopadhyay (2019), we initially implemented each sequential MRPE methodology $\mathcal{Q}_{1,j}, j = 0, 1, ..., 5$ based on the stopping rules given by (4.2.4), (4.3.14), and (4.3.16) respectively. To be more specific, we generated pseudo random samples from two independent populations, $N(5, 4^2)$ and $N(3, 4^2)$. We also fixed the weight function $A = 1$, the pilot sample size $m = 10$, and $\lambda = 2$ as needed, while selecting a wide range of c values including 0.0064, 0.0016 and 0.0001778 so that the optimal fixed sample sizes n^* came out to be 50 (small), 100(moderate) and 300 (large) according to (4.2.3). Throughout this section, we have used a unified range of notations that is set rather precisely as explained in Table 4.1.

Each row in Table 4.2 presents averages of simulated performances obtained from $T(=10000)$ independent replications. As reflected in Table 4.2, while all methodologies $\mathcal{Q}_{1,0}$-$\mathcal{Q}_{1,5}$ oversampled slightly compared with n^*, the estimated average sample sizes ($\overline{n}$) stay close to the optimal sample sizes (n^*), literally within one half of one observation, with small associated values of estimated s.e.'s ($s(\overline{n})$). The strategy $\mathcal{P}_{1,1}$, which involves the unbiased estimator of σ seems to perform worst as it leads to the largest $\overline{n}$ values.

Table 4.2. Simulations from $N(5, 4^2)$ and $N(3, 4^2)$ under $A = 1$, $m = 10$, and $\lambda = 2$ as needed, with 10000 runs implementing $\mathcal{Q}_{1,0}$-$\mathcal{Q}_{1,5}$ from (4.3.16)

n^* c	$\mathcal{Q}$	$\overline{n}$	$s(\overline{n})$	$\widehat{\kappa}$	$s(\widehat{\kappa})$	$\widehat{\omega}$	$s(\widehat{\omega})$	γ	$\widehat{\omega}/c$
50	$\mathcal{Q}_{1,0}$	50.2044	0.03615	0.99227	0.000720	0.003410	0.0000515	0.5	0.533
0.0064	$\mathcal{Q}_{1,1}$	50.4685	0.03603	0.99259	0.000716	0.003388	0.0000518	0.5	0.529
	$\mathcal{Q}_{1,2}$	50.3335	0.03666	0.99238	0.000722	0.003500	0.0000536	0.511	0.547
	$\mathcal{Q}_{1,3}$	50.3347	0.03862	0.99277	0.000723	0.003880	0.0000580	0.571	0.606
	$\mathcal{Q}_{1,4}$	50.3479	0.03792	0.99254	0.000727	0.003747	0.0000571	0.541	0.585
	$\mathcal{Q}_{1,5}$	50.3455	0.03787	0.99262	0.000726	0.003739	0.0000563	0.541	0.584
100	$\mathcal{Q}_{1,0}$	100.2279	0.05029	0.99646	0.000502	0.000816	0.0000117	0.5	0.510
0.0016	$\mathcal{Q}_{1,1}$	100.4922	0.05025	0.99650	0.000501	0.000816	0.0000121	0.5	0.510
	$\mathcal{Q}_{1,2}$	100.3853	0.05094	0.99650	0.000502	0.000835	0.0000119	0.511	0.522
	$\mathcal{Q}_{1,3}$	100.3828	0.05362	0.99668	0.000500	0.000925	0.0000133	0.571	0.578
	$\mathcal{Q}_{1,4}$	100.3889	0.05212	0.99660	0.000499	0.000875	0.0000127	0.541	0.547
	$\mathcal{Q}_{1,5}$	100.3681	0.05211	0.99662	0.000500	0.000877	0.0000125	0.541	0.548
300	$\mathcal{Q}_{1,0}$	300.2097	0.08619	0.99873	0.000287	0.000088	0.0000013	0.5	0.495
0.000178	$\mathcal{Q}_{1,1}$	300.4593	0.08620	0.99873	0.000287	0.000088	0.0000013	0.5	0.495
	$\mathcal{Q}_{1,2}$	300.3387	0.08691	0.99873	0.000286	0.000090	0.0000013	0.511	0.506
	$\mathcal{Q}_{1,3}$	300.3289	0.09213	0.99877	0.000287	0.000101	0.0000015	0.571	0.568
	$\mathcal{Q}_{1,4}$	300.3347	0.08956	0.99876	0.000287	0.000095	0.0000014	0.541	0.534
	$\mathcal{Q}_{1,5}$	300.3377	0.08911	0.99877	0.000286	0.000094	0.0000014	0.541	0.529

Table 4.3. Simulations from $N(5, 10^2)$ and $N(3, 10^2)$ under $A = 1$, $m = 10$, and $\lambda = 2$ as needed, with 10000 runs implementing $\mathcal{Q}_{1,0}$-$\mathcal{Q}_{1,5}$ from (4.3.16)

n^* c	$\mathcal{Q}$	$\overline{n}$	$s(\overline{n})$	$\widehat{\kappa}$	$s(\widehat{\kappa})$	$\widehat{\omega}$	$s(\widehat{\omega})$	γ	$\widehat{\omega}/c$
50	$\mathcal{Q}_{1,0}$	50.2044	0.03615	0.99227	0.000720	0.02131	0.000322	0.5	0.533
0.04	$\mathcal{Q}_{1,1}$	50.4685	0.03603	0.99259	0.000716	0.02118	0.000323	0.5	0.530
	$\mathcal{Q}_{1,2}$	50.3321	0.03666	0.99238	0.000722	0.02186	0.000331	0.511	0.547
	$\mathcal{Q}_{1,3}$	50.3338	0.03880	0.99270	0.000726	0.02447	0.000369	0.571	0.612
	$\mathcal{Q}_{1,4}$	50.3371	0.03785	0.99248	0.000724	0.02329	0.000355	0.541	0.582
	$\mathcal{Q}_{1,5}$	50.3455	0.03794	0.99264	0.000726	0.02347	0.000356	0.541	0.587
100	$\mathcal{Q}_{1,0}$	100.2279	0.05029	0.99646	0.000502	0.00510	0.000073	0.5	0.510
0.01	$\mathcal{Q}_{1,1}$	100.4922	0.05025	0.99650	0.000501	0.00510	0.000076	0.5	0.510
	$\mathcal{Q}_{1,2}$	100.3853	0.05091	0.99651	0.000502	0.00521	0.000074	0.511	0.521
	$\mathcal{Q}_{1,3}$	100.3828	0.05347	0.99669	0.000499	0.00575	0.000083	0.571	0.575
	$\mathcal{Q}_{1,4}$	100.3889	0.05208	0.99660	0.000500	0.00546	0.000079	0.541	0.546
	$\mathcal{Q}_{1,5}$	100.3681	0.05206	0.99664	0.000499	0.00547	0.000078	0.541	0.547
300	$\mathcal{Q}_{1,0}$	300.2097	0.08619	0.99873	0.000287	0.00055	0.000008	0.5	0.495
0.00111	$\mathcal{Q}_{1,1}$	300.4593	0.08620	0.99873	0.000287	0.00055	0.000008	0.5	0.495
	$\mathcal{Q}_{1,2}$	300.3387	0.08692	0.99873	0.000286	0.00056	0.000008	0.511	0.504
	$\mathcal{Q}_{1,3}$	300.3218	0.09181	0.99878	0.000286	0.00063	0.000009	0.571	0.567
	$\mathcal{Q}_{1,4}$	300.3347	0.08955	0.99876	0.000287	0.00059	0.000009	0.541	0.531
	$\mathcal{Q}_{1,5}$	300.3377	0.08908	0.99877	0.000286	0.00059	0.000009	0.541	0.531

In this sense, our methodologies $\mathcal{Q}_{1,0}$ and $\mathcal{Q}_{1,2}$-$\mathcal{Q}_{1,5}$ appear more "robust". The estimated risk efficiency values ($\widehat{\kappa}$) shown in the 5th column are all close to 1, whereas larger (smaller) the sample size (c) lead to $\widehat{\kappa}$ nearly 1. Furthermore, the values of $\widehat{\omega}/c$ provided in the last column (column 10) are all comparable to the theoretical approximation γ (column 9) corresponding to the methodologies $\mathcal{Q}_{1,0}$-$\mathcal{Q}_{1,5}$. Indeed, columns 9-10 begin to compare more favorably as we increased n^* values. That said, each methodology performed remarkably well whereas there was little to no significant differences among their overall characteristics.

We may add one additional note regarding our implementation of the strategy $\mathcal{Q}_{1,1}$ from (4.3.14): We did so first with (i) the exact expression of a_n from (4.3.13), and then (ii) a_n replaced by its large-sample approximation, namely $1 + \frac{1}{8}n^{-1}$, from (4.3.18). We ran simulations of $\mathcal{Q}_{1,1}$ by preserving significantly more decimal places under both (i) and (ii), but no appreciable difference was detected across the board. This may not be particularly surprising, however, because in our attempt to highlight asymptotics and how quickly asymptotics may kick in, we fixed the smallest n^* value 50. In Table 4.2, performances reported under $\mathcal{Q}_{1,1}$ were compiled when we replaced a_n by its large-sample approximation $1 + \frac{1}{8}n^{-1}$ from (4.3.18).

From (4.3.23), we can claim that $V_{\boldsymbol{\theta}}[N_{\mathcal{Q}_{1,j}}] \approx \frac{1}{2}\gamma_j n^*$ which can also be approximated very closely with $\frac{1}{2}\gamma_j \overline{n}$ because $\overline{n}$ is extremely close to n^*. On the other hand, we can certainly claim that $V_{\boldsymbol{\theta}}[N_{\mathcal{Q}_{1,j}}] \approx Ts^2(\overline{n})$ because $T(= 10000)$ is very large, $j = 0, 1, ..., 5$. A thorough examination over Tables 4.2 and 4.3 empirically validates the following approximation:

$$\tfrac{1}{2}\gamma_j \overline{n} \approx Ts^2(\overline{n}), \tag{4.3.25}$$

which speaks to a level of high accuracy of our simulated results.

One may criticize by suggesting that different (larger) common normal variance should be considered to further compare these methodologies $\mathcal{Q}_{1,0}$-$\mathcal{Q}_{1,5}$. So, we decided to increase σ from 4 to 10 keeping A, m, λ the same, picked the c values 0.04, 0.01 and 0.00111 so that the n^* values came out to be 50 (small), 100 (moderate) and 300 (large). The averages from simulated performances with $T = 10000$ replications are summarized in Table 4.3. The overall conclusions drawn from Table 4.3 are no different from those summarized from Table 4.2.

But, we must emphasize the point that both Tables 4.2 and 4.3 show averages from $T = 10000$ independent replications, and there was nothing hidden about that. And yet, if we had expected to see anything particularly special in Table 4.3 given the backdrop from Table 4.2, that may not be entirely fair. Over 10000 replications, even if there were some truly local key features present with very infrequent occurrences, these will be washed out very fast before we can hope to catch them by the huge weight of the majority rule.

Table 4.4. Simulations from $N(5, 10^2)$ and $N(3, 10^2)$ under $A = 1$, $m = 10$, and $\lambda = 2$ as needed, with 10 runs implementing $\mathcal{Q}_{1,0}$-$\mathcal{Q}_{1,5}$ from (4.3.16)

n^* c	$\mathcal{Q}$	$\overline{n}$	$s(\overline{n})$	$\widehat{\kappa}$	$s(\widehat{\kappa})$	$\widehat{\omega}$	$s(\widehat{\omega})$	γ	$\widehat{\omega}/c$
50	$\mathcal{Q}_{1,0}$	50.3	1.29	0.99567	0.025467	0.02422	0.008266	0.5	0.606
0.04	$\mathcal{Q}_{1,1}$	50.5	1.38	0.99182	0.027083	0.02857	0.013231	0.5	0.714
	$\mathcal{Q}_{1,2}$	50.5	1.42	0.99338	0.027248	0.03006	0.012985	0.511	0.752
	$\mathcal{Q}_{1,3}$	50.8	1.50	0.99870	0.027058	0.03114	0.012320	0.571	0.779
	$\mathcal{Q}_{1,4}$	51.1	1.59	1.00123	0.029073	0.03567	0.015229	0.541	0.892
	$\mathcal{Q}_{1,5}$	50.6	1.35	0.99609	0.026229	0.02635	0.009501	0.541	0.659
100	$\mathcal{Q}_{1,0}$	102.1	1.77	1.01568	0.016879	0.00645	0.002228	0.5	0.645
0.01	$\mathcal{Q}_{1,1}$	102.3	1.67	1.01599	0.016780	0.00595	0.001919	0.5	0.595
	$\mathcal{Q}_{1,2}$	102.4	1.66	1.01576	0.017000	0.00588	0.002195	0.511	0.588
	$\mathcal{Q}_{1,3}$	102.7	2.04	1.01269	0.017983	0.00814	0.004119	0.571	0.814
	$\mathcal{Q}_{1,4}$	102.6	1.83	1.01507	0.017031	0.00700	0.002474	0.541	0.700
	$\mathcal{Q}_{1,5}$	102.5	1.93	1.01497	0.017557	0.00752	0.002765	0.541	0.752
300	$\mathcal{Q}_{1,0}$	302.4	1.81	1.00642	0.006132	0.00026	0.000081	0.5	0.234
0.00111	$\mathcal{Q}_{1,1}$	302.7	1.83	1.00637	0.005988	0.00027	0.000089	0.5	0.243
	$\mathcal{Q}_{1,2}$	302.3	1.98	1.00634	0.005790	0.00029	0.000097	0.511	0.261
	$\mathcal{Q}_{1,3}$	302.4	2.76	1.00640	0.006404	0.00054	0.000173	0.571	0.486
	$\mathcal{Q}_{1,4}$	302.0	2.48	1.00645	0.006082	0.00043	0.000116	0.541	0.387
	$\mathcal{Q}_{1,5}$	302.7	2.28	1.00616	0.005874	0.00039	0.000139	0.541	0.351

Table 4.5. Simulations from $N(5,10^2)$ and $N(3,10^2)$ under $A=1$, $m=10$, and $\lambda=2$ as needed, with 10 runs implementing $\mathcal{Q}_{1,0}$-$\mathcal{Q}_{1,5}$ from (4.3.16): terminal sample sizes

n^*	c	$\mathcal{Q}$	n_t
50	0.04	$\mathcal{Q}_{1,0}$	49, 51, 51, 52, 45, 53, 44, 47, 54, 57
		$\mathcal{Q}_{1,1}$	50, 51, 52, 52, 46, 53, 42, 48, 53, 58
		$\mathcal{Q}_{1,2}$	49, 50, 53, 53, 46, 53, 42, 48, 53, 58
		$\mathcal{Q}_{1,3}$	47, 50, 52, 53, 45, 54, 45, 48, 54, 60
		$\mathcal{Q}_{1,4}$	48, 50, 52, 53, 44, 54, 45, 49, 55, 61
		$\mathcal{Q}_{1,5}$	49, 50, 53, 52, 45, 53, 44, 48, 54, 58
100	0.01	$\mathcal{Q}_{1,0}$	99, 105, 98, 90, 109, 106, 99, 105, 106, 104
		$\mathcal{Q}_{1,1}$	99, 105, 99, 91, 109, 106, 99, 105, 106, 104
		$\mathcal{Q}_{1,2}$	99, 105, 99, 92, 111, 104, 99, 104, 106, 105
		$\mathcal{Q}_{1,3}$	97, 105, 99, 94, 116, 101, 97, 105, 106, 107
		$\mathcal{Q}_{1,4}$	98, 105, 98, 92, 112, 104, 99, 105, 107, 106
		$\mathcal{Q}_{1,5}$	97, 105, 98, 92, 113, 105, 98, 105, 106, 106
300	0.00111	$\mathcal{Q}_{1,0}$	301, 308, 308, 310, 308, 296, 302, 298, 294, 299
		$\mathcal{Q}_{1,1}$	301, 308, 309, 310, 309, 297, 302, 297, 295, 299
		$\mathcal{Q}_{1,2}$	301, 308, 309, 311, 309, 298, 298, 295, 295, 299
		$\mathcal{Q}_{1,3}$	298, 310, 311, 316, 310, 294, 298, 291, 294, 302
		$\mathcal{Q}_{1,4}$	298, 310, 310, 312, 310, 297, 296, 294, 291, 302
		$\mathcal{Q}_{1,5}$	300, 308, 310, 314, 309, 296, 300, 292, 296, 302

So, were there any hidden features under the scenario highlighted in Table 4.3 which possibly stayed out of sight? We will never know from 10000 replications. Table 4.4 is a repeat from Table 4.3 except that Table 4.4 summarizes averages from 10 independent replications under each row. Obviously, the entries from Tables 4.3 and 4.4 tend to tell different stories. These also amount to drawing different conclusions from Tables 4.2 and 4.4. We just get different feelings by placing Table 4.2 side-by-side Table 4.4.

To be more specific, we may want to just zoom into 10 observed terminal sample sizes from Table 4.5 corresponding to each row of averages shown earlier in Table 4.4. Table 4.5 shows the sample sizes upon termination under each methodology $\mathcal{Q}_{1,0}$-$\mathcal{Q}_{1,5}$ corresponding to the rows from Table 4. We see that under these 10 runs, while all these methodologies continued to produce comparable final or terminal sample sizes, there are noticeable discrepancies between n^* and $\overline{n}$ values. In particular, when the optimal sample size is moderate (100) or large (300), $\overline{n}$ values exceeded n^* by more than 2 observations. Also, under $\mathcal{Q}_{1,2}$-$\mathcal{Q}_{1,5}$, which are constructed based on unbiased estimators of σ involving GMD and MAD, resulted in larger s.e. values $s(\overline{n})$ for $\overline{n}$ shown in column 5 (Table 4.4). While this may appear bad, $\mathcal{Q}_{1,2}$-$\mathcal{Q}_{1,5}$ produced final sample sizes which were nearer to the optimal sample sizes considering the distance between n^* and $\overline{n}$ after factoring in variability captured by $s(\overline{n})$.

We had looked at side-by-side boxplots prepared from 10 observed terminal sample sizes in each block ($n^* = 50, 100, 300$) within Table 4.5 across the methodologies $\mathcal{Q}_{1,0}$-$\mathcal{Q}_{1,5}$. The distributions of the final sample sizes were clearly right-skewed. For moderate and large sample sizes, the methodologies involving MAD ($\mathcal{Q}_{1,3}$-$\mathcal{Q}_{1,5}$) tended to yield final sample sizes with slightly larger spread. This is consistent with our previous analysis based on $s(\overline{n})$. However, the medians of the final sample sizes from $\mathcal{Q}_{1,2}$-$\mathcal{Q}_{1,5}$ in Table 4.5 turned out to be very close to (or at) the n^* values. In this sense, it may be reasonable to claim that the GMD- or MAD-based methodologies looks practically more robust than the customary methodologies ($\mathcal{Q}_{1,0}$-$\mathcal{Q}_{1,1}$) based on sample standard deviation(s).

Table 4.6. Simulations from $N(5, 4^2)$ and $N(3, 4^2)$ under $A = 1$, $m = 10$, and $\lambda = 2$ as needed, with 10000 runs implementing $\mathcal{Q}_{1,6}$-$\mathcal{Q}_{1,13}$ from (4.3.17)

n^* c	$\mathcal{Q}$	$\overline{n}$	$s(\overline{n})$	$\widehat{\kappa}$	$s(\widehat{\kappa})$	$\widehat{\omega}$	$s(\widehat{\omega})$	$\widehat{\omega}/c$
50	$\mathcal{Q}_{1,6}$	50.3627	0.04089	0.99321	0.000721	0.004346	0.0000651	0.679
0.0064	$\mathcal{Q}_{1,7}$	50.3532	0.04153	0.99332	0.000721	0.004490	0.0000678	0.702
	$\mathcal{Q}_{1,8}$	50.3560	0.04065	0.99312	0.000722	0.004319	0.0000670	0.675
	$\mathcal{Q}_{1,9}$	50.3504	0.04187	0.99303	0.000727	0.004600	0.0000749	0.719
	$\mathcal{Q}_{1,10}$	50.3833	0.03664	0.99230	0.000725	0.003494	0.0000511	0.546
	$\mathcal{Q}_{1,11}$	50.3886	0.03736	0.99255	0.000723	0.003616	0.0000523	0.565
	$\mathcal{Q}_{1,12}$	50.3884	0.03709	0.99240	0.000725	0.003575	0.0000521	0.559
	$\mathcal{Q}_{1,13}$	50.3911	0.03677	0.99258	0.000720	0.003520	0.0000518	0.550
100	$\mathcal{Q}_{1,6}$	100.4155	0.05695	0.99681	0.000500	0.001044	0.0000151	0.653
0.0016	$\mathcal{Q}_{1,7}$	100.4064	0.05841	0.99684	0.000503	0.001011	0.0000157	0.632
	$\mathcal{Q}_{1,8}$	100.3732	0.05643	0.99682	0.000504	0.001028	0.0000149	0.643
	$\mathcal{Q}_{1,9}$	100.3529	0.05697	0.99687	0.000504	0.001047	0.0000148	0.654
	$\mathcal{Q}_{1,10}$	100.4099	0.05069	0.99646	0.000503	0.000829	0.0000118	0.518
	$\mathcal{Q}_{1,11}$	100.4231	0.05158	0.99659	0.000501	0.000857	0.0000124	0.536
	$\mathcal{Q}_{1,12}$	100.4197	0.05118	0.99652	0.000503	0.000844	0.0000120	0.528
	$\mathcal{Q}_{1,13}$	100.4243	0.05102	0.99656	0.000501	0.000837	0.0000118	0.523
300	$\mathcal{Q}_{1,6}$	300.3819	0.09863	0.99878	0.000290	0.000116	0.0000017	0.653
0.000178	$\mathcal{Q}_{1,7}$	300.3849	0.10010	0.99884	0.000289	0.000119	0.0000017	0.669
	$\mathcal{Q}_{1,8}$	300.2707	0.09661	0.99883	0.000288	0.000111	0.0000016	0.624
	$\mathcal{Q}_{1,9}$	300.2920	0.09901	0.99885	0.000290	0.000116	0.0000017	0.653
	$\mathcal{Q}_{1,10}$	300.3739	0.08693	0.99872	0.000288	0.000090	0.0000013	0.506
	$\mathcal{Q}_{1,11}$	300.3751	0.08871	0.99873	0.000287	0.000093	0.0000013	0.523
	$\mathcal{Q}_{1,12}$	300.3722	0.08757	0.99873	0.000288	0.000092	0.0000013	0.518
	$\mathcal{Q}_{1,13}$	300.3708	0.08757	0.99873	0.000286	0.000091	0.0000013	0.512

We also ran simulations for the other selected estimation strategies $\mathcal{O}_{1,6}$-$\mathcal{O}_{1,13}$ coming from (4.3.17) under considered design parameters similar to those that we had incorporated in the constructions of our previous Table 4.2 by averaging performances obtained from 10000 replications. A summary is shown in Table 4.6 much in the same spirit as we did so earlier in Table 4.2. We find that all the methodologies $\mathcal{Q}_{1,6}$-$\mathcal{Q}_{1,13}$ from (4.3.17) perform remarkably well with nearly no substantial differences among them.

4.3.6. Real Data Illustrations

We will illustrate our MRPE strategies $\mathcal{Q}_{1,0}$-$\mathcal{Q}_{1,5}$ with the help of H-1B petition data largely disseminated through the website:

https://www.foreignlaborcert.doleta.gov/performancedata.cfm,

granting public access for data analysis, clearly stated in the following quote from this site. It says, "Disclosure data consists of selected information extracted from non-immigrant and immigrant application tables within the Office of Foreign Labor Certification's case management systems. The data sets provide public access to the latest quarterly and annual data in easily accessible formats for the purpose of performing in-depth longitudinal research and analysis." So, our illustrations and analysis using data from this open-access site are clearly allowed by this disclosure.

H-1B visa is an employment-based, non-immigrant visa category for temporary foreign workers in the United States. The Office of Foreign Labor Certification's iCERT Visa Portal

System updates H-1B petition data every year. This public disclosure file includes administrative date from employer's Labor Condition Application(LCA) and provides key insights into prevailing wages for job titles that are being sponsored by US employers under H-1B category.

Among different job titles, "Programmer Analyst" and "Data Analyst" are the most popular ones. In this illustration, we used three different subsets selected from full dataset and implemented two-sample comparison defined in Section 4.2.1. More specifically, we considered the following three groups:

$$\begin{aligned} &D_1\text{:} \quad \text{prevailing wage of "Sr. Programmer Analyst" in the year 2011;} \\ &D_2\text{:} \quad \text{prevailing wage of "Sr. Programmer Analyst" in the year 2012;} \\ &D_3\text{:} \quad \text{prevailing wage of "Data Analyst" in the year 2011.} \end{aligned} \tag{4.3.26}$$

Since distributions of the prevailing wage tended to be right-skewed, we performed log-transformation on the data, and conducted the Shapiro-normality tests on three full datasets. The associated

p-values of Shapiro-test of normality were: 0.1, 0.2633, and 0.6452,

respectively.

Table 4.7. MRPE of $\delta_{1,3}(\equiv \mu_1 - \mu_3)$ when $\sigma_1 = \sigma_3 = \sigma$

	$n^* = 30$		$n^* = 50$		$n^* = 100$	
Method	n	T_n	n	T_n	n	T_n
$\mathcal{Q}_{1,0}$	33	0.0929	43	0.1432	93	0.0958
$\mathcal{Q}_{1,1}$	39	0.0944	53	0.0823	102	0.0922
$\mathcal{Q}_{1,2}$	31	0.1129	50	0.1010	93	0.1002
$\mathcal{Q}_{1,3}$	26	0.0895	45	0.0878	92	0.1047
$\mathcal{Q}_{1,4}$	31	0.1042	46	0.0846	91	0.0996
$\mathcal{Q}_{1,5}$	29	0.0794	47	0.0840	93	0.0875

Suppose we wanted to estimate the difference of the prevailing wage between Data Analyst (D_3) and Sr. Programmer Analyst (D_1) in the year 2011, that is, we needed to look at the datasets D_1 and D_3.There were 232 and 218 observations in D_1 and D_3, respectively. These two groups shared the same variance (the p-value of F-test was 0.3085).

In order to incorporate the MRPE problem for the difference of means, we randomly sampled $m = 20$ observations from both groups initially and used them as pilot data. Then sequentially, we sampled one observation (without replacement) at-a-time from both groups and checked with the stopping rules $\mathcal{Q}_{1,0}$-$\mathcal{Q}_{1,5}$ defined in (4.2.4) and (4.3.16). Once the sampling terminated, the terminal sample size n along with the sample means of both groups, $\overline{X}_{1n}$ and $\overline{X}_{3n}$, were obtained. The MRPE of $\delta_{1,3}(\equiv \mu_1 - \mu_3)$ was $T_n \equiv \overline{X}_{1n} - \overline{X}_{3n}$. Simple random sampling without replacement may cause very weak dependence within the sequence of gathered observations, but any impact was visibly negligible.

In this illustration, instead of varying c, the cost sampling per unit, we varied the optimal fixed sample size n^* (involving full datasets' common variance σ^2). We never use the magnitude of n^* anywhere in our implementation of the purely sequential estimation strategies $\mathcal{Q}_{1,0}$-$\mathcal{Q}_{1,5}$. We fixed $A = 1$ for all such studies. Table 4.7 provides terminal sample size value (n) along with the terminal estimated value (T_n) for $\delta_{1,3}$ based on a single run each. The values of n^*, which stay unknown throughout, are shown in Table 4.7 to gauge how close (or distant) a single observed n may be to (or from) n^*.

Table 4.7 shows clearly that the estimation results were consistent across different methodologies under consideration. The terminal sample sizes were close to the optimal fixed-sample sizes which allows us some added confidence behind our proposed methodologies even though, strictly speaking, we do not know or utilize the n^* value. $\mathcal{O}_{1,1}$ showed potential oversampling while $\mathcal{Q}_{1,3}$ and $\mathcal{Q}_{1,5}$ tended to stop earlier, but again we should refrain from making too many conclusive statements based on single runs! However, each result corresponds to a single random run only.

Considering the whole datasets D_1, D_3, the true mean difference turned out to be $\delta_{1,3} = 0.0958$, which we treated as completely unknown at the beginning. But, Table 4.7 clearly points in the direction that all MRPE methodologies under consideration did end up with reasonably precise estimates of $\delta_{1,3}$.

4.4. Unequal and Unknown Variances

We now go back to Section 4.2.2 and develop MRPE strategies when the two treatment variances are unknown and unequal. From Section 4.2.2, we recall that we begin with pilot data $\{X_{i1}, X_{i2}, ..., X_{im}, i = 1, 2\}, m \geq 2$. After this initial stage, we proceed with purely sequential sampling on the two treatments under consideration in a parallel fashion, independently of each other. It is possible that sampling from one arm (say, treatment 1) quits first, but sampling from the other arm (treatment 2) may continue until termination of sampling from this arm. We must stop gathering data as soon as sampling from both arms terminates.

That is exactly how the MRPE estimation strategy $\mathcal{Q}_{2,0}$ from (4.2.9) was implemented and we note that we may equivalently express $N_{i,\mathcal{Q}_{2,0}} = J_i + 1$ w.p.1 where we denote:

$$J_i(c) \equiv J_i = \inf\{n \geq m-1: \quad n^3(1+n^{-1})^2 n_i^{*-2} \geq \Sigma_{j=1}^{n} Y_{i,j}\}, \\ \text{the } Y_{i,j}\text{'s being i.i.d. } \chi_1^2,\ i=1,2. \tag{4.4.1}$$

4.4.1. Summary from Nonlinear Renewal Theory

Now, we apply the nonlinear renewal theory from Woodroofe (1977) and we identify with the set of notation from Mukhopadhyay and Solanky (1994, Section 2.4.2):

$$\delta = 3,\ L_0 = 2,\ \theta = 1,\ \tau^2 = 2,\ b = \tfrac{1}{2},\ \beta^* = \tfrac{1}{2},\ h^* = n_i^{*-2},\ n_0^* = n_i^*, \text{ and } p = \tfrac{1}{2}, \tag{4.4.2}$$

and

$$\nu = \tfrac{3}{2} - D \text{ where } D = \Sigma_{n=1}^{\infty} n^{-1} E\left[\max\left(0, \chi_n^2 - 3n\right)\right] \text{ and } \eta = \tfrac{1}{2}\nu - \tfrac{7}{4} = -1 - \tfrac{1}{2}D, \tag{4.4.3}$$

in order to handle J_i from (4.1), $i = 1, 2$.

Up to 6 decimal places, we evaluated:

$$D \approx 0.235365 \Rightarrow \eta = -1 - \tfrac{1}{2}D \approx -1.117683. \tag{4.4.4}$$

From the established properties of J (Mukhopadhyay and Solanky 1994, Theorem 2.4.8), we conclude as $c \to 0$:

$$\begin{aligned} &\text{a)} \quad P_{\boldsymbol{\theta}}\{N_{i,\mathcal{Q}_{2,0}} \le \varepsilon n^*\} = O(n^{*-(m-1)}) \text{ if } 0 < \varepsilon < 1,\ m \ge 2; \\ &\text{b)} \quad N^*_{i,\mathcal{Q}} \equiv n_i^{*-1/2}(N_{i,\mathcal{Q}_{2,0}} - n_i^*) \xrightarrow{\pounds} N(0, \tfrac{1}{2}) \text{ if } m \ge 2; \\ &\text{c)} \quad \left| N^*_{i,\mathcal{Q}_{2,0}} \right|^r \text{ is uniformly integrable if } m > 1 + \tfrac{1}{2}r,\ r > 0; \end{aligned} \tag{4.4.5}$$

$i = 1, 2$.

Next, we may easily show the following first-order results as $c \to 0$:

$$\begin{aligned} &\text{a)} \quad E_{\boldsymbol{\theta}}[N^*_{i,\mathcal{Q}_{2,0}} n_i^{*-1}] \to 1 \text{ if } m \ge 2 \ [\textit{asymptotic first-order efficiency}]; \\ &\text{b)} \quad E_{\boldsymbol{\theta}}[n_i^* N^{-1}_{i,\mathcal{Q}_{2,0}}] \to 1 \text{ if } m \ge 3; \\ &\text{c)} \quad \kappa_{\mathcal{Q}_{2,0}}(c) \to 1 \text{ if } m \ge 3 \ [\textit{asymptotic risk efficiency}]; \end{aligned} \tag{4.4.6}$$

$i = 1, 2$, with the risk efficiency, $\kappa_{\mathcal{Q}_{2,0}}(c)$, coming from (4.2.10).

Again, we can claim:

$$E_{\boldsymbol{\theta}}[N_{i,\mathcal{Q}_{2,0}}] = n^* + \eta + 1 + o(1) = n^* - 0.117683 + o(1) \text{ if } m \ge 3, \tag{4.4.7}$$

with η coming from (4.4.3). The equation (4.4.7) shows that the purely sequential estimation strategy (4.2.9) is asymptotically second-order efficient in the sense of Ghosh and Mukhopadhyay (1981).

Now, we state the asymptotic second-order expansion of the associated regret function defined in (4.2.10) as $c \to 0$:

$$\omega_{\mathcal{Q}_{2,0}}(c) \equiv c\Sigma_{i=1}^{2} E_{\boldsymbol{\theta}} \left[\frac{(N_{i,\mathcal{Q}_{2,0}} - n_i^*)^2}{N_{i,\mathcal{Q}_{2,0}}} \right] = c + o(c) \text{ if } m \geq 4, \tag{4.4.8}$$

after combining parts (a), (b) and (c) from (4.4.5).

4.4.2. Stopping Rules with Selected Combinations from $\mathbf{S_{in}}$, GMD, and MAD

In the spirits of (4.2.9) and (4.3.12), let us begin by putting forward a kind of generic purely sequential estimation strategy as follows:

$$\text{Methodology } \mathcal{Q}_2 \begin{cases} \text{The total sample size } N_{\mathcal{Q}_2}(c) \equiv N_{\mathcal{Q}_2} = N_{1,\mathcal{Q}_2} + N_{2,\mathcal{Q}_2} \text{ where} \\ N_{i,\mathcal{Q}_2} = \inf \left\{ n \geq m : \quad n \geq (Ac^{-1})^{1/2} \left(W_{in} + n^{-\lambda} \right) \right\}, \, \lambda > \frac{1}{2}, \text{ and} \\ \text{estimating } \delta \text{ with a terminal estimator, } T_{\mathbf{N}_{\mathcal{Q}_2}} (\equiv T_{\mathbf{N}} = \overline{X}_{1N_{1,\mathcal{Q}_2}} - \overline{X}_{2N_{2,\mathcal{Q}_2}}), \\ \text{from fully accrued data } \{N_{i,\mathcal{Q}_2}, X_{i1}, X_{i2}, ..., X_{iN_{i,\mathcal{Q}_2}}, i = 1, 2\}. \end{cases} \tag{4.4.9}$$

Here, "W_{in}" stands for a consistent estimator of σ_i. Indeed, under the methodology $\mathcal{Q}_2$, we expect $P_{\boldsymbol{\theta}}\{N_{\mathcal{Q}_2} < \infty\} = 1$ to hold, so that the terminal estimator $T_{\mathbf{N}_{\mathcal{Q}_2}}$ will be meaningful. Recall that (4.2.9) gave the methodology $\mathcal{Q}_2 \equiv \mathcal{Q}_{2,0}$ with $W_{in} = S_{in}$ which was consistent, but biased, for σ_i. Note, however, that in the definition of $\mathcal{Q}_{2,0}$, we did not use a fudge factor such as $n^{-\lambda}$.

With $n \geq 2$, we can easily verify:

$$E_{\boldsymbol{\theta}} \left[a_n S_{in} \right] = \sigma_i \text{ where } a_n = (\tfrac{1}{2}n - \tfrac{1}{2})^{1/2} \Gamma \left(\tfrac{1}{2}n - \tfrac{1}{2} \right) \left\{ \Gamma \left(\tfrac{1}{2}n \right) \right\}^{-1}. \tag{4.4.10}$$

Then, the original estimation strategy (4.2.9) may be modified as follows:

Methodology $\mathcal{Q}_{2,1}$: The total sample size $N_{\mathcal{Q}_{2,1}}(c) \equiv N = N_{1,\mathcal{Q}_{2,1}} + N_{2,\mathcal{Q}_{2,1}}$ where

$$N_{i,\mathcal{Q}_{2,1}}(c) \equiv N_i = \inf\{n \geq m : \quad n \geq (Ac^{-1})^{1/2} a_n S_{in}\},\ i = 1, 2. \tag{4.4.11}$$

with the terminal estimator $T_{\mathbf{N}}(= \overline{X}_{1N_{1,\mathcal{Q}_{2,1}}} - \overline{X}_{2N_{2,\mathcal{Q}_{2,1}}})$ for δ based on fully accrued data upon termination:

$$\{N_{i,\mathcal{Q}_{2,1}}, X_{i1}, X_{i2}, ..., X_{iN_{i,\mathcal{Q}_{2,1}}}, i = 1, 2\}.$$

Next, in order to explore other tangible choices of W_{in_i}, we recall (4.3.15) and modify notations as follows:

$$G_{in_i} = \frac{1}{2}\sqrt{\pi}\binom{n_i}{2}^{-1} \underset{1 \leq j < k \leq n_i}{\Sigma\Sigma} |X_{ij} - X_{ik}|, \text{ and } M_{in_i} = \sqrt{\frac{\pi n_i}{2(n_i-1)}} n_i^{-1} \Sigma_{k=1}^{n_i} \left|X_{ik} - \overline{X}_{in_i}\right|, \tag{4.4.12}$$

with $i = 1, 2$. The G_{in_i}'s and M_{in_i}'s from (4.4.12) respectively correspond to unbiased and consistent estimators of σ_i derived from GMD and MAD respectively based on random samples from the i^{th} treatment, $i = 1, 2$.

We will successively substitute W_{in} in (4.4.9), the generic statistic that is used to define the first-time boundary crossing, with the following unbiased and consistent estimators of

σ_i's. That way, we will introduce a rather large array of newer MRPE strategies:

$$
\begin{array}{ll}
\text{Methodology } \mathcal{Q}_2 \equiv \mathcal{Q}_{2,0}: & \text{Implement (4.2.9);} \\
\text{Methodology } \mathcal{Q}_2 \equiv \mathcal{Q}_{2,1}: & \text{Replace } S_{in} \text{ in (4.2.9) with } a_n S_{in}, \text{ with } a_n \text{ from (4.4.10);} \\
\text{Methodology } \mathcal{Q}_2 \equiv \mathcal{Q}_{2,2}: & \text{Replace } W_{in} \text{ in (4.4.9) with } G_{in} \text{ from (4.4.12);} \\
\text{Methodology } \mathcal{Q}_2 \equiv \mathcal{Q}_{2,3}: & \text{Replace } W_{in} \text{ in (4.4.9) with } M_{in} \text{ from (4.4.12);} \\
\text{Methodology } \mathcal{Q}_2 \equiv \mathcal{Q}_{2,4}: & \text{Replace } W_{1n} \text{ with } G_{1n}, W_{2n} \text{ with } M_{2n} \text{ from (4.4.12);} \\
\text{Methodology } \mathcal{Q}_2 \equiv \mathcal{Q}_{2,5}: & \text{Replace } W_{1n} \text{ with } M_{1n}, W_{2n} \text{ with } G_{2n} \text{ from (4.4.12).}
\end{array} \tag{4.4.13}
$$

Theoretically, we surely could include other estimators of σ_i's in the spirit of those cited in (4.3.17), but for brevity, we choose not to. The reason for such non-inclusion is driven by the fact that such more complicated convex combinations did not lead to any drastically different and/or interesting feature in comparison with those from (4.3.16). Data analysis from Section 4.3.5 supported such a general sentiment.

The estimation strategies from (4.4.9) are all perfectly logical in the sense that the result $P_{\boldsymbol{\theta}}\{N_{\mathcal{Q}_2} < \infty\} = 1$ holds while using W_{in}'s from (4.4.13). Obviously, $I(N_{\mathcal{Q}_2} = n)$ and T_n, are independently distributed for all fixed $n \geq m$. One may additionally refer to Mukhopadhyay (2000, pp. 324-327).

All sufficient conditions (C1)-(C7) from Hu and Mukhopadhyay (2019) are easily verified in the context of (4.4.13). Hence, in the case of each associated methodology from (4.4.13), we are able to obtain the asymptotic second-order expansion of the regret function, namely,

$$\omega_{\mathcal{Q}_{2,j}}(c) \equiv c\Sigma_{i=1}^{2} E_{\boldsymbol{\theta}}\left[\frac{(N_{i,\mathcal{Q}_{2,j}}-n_i^*)^2}{N_{i,\mathcal{Q}_{2,j}}}\right] = \gamma_j c + o(c) \begin{cases} \text{if } m \geq 4 & \text{when } j = 0,1 \\ \text{if } m \geq 2 & \text{when } j = 2,...,5, \end{cases} \tag{4.4.14}$$

with appropriate expressions for "γ_j" which can be made explicit.

In the contexts of $\mathcal{Q}_{2,j}, j = 2,...,5$, we will not require any additional condition on m beyond $m \geq 2$ since we included a term $n^{-\lambda}$ on the right-hand side of our definition of the boundary crossing criterion in (4.4.9). What that does is this: We can claim that $N \geq (A/c)^{1/2(1+\lambda)}$ holds w.p.1 whereas this lower bound also goes to ∞ as $c \to 0$.

Now, we revert back to the basic methodologies, namely $\mathcal{Q}_{2,j}$, $j = 0,1,...,5$, from our previous list in (4.4.13) so that we can explicitly provide respective expressions for "γ_j" under each scenario in view of the recently published work of Hu and Mukhopadhyay (2019). We simply report the following conclusions without going into substantial details from behind the scene:

$$\begin{aligned} &\text{a)} \quad \omega_{\mathcal{Q}_{2,0}}(c) = c + o(c) \text{ if } m \geq 4; \\ &\text{b)} \quad \omega_{\mathcal{Q}_{2,1}}(c) = c + o(c) \text{ if } m \geq 4; \\ &\text{c)} \quad \omega_{\mathcal{Q}_{2,2}}(c) = \tfrac{1}{3}(2\pi + 12\sqrt{3} - 24)c + o(c) \approx 1.023c + o(c); \\ &\text{d)} \quad \omega_{\mathcal{Q}_{2,3}}(c) = (\pi - 2)\, c + o(c) \approx 1.142c + o(c); \\ &\text{e)} \quad \omega_{\mathcal{Q}_{2,4}}(c) = \tfrac{1}{6}(5\pi + 12\sqrt{3} - 30)c + o(c) \approx 1.082c + o(c); \\ &\text{f)} \quad \omega_{\mathcal{Q}_{2,5}}(c) = \tfrac{1}{6}(5\pi + 12\sqrt{3} - 30)c + o(c) \approx 1.082c + o(c). \end{aligned} \tag{4.4.15}$$

4.4.3. Simulations

Analogous to our approach explained in Section 4.3.5, we went ahead and implemented each MRPE strategy $\mathcal{Q}_{2,j}, j = 0,1,...,5$ based on various stopping rules given by (4.4.13) respectively in the normal case. More specifically, we generated pseudorandom samples from two independent populations, $N(5, 2^2)$ and $N(3, 3^2)$. We fixed $A = 1$, the pilot sample size $m = 10$, and $\lambda = 2$ whenever needed, as we went through a selection of c values 0.0016, 0.0004 and 0.0000444 so that we could determine the optimal fixed sample sizes $(n_1^*, n_2^*, n^*) = (50, 75, 125)$, $(100, 150, 250)$ and $(300, 450, 750)$ according to (2.8).

Table 4.8(a). Simulations from $N(5, 2^2)$ and $N(3, 3^2)$ under $A = 1$, $m = 10$, and $\lambda = 2$ as needed with 10000 runs implementing $\mathcal{Q}_{2,0}$-$\mathcal{Q}_{2,5}$ from (4.4.13)

$\mathbf{n}^* = (n_1^*, n_2^*)$ c	$\mathcal{Q}$	$\overline{n}_1$ $s(\overline{n}_1)$	$\overline{n}_2$ $s(\overline{n}_2)$	$\widehat{\kappa}$ $s(\widehat{\kappa})$	$\widehat{\omega}$ $s(\widehat{\omega})$	γ $\widehat{\omega}/c$
$(50, 75)$	$\mathcal{Q}_{2,0}$	49.8192	74.8701	0.98752	0.001849	1
0.0016		0.05278	0.06300	0.001344	0.0000226	1.156
	$\mathcal{Q}_{2,1}$	50.1057	75.1296	0.98787	0.001785	1
		0.05184	0.06298	0.001330	0.0000214	1.115
	$\mathcal{Q}_{2,2}$	50.1158	75.1473	0.98779	0.001848	1.023
		0.05296	0.06360	0.001340	0.0000220	1.155
	$\mathcal{Q}_{2,3}$	50.1317	75.1428	0.98820	0.002029	1.142
		0.05494	0.06740	0.001339	0.0000243	1.268
	$\mathcal{Q}_{2,4}$	50.1158	75.1428	0.98786	0.001959	1.082
		0.05296	0.06740	0.001345	0.0000236	1.224
	$\mathcal{Q}_{2,5}$	50.1317	75.1473	0.98812	0.001918	1.082
		0.05494	0.06360	0.001334	0.0000231	1.199
$(100, 150)$	$\mathcal{Q}_{2,0}$	99.8528	149.9171	0.99405	0.000421	1
0.0004		0.07182	0.08743	0.000924	0.0000046	1.053
	$\mathcal{Q}_{2,1}$	100.1033	150.1817	0.99408	0.000419	1
		0.07201	0.08714	0.000924	0.0000045	1.047
	$\mathcal{Q}_{2,2}$	100.1156	150.2004	0.99416	0.000429	1.023
		0.07327	0.08787	0.000926	0.0000045	1.073

Table 4.8(b). Simulations from $N(5, 2^2)$ and $N(3, 3^2)$ under $A = 1$, $m = 10$, and $\lambda = 2$ as needed with 10000 runs implementing $\mathcal{Q}_{2,0}$-$\mathcal{Q}_{2,5}$ from (4.4.13)

$\mathbf{n}^* = (n_1^*, n_2^*)$ c	$\mathcal{Q}$	$\overline{n}_1$ $s(\overline{n}_1)$	$\overline{n}_2$ $s(\overline{n}_2)$	$\widehat{\kappa}$ $s(\widehat{\kappa})$	$\widehat{\omega}$ $s(\widehat{\omega})$	γ $\widehat{\omega}/c$
$(100, 150)$	$\mathcal{Q}_{2,3}$	100.1076	150.1684	0.99435	0.000478	1.142
0.0004		0.07702	0.09320	0.000928	0.0000050	1.194
	$\mathcal{Q}_{2,4}$	100.1156	150.1684	0.99425	0.000455	1.082
		0.07327	0.09320	0.000927	0.0000048	1.137
	$\mathcal{Q}_{2,5}$	100.1076	150.2004	0.99426	0.000452	1.082
		0.07702	0.08787	0.000927	0.0000048	1.130
$(300, 450)$	$\mathcal{Q}_{2,0}$	299.7950	449.8081	0.99777	0.000044	1
0.0000444		0.12263	0.14912	0.000526	0.0000005	1.003
	$\mathcal{Q}_{2,1}$	300.0422	450.0415	0.99776	0.000045	1
		0.12308	0.15021	0.000529	0.0000005	1.013
	$\mathcal{Q}_{2,2}$	300.0604	450.0602	0.99777	0.000046	1.023
		0.12415	0.15112	0.000527	0.0000005	1.026
	$\mathcal{Q}_{2,3}$	300.0436	450.0282	0.99785	0.000051	1.142
		0.13102	0.15979	0.000528	0.0000005	1.147
	$\mathcal{Q}_{2,4}$	300.0604	450.0282	0.99782	0.000048	1.082
		0.12415	0.15979	0.000528	0.0000005	1.087
	$\mathcal{Q}_{2,5}$	300.0436	450.0602	0.99781	0.000048	1.082
		0.13102	0.15112	0.000527	0.0000005	1.086

Table 4.9(a). Simulations from $N(5, 2^2)$ and $N(3, 8^2)$ under $A = 1$, $m = 10$, and $\lambda = 2$ as needed with 10000 runs implementing $\mathcal{Q}_{2,0}$-$\mathcal{Q}_{2,5}$ from (4.4.13)

$\mathbf{n}^* = (n_1^*, n_2^*)$ c	$\mathcal{Q}$	$\overline{n}_1$ $s(\overline{n}_1)$	$\overline{n}_2$ $s(\overline{n}_2)$	$\widehat{\kappa}$ $s(\widehat{\kappa})$	$\widehat{\omega}$ $s(\widehat{\omega})$	γ $\widehat{\omega}/c$
$(50, 200)$	$\mathcal{Q}_{2,0}$	49.8192	199.9278	0.99400	0.001779	1
0.0016		0.05278	0.10018	0.001164	0.0000211	1.112
	$\mathcal{Q}_{2,1}$	50.1057	200.1781	0.99409	0.001728	1
		0.05184	0.10065	0.001150	0.0000201	1.080
	$\mathcal{Q}_{2,2}$	50.1158	200.2126	0.99413	0.001778	1.023
		0.05296	0.10100	0.001159	0.0000211	1.112
	$\mathcal{Q}_{2,3}$	50.1317	200.1968	0.99440	0.001947	1.142
		0.05494	0.10683	0.001153	0.0000229	1.217
	$\mathcal{Q}_{2,4}$	50.1158	200.1968	0.99423	0.001877	1.082
		0.05296	0.10683	0.001159	0.0000220	1.173
	$\mathcal{Q}_{2,5}$	50.1317	200.2126	0.99429	0.001848	1.082
		0.05494	0.10100	0.001152	0.0000220	1.155
$(100, 400)$	$\mathcal{Q}_{2,0}$	99.8528	399.8625	0.99689	0.000412	1
0.0004		0.07182	0.14074	0.000799	0.0000044	1.031
	$\mathcal{Q}_{2,1}$	100.1033	400.0918	0.99685	0.000416	1
		0.07201	0.14214	0.000802	0.0000045	1.041
	$\mathcal{Q}_{2,2}$	100.1156	400.1136	0.99689	0.000425	1.023
		0.07327	0.14261	0.000805	0.0000045	1.062

Table 4.9(b). Simulations from $N(5, 2^2)$ and $N(3, 8^2)$ under $A = 1$, $m = 10$, and $\lambda = 2$ as needed with 10000 runs implementing $\mathcal{Q}_{2,0}$-$\mathcal{Q}_{2,5}$ from (4.4.13)

$\mathbf{n}^* = (n_1^*, n_2^*)$ c	$\mathcal{Q}$	$\overline{n}_1$ $s(\overline{n}_1)$	$\overline{n}_2$ $s(\overline{n}_2)$	$\widehat{\kappa}$ $s(\widehat{\kappa})$	$\widehat{\omega}$ $s(\widehat{\omega})$	γ $\widehat{\omega}/c$
$(100, 400)$	$\mathcal{Q}_{2,3}$	100.1076	400.0842	0.99701	0.000469	1.142
0.0004		0.07702	0.14991	0.000805	0.0000049	1.172
	$\mathcal{Q}_{2,4}$	100.1156	400.0842	0.99697	0.000446	1.082
		0.07327	0.14991	0.000805	0.0000047	1.115
	$\mathcal{Q}_{2,5}$	100.1076	400.1136	0.99694	0.000448	1.082
		0.07702	0.14261	0.000806	0.0000048	1.119
$(300, 1200)$	$\mathcal{Q}_{2,0}$	299.7950	1199.943	0.99898	0.000045	1
0.0000444		0.12263	0.24482	0.000457	0.0000005	1.006
	$\mathcal{Q}_{2,1}$	300.0422	1200.186	0.99897	0.000045	1
		0.12308	0.24629	0.000458	0.0000005	1.016
	$\mathcal{Q}_{2,2}$	300.0604	1200.252	0.99899	0.000046	1.023
		0.12415	0.24740	0.000457	0.0000005	1.027
	$\mathcal{Q}_{2,3}$	300.0436	1200.315	0.99902	0.000051	1.142
		0.13102	0.26275	0.000458	0.0000005	1.152
	$\mathcal{Q}_{2,4}$	300.0604	1200.315	0.99900	0.000049	1.082
		0.12415	0.26275	0.000457	0.0000005	1.093
	$\mathcal{Q}_{2,5}$	300.0436	1200.252	0.99901	0.000048	1.082
		0.13102	0.24740	0.000457	0.0000005	1.087

Table 4.8 presents summaries from our simulated performances obtained from $T(= 10000)$ replications in the construction of each row. The notations used are analogous to those from Table 1, but with obvious minor improvisations. The methodology $\mathcal{Q}_{2,0}$ appeared to undersample slightly. All other methodologies $\mathcal{Q}_{2,1}$-$\mathcal{Q}_{2,5}$ oversampled slightly, but the average

estimated sample sizes $\overline{n}_1, \overline{n}_2$ clearly lie within a narrow band of the optimal fixed sample sizes n_1^*, n_2^* respectively.

The estimated risk efficiency values $\widehat{\kappa}$ shown in column 5 all appear close to 1 whereas the values of $\widehat{\omega}$, the estimated regrets, from column 7 came out reasonably close to their corresponding theoretical approximations from (4.4.15) provided by γ. As expected, larger (or smaller) the sample size (or c) is, the closer is $\widehat{\kappa}$ (or $\widehat{\omega}/c$) is to 1 (or γ). Overall, there is little to nearly no significant difference among these methodologies $\mathcal{Q}_{2,0}$-$\mathcal{Q}_{2,5}$ from (4.4.13).

One may suggest that we should have fixed more disparate population standard deviations and compared these methodologies $\mathcal{Q}_{2,0}$-$\mathcal{Q}_{2,5}$. In that vein, we fixed $\sigma_2 = 8$ instead of 3, but continued with the other choices made earlier: $A = 1$, $m = 10$, $\mu_1 = 5$, $\sigma_1 = 2$, $\mu_2 = 3$, and $c = 0.0016, 0.0004$, and 0.0000444 so that $\mathbf{n}^*$ varied through $(50, 200)$, $(100, 400)$ and $(300, 1200)$, respectively. Table 4.9 summarizes average performances in each row under 10000 replications. Overall sentiments gained from Tables 4.8 and 4.9 are very similar.

A large number of replications may tend to overly smooth subtle local features. Tables 4.9 and 4.10 may superficially look similar, but Table 4.10 shows average performances from 10 independent replications only. In the spirit of Table 4.5, our Table 4.11 shows the sample size pairs upon termination under each methodology $\mathcal{Q}_{2,0}$-$\mathcal{Q}_{2,5}$ corresponding to the rows from Table 4.10. When only 10 replications were run, we saw a tiny bit of difference in the methodologies $\mathcal{Q}_{2,0}$ and $\mathcal{Q}_{2,1}$, no matter whether $n^*(= n_1^* + n_2^*)$ was small, moderate or large. Tables 4.12 and 4.13 report average performances from 10 independent replications only similar to those summarized in Tables 4.10 and 4.11, but now we increased σ_2 to 100.

Table 4.10(a). Simulations from $N(5, 2^2)$ and $N(3, 8^2)$ under $A = 1$, $m = 10$, and $\lambda = 2$ as needed with 10 runs implementing $\mathcal{Q}_{2,0}$-$\mathcal{Q}_{2,5}$ from (4.4.13)

$\mathbf{n}^* = (n_1^*, n_2^*)$ c	$\mathcal{Q}$	$\overline{n}_1$ $s(\overline{n}_1)$	$\overline{n}_2$ $s(\overline{n}_2)$	$\widehat{\kappa}$ $s(\widehat{\kappa})$	$\widehat{\omega}$ $s(\widehat{\omega})$	γ $\widehat{\omega}/c$
$(50, 200)$	$\mathcal{Q}_{2,0}$	49.8	204.0	1.00894	0.001660	1
0.0016		1.356	3.724	0.033056	0.0004084	1.038
	$\mathcal{Q}_{2,1}$	49.8	204.0	1.00908	0.001651	1
		1.340	3.724	0.032929	0.0004047	1.032
	$\mathcal{Q}_{2,2}$	49.8	204.7	1.00827	0.001605	1.023
		1.420	3.474	0.032159	0.0003594	1.003
	$\mathcal{Q}_{2,3}$	49.3	204.9	1.00808	0.001985	1.142
		1.680	3.787	0.031370	0.0004650	1.240
	$\mathcal{Q}_{2,4}$	49.8	204.9	1.00782	0.001747	1.082
		1.420	3.787	0.032429	0.0004541	1.092
	$\mathcal{Q}_{2,5}$	49.3	204.7	1.00853	0.001843	1.082
		1.680	3.474	0.031090	0.0003731	1.152
$(100, 400)$	$\mathcal{Q}_{2,0}$	99.9	406.5	1.01026	0.000484	1
0.0004		2.496	4.911	0.027733	0.0001286	1.211
	$\mathcal{Q}_{2,1}$	99.9	406.6	1.01016	0.000487	1
		2.496	4.924	0.027726	0.0001284	1.217
	$\mathcal{Q}_{2,2}$	100.7	407.8	1.01155	0.000466	1.023
		2.629	4.255	0.027683	0.0001228	1.165

Table 4.10(b). Simulations from $N(5, 2^2)$ and $N(3, 8^2)$ under $A = 1$, $m = 10$, and $\lambda = 2$ as needed with 10 runs implementing $\mathcal{Q}_{2,0}$-$\mathcal{Q}_{2,5}$ from (4.4.13)

$\mathbf{n}^* = (n_1^*, n_2^*)$ c	$\mathcal{Q}$	$\overline{n}_1$ $s(\overline{n}_1)$	$\overline{n}_2$ $s(\overline{n}_2)$	$\widehat{\kappa}$ $s(\widehat{\kappa})$	$\widehat{\omega}$ $s(\widehat{\omega})$	γ $\widehat{\omega}/c$
$(100, 400)$	$\mathcal{Q}_{2,3}$	100.9	408.2	1.01153	0.000532	1.142
0.0004		2.861	4.614	0.027327	0.0001471	1.329
	$\mathcal{Q}_{2,4}$	100.7	408.2	1.01148	0.000497	1.082
		2.629	4.614	0.027910	0.0001378	1.243
	$\mathcal{Q}_{2,5}$	100.9	407.8	1.01160	0.000500	1.082
		2.861	4.255	0.027096	0.0001331	1.250
$(300, 1200)$	$\mathcal{Q}_{2,0}$	306.5	1212.6	1.01212	0.000042	1
0.0000444		3.481	6.725	0.012842	0.0000115	0.938
	$\mathcal{Q}_{2,1}$	306.8	1213.3	1.01228	0.000046	1
		3.732	7.018	0.013718	0.0000128	1.043
	$\mathcal{Q}_{2,2}$	307.6	1213.7	1.01218	0.000052	1.023
		3.685	7.844	0.014193	0.0000140	1.163
	$\mathcal{Q}_{2,3}$	308.2	1213.8	1.01220	0.000055	1.142
		3.530	8.434	0.013806	0.0000148	1.231
	$\mathcal{Q}_{2,4}$	307.6	1213.8	1.01216	0.000055	1.082
		3.685	8.434	0.014214	0.0000148	1.235
	$\mathcal{Q}_{2,5}$	308.2	1213.7	1.01223	0.000052	1.082
		3.530	7.844	0.013784	0.0000140	1.159

Table 4.11(a). Simulations from $N(5, 2^2)$ and $N(3, 8^2)$ under $A = 1$, $m = 10$,and $\lambda = 2$ as needed with 10 runs implementing (4.4.13): terminal sample sizes

(n_1^*, n_2^*) c	$\mathcal{Q}$	(n_{1t}, n_{2t})
$(50, 200)$	$\mathcal{Q}_{2,0}$	$(45, 197), (53, 209), (53, 195), (46, 179), (50, 218)$
0.0016		$(55, 214), (49, 197), (56, 211), (47, 211), (44, 209)$
	$\mathcal{Q}_{2,1}$	$(45, 197), (53, 209), (54, 195), (46, 179), (50, 218)$
		$(55, 214), (49, 197), (55, 211), (47, 211), (44, 209)$
	$\mathcal{Q}_{2,2}$	$(45, 196), (53, 210), (54, 196), (45, 183), (50, 221)$
		$(55, 210), (49, 199), (56, 210), (47, 212), (44, 210)$
	$\mathcal{Q}_{2,3}$	$(44, 192), (48, 209), (56, 194), (45, 186), (51, 226)$
		$(55, 208), (46, 199), (58, 210), (46, 213), (44, 212)$
	$\mathcal{Q}_{2,4}$	$(45, 192), (53, 209), (54, 194), (45, 186), (50, 226)$
		$(55, 208), (49, 199), (56, 210), (47, 213), (44, 212)$
	$\mathcal{Q}_{2,5}$	$(44, 196), (48, 210), (56, 196), (45, 183), (51, 221)$
		$(55, 210), (46, 199), (58, 210), (46, 212), (44, 210)$
$(100, 400)$	$\mathcal{Q}_{2,0}$	$(98, 405), (100, 375), (104, 431), (103, 402), (91, 423)$
0.0004		$(105, 406), (87, 406), (92, 413), (107, 392), (112, 412)$
	$\mathcal{Q}_{2,1}$	$(98, 405), (100, 375), (104, 431), (103, 402), (91, 423)$
		$(105, 406), (87, 406), (92, 413), (107, 392), (112, 413)$
	$\mathcal{Q}_{2,2}$	$(97, 407), (100, 379), (105, 430), (105, 408), (90, 421)$
		$(105, 413), (88, 408), (95, 405), (107, 398), (115, 409)$

Table 4.11(b). Simulations from $N(5, 2^2)$ and $N(3, 8^2)$ under $A = 1$, $m = 10$,and $\lambda = 2$ as needed with 10 runs implementing (4.4.13): terminal sample sizes

(n_1^*, n_2^*) c	$\mathcal{Q}$	(n_{1t}, n_{2t})
$(100, 400)$	$\mathcal{Q}_{2,3}$	$(92, 401), (100, 379), (104, 433), (105, 409), (90, 425)$
0.0004		$(104, 415), (94, 408), (92, 402), (110, 403), (118, 407)$
	$\mathcal{Q}_{2,4}$	$(97, 401), (100, 379), (105, 433), (105, 409), (90, 425)$
		$(105, 415), (88, 408), (95, 402), (107, 403), (115, 407)$
	$\mathcal{Q}_{2,5}$	$(92, 407), (100, 379), (104, 430), (105, 408), (90, 421)$
		$(104, 413), (94, 408), (92, 405), (110, 398), (118, 409)$
$(300, 1200)$	$\mathcal{Q}_{2,0}$	$(301, 1216), (299, 1234), (288, 1210), (325, 1198), (306, 1185)$
0.0000444		$(312, 1225), (299, 1234), (314, 1190), (302, 1190), (319, 1244)$
	$\mathcal{Q}_{2,1}$	$(301, 1216), (299, 1236), (288, 1209), (325, 1198), (306, 1185)$
		$(312, 1225), (300, 1235), (319, 1192), (298, 1189), (320, 1248)$
	$\mathcal{Q}_{2,2}$	$(301, 1220), (301, 1239), (291, 1212), (326, 1185), (306, 1186)$
		$(312, 1226), (302, 1238), (321, 1192), (296, 1188), (320, 1251)$
	$\mathcal{Q}_{2,3}$	$(297, 1219), (299, 1249), (296, 1209), (324, 1178), (309, 1189)$
		$(311, 1225), (304, 1240), (323, 1194), (298, 1186), (321, 1249)$
	$\mathcal{Q}_{2,4}$	$(301, 1219), (301, 1249), (291, 1209), (326, 1178), (306, 1189)$
		$(312, 1225), (302, 1240), (321, 1194), (296, 1186), (320, 1249)$
	$\mathcal{Q}_{2,5}$	$(297, 1220), (299, 1239), (296, 1212), (324, 1185), (309, 1186)$
		$(311, 1226), (304, 1238), (323, 1192), (298, 1188), (321, 1251)$

Table 4.12(a). Simulations from $N(5, 2^2)$ and $N(3, 100^2)$ under $A = 1$, $m = 10$, and $\lambda = 2$ as needed with 10 runs implementing $\mathcal{Q}_{2,0}$-$\mathcal{Q}_{2,5}$ from (4.4.13)

$\mathbf{n}^* = (n_1^*, n_2^*)$ c	$\mathcal{Q}$	$\overline{n}_1$ $s(\overline{n}_1)$	$\overline{n}_2$ $s(\overline{n}_2)$	$\widehat{\kappa}$ $s(\widehat{\kappa})$	$\widehat{\omega}$ $s(\widehat{\omega})$	γ $\widehat{\omega}/c$
$(50, 2500)$	$\mathcal{Q}_{2,0}$	49.8	2508.3	1.00265	0.001168	1
0.0016		1.356	10.091	0.027649	0.0002867	0.730
	$\mathcal{Q}_{2,1}$	49.8	2508.6	1.00261	0.001175	1
		1.340	10.207	0.027510	0.0002829	0.734
	$\mathcal{Q}_{2,2}$	49.8	2508.8	1.00257	0.001310	1.023
		1.420	10.771	0.027598	0.0003750	0.819
	$\mathcal{Q}_{2,3}$	49.3	2510.8	1.00250	0.001807	1.142
		1.680	12.496	0.026374	0.0004870	1.129
	$\mathcal{Q}_{2,4}$	49.8	2510.8	1.00248	0.001569	1.082
		1.420	12.496	0.027626	0.0004766	0.981
	$\mathcal{Q}_{2,5}$	49.3	2508.8	1.00259	0.001547	1.082
		1.680	10.771	0.026345	0.0003520	0.967
$(100, 5000)$	$\mathcal{Q}_{2,0}$	99.9	5024.8	1.00459	0.000375	1
0.0004		2.496	11.535	0.024926	0.0000977	0.939
	$\mathcal{Q}_{2,1}$	99.9	5024.9	1.00458	0.000377	1
		2.496	11.595	0.024927	0.0000978	0.942
	$\mathcal{Q}_{2,2}$	100.7	5027.3	1.00478	0.000413	1.023
		2.629	12.090	0.025569	0.0001028	1.033

Table 4.12(b). Simulations from $N(5, 2^2)$ and $N(3, 100^2)$ under $A = 1$, $m = 10$, and $\lambda = 2$ as needed with 10 runs implementing $\mathcal{Q}_{2,0}$-$\mathcal{Q}_{2,5}$ from (4.4.13)

$\mathbf{n}^* = (n_1^*, n_2^*)$ c	$\mathcal{Q}$	$\overline{n}_1$ $s(\overline{n}_1)$	$\overline{n}_2$ $s(\overline{n}_2)$	$\widehat{\kappa}$ $s(\widehat{\kappa})$	$\widehat{\omega}$ $s(\widehat{\omega})$	γ $\widehat{\omega}/c$
$(100, 5000)$	$\mathcal{Q}_{2,3}$	100.9	5028.0	1.00474	0.000500	1.142
0.0004		2.861	14.628	0.024943	0.0001199	1.249
	$\mathcal{Q}_{2,4}$	100.7	5028.0	1.00474	0.000465	1.082
		2.629	14.628	0.025580	0.0001084	1.164
	$\mathcal{Q}_{2,5}$	100.9	5027.3	1.00479	0.000447	1.082
		2.861	12.090	0.024932	0.0001150	1.118
$(300, 15000)$	$\mathcal{Q}_{2,0}$	306.5	15012.7	1.00118	0.000045	1
0.0000444		3.481	29.231	0.011695	0.0000118	1.002
	$\mathcal{Q}_{2,1}$	306.8	15012.8	1.00118	0.000047	1
		3.732	29.151	0.012578	0.0000123	1.062
	$\mathcal{Q}_{2,2}$	307.6	15018.2	1.00114	0.000048	1.023
		3.685	29.038	0.013078	0.0000124	1.090
	$\mathcal{Q}_{2,3}$	308.2	15025.1	1.00113	0.000051	1.142
		3.530	30.350	0.012630	0.0000131	1.152
	$\mathcal{Q}_{2,4}$	307.6	15025.1	1.00112	0.000051	1.082
		3.685	30.350	0.013075	0.0000132	1.155
	$\mathcal{Q}_{2,5}$	308.2	15018.2	1.00114	0.000048	1.082
		3.530	29.038	0.012633	0.0000124	1.087

Table 4.13(a). Simulations from $N(5, 2^2)$ and $N(3, 100^2)$ under $A = 1$, $m = 10$,and $\lambda = 2$ as needed with 10 runs implementing (4.4.13): terminal sample sizes

(n_1^*, n_2^*) c	$\mathcal{Q}$	(n_{1t}, n_{2t})
$(50, 2500)$	$\mathcal{Q}_{2,0}$	$(45, 2551), (53, 2507), (53, 2517), (46, 2508), (50, 2536)$
0.0016		$(55, 2496), (49, 2543), (56, 2482), (47, 2442), (44, 2501)$
	$\mathcal{Q}_{2,1}$	$(45, 2551), (53, 2508), (54, 2517), (46, 2508), (50, 2536)$
		$(55, 2497), (49, 2545), (55, 2480), (47, 2442), (44, 2502)$
	$\mathcal{Q}_{2,2}$	$(45, 2558), (53, 2499), (54, 2517), (45, 2514), (50, 2534)$
		$(55, 2503), (49, 2543), (56, 2484), (47, 2436), (44, 2500)$
	$\mathcal{Q}_{2,3}$	$(44, 2569), (48, 2491), (56, 2517), (45, 2518), (51, 2535)$
		$(55, 2531), (46, 2533), (58, 2503), (46, 2420), (44, 2491)$
	$\mathcal{Q}_{2,4}$	$(45, 2569), (53, 2491), (54, 2517), (45, 2518), (50, 2535)$
		$(55, 2531), (49, 2533), (56, 2503), (47, 2420), (44, 2491)$
	$\mathcal{Q}_{2,5}$	$(44, 2558), (48, 2499), (56, 2517), (45, 2514), (51, 2534)$
		$(55, 2503), (46, 2543), (58, 2484), (46, 2436), (44, 2500)$
$(100, 5000)$	$\mathcal{Q}_{2,0}$	$(98, 5058), (100, 5026), (104, 5032), (103, 5024), (91, 4944)$
0.0004		$(105, 5030), (87, 5033), (92, 5085), (107, 5006), (112, 5010)$
	$\mathcal{Q}_{2,1}$	$(98, 5058), (100, 5026), (104, 5032), (103, 5025), (91, 4943)$
		$(105, 5030), (87, 5033), (92, 5085), (107, 5007), (112, 5010)$
	$\mathcal{Q}_{2,2}$	$(97, 5058), (100, 5032), (105, 5037), (105, 5029), (90, 4934)$
		$(105, 5019), (88, 5039), (95, 5083), (107, 5020), (115, 5022)$

Table 4.13(b). Simulations from $N(5, 2^2)$ and $N(3, 100^2)$ under $A = 1$, $m = 10$,and $\lambda = 2$ as needed with 10 runs implementing (4.4.13): terminal sample sizes

(n_1^*, n_2^*) c	$\mathcal{Q}$	(n_{1t}, n_{2t})
$(100, 5000)$	$\mathcal{Q}_{2,3}$	$(92, 5059)$, $(100, 5036)$, $(104, 5063)$, $(105, 5037)$, $(90, 4911)$
0.0004		$(104, 5002)$, $(94, 5038)$, $(92, 5074)$, $(110, 5043)$, $(118, 5017)$
	$\mathcal{Q}_{2,4}$	$(97, 5059)$, $(100, 5036)$, $(105, 5063)$, $(105, 5037)$, $(90, 4911)$
		$(105, 5002)$, $(88, 5038)$, $(95, 5074)$, $(107, 5043)$, $(115, 5017)$
	$\mathcal{Q}_{2,5}$	$(92, 5058)$, $(100, 5032)$, $(104, 5037)$, $(105, 5029)$, $(90, 4934)$
		$(104, 5019)$, $(94, 5039)$, $(92, 5083)$, $(110, 5020)$, $(118, 5022)$
$(300, 15000)$	$\mathcal{Q}_{2,0}$	$(301, 15114)$, $(299, 14996)$, $(288, 15126)$, $(325, 15058)$, $(306, 15043)$
0.0000444		$(312, 14842)$, $(299, 15059)$, $(314, 14886)$, $(302, 15040)$, $(319, 14963)$
	$\mathcal{Q}_{2,1}$	$(301, 15115)$, $(299, 14999)$, $(288, 15124)$, $(325, 15057)$, $(306, 15044)$
		$(312, 14842)$, $(300, 15059)$, $(319, 14886)$, $(298, 15039)$, $(320, 14963)$
	$\mathcal{Q}_{2,2}$	$(301, 15125)$, $(301, 14982)$, $(291, 15141)$, $(326, 15065)$, $(306, 15049)$
		$(312, 14871)$, $(302, 15060)$, $(321, 14888)$, $(296, 15045)$, $(320, 14956)$
	$\mathcal{Q}_{2,3}$	$(297, 15155)$, $(299, 14952)$, $(296, 15150)$, $(324, 15092)$, $(309, 15049)$
		$(311, 14910)$, $(304, 15034)$, $(323, 14875)$, $(298, 15062)$, $(321, 14972)$
	$\mathcal{Q}_{2,4}$	$(301, 15155)$, $(301, 14952)$, $(291, 15150)$, $(326, 15092)$, $(306, 15049)$
		$(312, 14910)$, $(302, 15034)$, $(321, 14875)$, $(296, 15062)$, $(320, 14972)$
	$\mathcal{Q}_{2,5}$	$(297, 15125)$, $(299, 14982)$, $(296, 15141)$, $(324, 15065)$, $(309, 15049)$
		$(311, 14871)$, $(304, 15060)$, $(323, 14888)$, $(298, 15045)$, $(321, 14956)$

We had looked at side-by-side boxplots prepared from 10 observed terminal sample size pairs under the configurations (i) $N(5, 2^2)$ and $N(3, 8^2)$ from Table 11 as well as (ii) $N(5, 2^2)$ and $N(3, 100^2)$ from Table 13. From the boxplots, we felt that for smaller n^* values, the methodologies $\mathcal{Q}_{2,2}$-$\mathcal{Q}_{2,5}$ led to larger final sample sizes on an average with regard to their means and medians. But, for larger n^* values, while they still led to larger means, the medians were nearly the same as (or even smaller than) those associated with the methodologies $\mathcal{Q}_{2,0}$ and $\mathcal{Q}_{2,1}$. Overall, $\mathcal{Q}_{2,3}$, the one based on MAD, seemed to have the largest spread amongst these methodologies.

As reflected in the side-by-side boxplots, the methodology $\mathcal{Q}_{2,4}$ ($\mathcal{Q}_{2,5}$) performed comparably with $\mathcal{Q}_{2,3}$ ($\mathcal{Q}_{2,2}$), especially when the two populations were $N(5, 2^2)$ and $N(3, 100^2)$, that is, when there was a huge difference between σ_1 and σ_2. This was perhaps due to the fact that the observations sampled from $N(3, 100^2)$ contributed to the combined final sample size in a big way.

In this latter case, if we focus on the last column of Table 4.12, we may note that under the configuration $\mathbf{n}^* = (50, 2500)$, the $\widehat{\omega}/c$ values were obviously smaller than the theoretical approximations given by γ associated with $\mathcal{Q}_{2,0}$ and $\mathcal{Q}_{2,1}$. When $\mathbf{n}^*$ increased to $(100, 5000)$ or $(300, 15000)$, the $\widehat{\omega}/c$ values began settling down more around γ.

Our sentiment is that when there is a huge difference between σ_1 and σ_2, a considerably larger combined sample size may be needed to ensure practical validity of the second-order regret property under $\mathcal{Q}_{2,0}$ or $\mathcal{Q}_{2,1}$. The MAD-based methodology $\mathcal{Q}_{2,3}$ appears more robust in the sense that the second-order term appears closer to the theoretical values. Under that backdrop, $\mathcal{Q}_{2,3}$ might be suggested for practical purposes.

4.4.4. Real Data Illustrations

We go back to the earlier data source we had incorporated in Section 4.3.6. Suppose that we want to estimate the change of the prevailing wage of "Sr. programmer analyst" between the years 2011 and 2012. That is, we now focus on datasets D_1 and D_2 highlighted in

(4.3.26) There were 232 and 371 observations in those groups respectively. The F-test on the variances came up with a p-value 0.003976. Thus, it seemed reasonable for us to postulate that these two datasets had different variances.

Next, in order to find the MRPE of the difference of means, $\delta \equiv \mu_1 - \mu_2$, we initially sampled $m = 20$ observations randomly from each group and used them as our pilot data. Then, sequentially we sampled one observation (without replacement) at-a-time from D_1 and checked with the stopping rule defined in (4.4.9). The terminal sample size n_1 along with $\overline{X}_{1n_1}$, the terminal sample mean of D_1, were observed. Simultaneously, we sequentially sampled one observation (without replacement) at-a-time from D_2 and checked with the stopping rule defined in (4.4.9).

Table 4.14. MRPE of $\delta_{1,2}(= \mu_1 - \mu_2)$ when $\sigma_1 \neq \sigma_2$

	$(n_1^*, n_2^*) = (30, 30)$		$(n_1^*, n_2^*) = (100, 30)$		$(n_1^*, n_2^*) = (30, 150)$		$(n_1^*, n_2^*) = (150, 150)$	
$\mathcal{Q}$	(n_1, n_2)	$T_{\mathbf{n}}$	(n_1, n_2)	$T_{\mathbf{n}}$	(n_1, n_2)	$T_{\mathbf{n}}$	(n_1, n_2)	$T_{\mathbf{n}}$
$\mathcal{Q}_{2,0}$	$(30, 20)$	-0.0191	$(100, 33)$	-0.0109	$(30, 143)$	-0.0174	$(145, 135)$	-0.0054
$\mathcal{Q}_{2,1}$	$(31, 30)$	0.0049	$(97, 31)$	-0.0006	$(31, 147)$	0.0016	$(152, 160)$	-0.0069
$\mathcal{Q}_{2,2}$	$(36, 31)$	-0.0010	$(107, 28)$	0.0203	$(36, 151)$	-0.0111	$(159, 141)$	0.0123
$\mathcal{Q}_{2,3}$	$(30, 34)$	0.0261	$(106, 34)$	0.0184	$(30, 147)$	0.0217	$(154, 145)$	-0.0053
$\mathcal{Q}_{2,4}$	$(30, 27)$	-0.0164	$(98, 31)$	-0.0183	$(30, 147)$	-0.0187	$(148, 134)$	-0.0082
$\mathcal{Q}_{2,5}$	$(29, 30)$	-0.0006	$(86, 34)$	-0.0039	$(29, 147)$	0.0028	$(145, 159)$	-0.0102

Simple random sampling without replacement may cause very weak dependence within the sequence of gathered observations, but any impact was visibly negligible. The terminal sample size n_2 along with $\overline{X}_{2n_2}$, the terminal sample mean of D_2, were observed.Finally, the MRPE of $\delta_{1,2}$ was $T_{\mathbf{n}} \equiv \overline{X}_{1n_1} - \overline{X}_{2n_2}$.

Table 4.14 shows that the terminal sample size pairs (n_1, n_2) were close to the pre-assigned optimal fixed sample size pairs (n_1^*, n_2^*). From the full dataset, we found $\delta_{1,2} = -0.00265$. Compared to this value, the MRPE $T_{\mathbf{n}}$ for $\delta_{1,2}$ tended to fluctuate around the true difference. But, then, we should reiterate that each $T_{\mathbf{n}}$ value reported in Table 4.14 was obtained from a single run of (4.4.9). That said, the observed terminal $T_{\mathbf{n}}$ values do not look like scattered all over the space.

4.5. Brief Conclusion

We began with two *independent* populations, $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$, with all four parameters assumed unknown, $(\mu_1, \mu_2) \in R^2$ and $(\sigma_1, \sigma_2) \in R^{+2}$ and revisited the MRPE problems for $\mu_1 - \mu_2 (\equiv \delta$, say) in our quest for comparing two treatments. Under two distinct scenarios of (i) equal but unknown variance and (ii) unequal and unknown variances, we proceeded to replace the multiples of sample standard deviations used in defining requisite boundary crossing conditions with the GMD's, the MAD's, along with a number of combinations of the sample standard deviations, the GMD's, and the MAD's. We have proved that the ensuing body of such newly developed purely sequential MRPE strategies have associated second-order regret expansions. Our analyses suggest that the GMD-based as well as the MAD-based strategies are better equipped to withstand occurrences of possible outlying observations. Our proposed theory and methodologies are amply supported by large-scale simulations and illustrations with real data.

Chapter 5

Future Work

5.1. Two Sample Comparison Under Sampling in Groups

In Chapter 4, we constructed the sequential methodologies of comparing two treatment means based on GMD and MAD. Sampling one (or one pair of) observation(s) is not essential in these methods. These methods can be reasonably expanded into the group sampling framework. The estimators defined in Chapter 2 and Chapter 3 provide the basic idea in defining the new class of estimators. We will provide the general procedures of conducting such two-sample comparisons and investigate the first-order and second-order properties.

5.2. MRPE For A Function Of Mean In Normal Population Under Sampling in Groups

In many cases, instead of estimating the population mean μ, people focus on the estimate of a parametric function $g(\mu)$. Mukhopadhyay and Wang (2019) developed a general sequential methodologies to get the minimum risk point estimation (MRPE) for a function of mean in a normal distribution. Under an appropriately formulated weighted squared error loss due to estimation of $g(\mu)$ plus linear cost of sampling, the purely sequential procedure has first-order and second-order asymptotic efficiency and risk efficiency property. This purely sequential sampling strategy can be easily extended to the sampling-in-group methodology. We will examine the asymptotic first-order and second order properties and perform large scale simulation as well as real world applications.

Chapter 6

Summary

Sequential analysis has been widely used in clinical trials, quality and reliability control, and environmental sampling. It's the concept of decision-making in real-time as data is collected, as opposed to retrospectively on a pre-fixed sample size, as is typically done. Under sequential analysis, an experimenter gathers information regarding an unknown parameter by observing random samples in successive steps. A common feature among such sampling designs is that the total number of observations collected at termination is a positive integer-valued random variable. Inference problems arose from many areas of interest including tests of hypothesis, point or interval estimation, regression analysis, multivariate analysis, selection and ranking, and multiple comparisons. When defining the stopping rule in sequential methodology, it's not necessary to limit ourselves in using certain estimators, or sampling a certain number of observations at-a-time. Results in this thesis show that the stopping rule can be quite flexible. Traditional purely sequential sampling, that is, sample one observation at-a-time, can be easily extended to a group sampling strategy with all good asymptotic properties maintained. This result can be directly applied when observations come in batches, or when sampling in groups is more efficient than sampling in individuals. With sequential design, it is possible to have optimization with high reliability and control the error (or even measure the error) at the same time.

List of Tables

Bibliography

Abramowitz, M. and Stegun, I. A. (1972). *Handbook of Mathematical Functions*, tenth printing, New York: Dover.

Albert, A. (1966). Fixed Size Confidence Ellipsoids for Linear Regression Parameters, The Annals of Mathematical Statistics, 37(6), pp.1602-1630.

Anscombe, F. J. (1952). Large-Sample Theory of Sequential Estimation, *Proceedings of Cambridge Philosophical Society* 48: 600-607.

Anscombe, F. J. (1953). Sequential Estimation, *Journal of Royal Statistical Society, Series B* 15: 1-29.

Aoshima, M., Mukhopadhyay, N., and Kobayashi, Y. (2011). Two-Stage Procedures for Estimating the Difference of Means When the Sampling Cost Is Different, *Sequential Analysis* 30: 160-171.

Barnard, G. A. (1946). Sequential Tests in Industrial Statistics, *Supplement to the Journal of the Royal Statistical Society,* 8(1), pp.1-21.

Basu, D. (1955). On Statistics Independent of a Complete Sufficient Statistic, *Sankhyā* 15: 377-380.

Bose, A. and Mukhopadhyay, N. (1995). A Note on Accelerated Sequential Estimation of the Mean of NEF-PVF Distributions, *Annals of the Institute of Statistical Mathematics*, 47(1), pp.99-104.

Chow, Y. S. and Robbins, H. (1965). On the Asymptotic Theory of Fixed-Width Sequential Confidence Intervals for the Mean, *Annals of Mathematical Statistics* 36: 457-462.

Costanza, M. C., Hamdy, H. I. and Son, M. S. (1986). Two-Stage Fixed Width Confidence Intervals for the Common Location Parameter of Several Exponential Distributions, *Communications in Statistics-Theory and Methods*, 15(8), pp.2305-2322.

Dantzig, G. B. (1940). On the Non-Existence of Tests of "Student's" Hypothesis Having Power Functions Independent of σ, *Annals of Mathematical Statistics* 11: 186-192.

Dodge, H. F. and Romig, H. G. (1929). A Method of Sampling Inspection, *The Bell System Technical Journal,* 8(4), pp.613-631.

Draxl, C., Hodge, B. M., Clifton, A. and McCaa, J. (2015a). Overview and Meteorological Validation of the Wind Integration National Dataset Toolkit (Technical Report, NREL/TP-5000-61740), Golden, Colorado: National Renewable Energy Laboratory.

Draxl, C., Hodge, B. M., Clifton, A. and McCaa, J. (2015b). The Wind Integration National Dataset (WIND) Toolkit, *Applied Energy* 151: 355366.

Dua, D. and Graff, C. (2019). UCI Machine Learning Repository (http://archive.ics.uci.edu/ml), University of California-Irvine, School of Information and Computer Science.

Ghosh, B. K. (1970). *Sequential Tests of Statistical Hypotheses* (No. 04; QA277, G4.), Reading: Addison-Wesley.,

Ghosh, B. K. and Sen, P. K. (1991). *Handbook of Sequential Analysis*, edited volume, New York: Dekker.

Ghosh, M. and Mukhopadhyay, N. (1975). Asymptotic Normality of Stopping Times in Sequential Analysis, unpublished report, Research and Training School, Indian Statistical Institute, Calcutta.

Ghosh, M. and Mukhopadhyay, N. (1976). On Two Fundamental Problems in Sequential Estimation, *Sankhyā, Series B* 38: 203-218.

Ghosh, M. and Mukhopadhyay, N. (1980). Sequential Point Estimation of the Difference of Two Normal Means, *Annals of Statistics* 8: 221-225.

Ghosh, M. and Mukhopadhyay, N. (1981). Consistency and Asymptotic Efficiency of Two-Stage and Sequential Procedures, *Sankhyā: The Indian Journal of Statistics, Series A,* 43: 220–227.

Ghosh, B. K. and Sen, P. K. (1991). *Handbook of Sequential Analysis*, edited volume, New York: Dekker.

Ghosh, M., Mukhopadhyay, N., and Sen, P. K. (1997). *Sequential Estimation*, New York: Wiley.

Ghurye, S. G. (1958). Note on Sufficient Statistics and Two-Stage Procedures, *Annals of Mathematical Statistics*, 29(1), pp.155-166.

Haldane, J. B. S. (1945). On A Method of Estimating Frequencies, *Biometrika*, 33(3), pp.222-225.

Herrey, E. M. (1965). Confidence Intervals Based on the Mean Absolute Deviation of a Normal Sample, *Journal of American Statistical Association* 60: 257-269.

https://www.nrel.gov/wind/data-tools.html.

https://www.foreignlaborcert.doleta.gov/performancedata.cfm: The Office of Foreign Labor Certification's iCErt Visa Portal System, 2020.

Hu, J. and Mukhopadhyay, N. (2019). Second-Order Asymptotics in a Class of Purely Sequential Minimum Risk Point Estimation (MRPE) Methodologies, *Japanese Journal of Statistics and Data Science* 2: 81-104.

King, J., Clifton, A., and Hodge, B. M. (2014). Validation of Power Output for the WIND Toolkit (Technical Report, NREL/TP-5D00-61714), Golden, Colorado: National Renewable Energy Laboratory.

Lai, T. L. (2001). Sequential Analysis: Some Classical Problems and New Challenges, *Statistica Sinica*, pp.303-351.

Lai, T. L. and Siegmund, D. (1977). A Nonlinear Renewal Theory with Applications to Sequential Analysis I, *Annals of Statistics* 5: 946-954.

Lai, T. L. and Siegmund, D. (1979). A Nonlinear Renewal Theory with Applications to Sequential Analysis II, *Annals of Statistics* 7: 60-76.

Lieberman-Cribbin, W., Draxl, C., and Clifton, A. (2014). Guide to Using the WIND Toolkit Validation Code (Technical Report, NREL/TP-5000-62595), Golden, Colorado: National Renewable Energy Laboratory.

Mahalanobis, P. C. (1940). A Sample Survey of the Acreage Under Jute in Bengal, *Sankhyā: The Indian Journal of Statistics*: pp.511-530.

Malinovsky, Y. and Zacks, S. (2018). Proportional Closeness Estimation of Probability of Contamination Under Group Testing, *Sequential Analysis* 37: 145-157.

Mukhopadhyay, N. (1974). Sequential estimation of location parameter in exponential distributions, *Calcutta Statistical Association Bulletin*, 23(1-4), pp.85-96.

Mukhopadhyay, N. (1976). Sequential Estimation of a Linear Function of Means of Three Normal Populations, *Journal of American Statistical Association* 71: 149-153.

Mukhopadhyay, N. (1977). Remarks on Sequential Estimation of a Linear Function of Two Means: The Normal Case, *Metrika* 24: 197-201.

Mukhopadhyay, N. (1988). Sequential Estimation Problems for Negative Exponential Populations, *Communications in Statistics-Theory & Methods, Series A* 17: 2471-2506.

Mukhopadhyay, N. (1990). Some Properties of A Three-Stage Procedure with Applications in Sequential Analysis, *Sankhyā: The Indian Journal of Statistics, Series A*, pp.218-231.

Mukhopadhyay, N. (1991). Parametric Sequential Point Estimation, in *Handbook of Sequential Analysis*, B. K. Ghosh and P. K. Sen, eds., pp. 245-267, New York: Dekker.

Mukhopadhyay, N. (2000). *Probability and Statistical Inference*, New York: Dekker.

Mukhopadhyay, N. and Al-Mousawi, J. S. (1986). Fixed-Size Confidence Regions for the Mean Vector of a Multinormal Distribution, *Sequential Analysis* 5: 139-168.

Mukhopadhyay, N. and Chattopadhyay, B. (2013). Two-Stage Fixed-Width Confidence Interval Procedures for the Mean of a Normal Distribution, *Sequential Analysis* 32: 134-157.

Mukhopadhyay, N. and Chattopadhyay, S. (1991). Sequential Methodologies for Comparing Exponential Mean Survival Times, *Sequential Analysis* 10: 139-148.

Mukhopadhyay, N. and Datta, S. (1996). On Sequential Fixed-Width Confidence Intervals for the Mean and Second-Order Expansions of the Associated Coverage Probabilities, *Annals of the Institute of Statistical Mathematics*, 48(3), pp.497-507.

Mukhopadhyay, N. and de Silva, B. M. (2009). *Sequential Methods and Their Applications*, Boca Raton: CRC.

Mukhopadhyay, N. and Hu, J. (2017). Confidence Intervals and Point Estimators for a Normal Mean Under Purely Sequential Strategies Involving Gini's Mean Difference and Mean Absolute Deviation, *Sequential Analysis* 36: 210-239.

Mukhopadhyay, N. and Moreno, M. (1991). Multistage and Sequential Minimum Risk Point Estimation Procedures for the Means of Generalized U-Statistics, *Sankhyā: The Indian Journal of Statistics, Series A* 53: 220-254.

Mukhopadhyay, N. and Purkayastha, S. (1994). On Sequential Estimation of the Difference of Means, *Statistics & Decisions* 12: 41-52.

Mukhopadhyay, N. and Solanky, T. K. S. (1994). *Multistage Selection and Ranking Procedures*, New York: Dekker.

Mukhopadhyay, N. and Wang, Z. (2020). Purely Sequential FWCI and MRPE Problems for the Mean of a Normal Population by Sampling in Groups with Illustrations Using Breast Cancer Data, *Sequential Analysis* 39: https://doi.org/10.1080/07474946.2020.1766893.

Mukhopadhyay, N., and Zhang, C. (2018). EDA on the Asymptotic Normality of the Standardized Sequential Stopping Times, Part-I: Parametric Models, *Sequential Analysis* 37: 342-374.

Mukhopadhyay, N., Datta, S., and Chattopadhyay, S. (2004). *Applied Sequential Methodologies*, edited volume, New York: Dekker.

Mukhopadhyay, N., de Silva, B. M., and Waikar, V. B. (2010). On Two-Stage Confidence Interval Procedures and Their Comparisons for Estimating the Difference of Normal Means, *Sequential Analysis* 29: 44-68.

Ray, W. D. (1957). Sequential Confidence Intervals for the Mean of a Normal Population with Unknown Variance, *Journal of Royal Statistical Society, Series B* 10: 133-143.

Robbins, H. (1959). Sequential Estimation of the Mean of a Normal Population, in *Probability and Statistics*, H. Cramér volume, U. Grenander, ed., pp. 235-245, Uppsala: Almquist and Wiksell.

Robbins, H., Simons, G., and Starr, N. (1967). A Sequential Analogue of the Behrens-Fisher Problem, *Annals of Mathematical Statistics* 38: 1384-1391.

Sen, P. K. (1981). *Sequential Nonparametrics*, New York: Wiley.

Shewhart, W. A. (1931). *Economic Control of Quality of Manufactured Product*, London: Macmillan And Co Ltd.

Siegmund, D. (1985). *Sequential Analysis: Tests and Confidence Intervals*, New York: Springer.

Srivastava, M. S. (1970). On a Sequential Analogue of the Behrens-Fisher Problem, *Journal of Royal Statistical Society*, *Series B* 32: 144-148.

Starr, N. (1966a). The Performance of a Sequential Procedure for the Fixed-Width Interval Estimation of the Mean, *Annals of Mathematical Statistics* 37: 36-50.

Starr, N. (1966b). On the Asymptotic Efficiency of a Sequential Procedure for Estimating the Mean, *Annals of Mathematical Statistics* 37: 1173-1185.

Starr, N. and Woodroofe, M. B. (1969). Remarks on Sequential Point Estimation, *Proceedings of the National Academy of Sciences, USA* 63: 285-288.

Starr, N. and Woodroofe, M. (1972). Further Remarks on Sequential Estimation: The Exponential Case, *Annals of Mathematical Statistics* 43: 1147-1154.

Stein, C. (1945). A Two Sample Test for a Linear Hypothesis Whose Power Is Independent of the Variance, *Annals of Mathematical Statistics* 16: 243-258.

Stein, C. (1949). Some Problems in Sequential Estimation (abstract), *Econometrica* 17: 77-78.

Thompson, W. R. (1933). On the Likelihood That One Unknown Probability Exceeds Another in View of the Evidence of Two Samples, *Biometrika*, pp.285-294.

Wald, A. (1948). Asymptotic Properties of the Maximum Likelihood Estimate of an Unknown Parameter of a Discrete Stochastic Process, *Annals of Mathematical Statistics*, pp.40-46.

Wiener, N. (1939). The Ergodic Theorem, *Duke Mathematical Journal* 5: 1-18.

Woodroofe, M. (1977). Second Order Approximations for Sequential Point and Interval Estimation, *Annals of Statistics* 5: 984-995.

Woodroofe, M. (1982). *Nonlinear Renewal Theory in Sequential Analysis*, Philadelphia: SIAM.

Woodroofe, M. (1987). Confidence Intervals with Fixed Proportional Accuracy, *Journal of Statistical Planning and Inference*, 15, pp.131-146.

Zacks, S. (2009). *Stage-Wise Adaptive Designs*, New York: Wiley.

Zacks, S. (2017). *Sample Path Analysis and Distributions of Boundary Crossing Times*, Lecture Notes in Mathematics, New York: Springer.