5-20-2005

Construction of Some Unbalanced Designs for the Partition Problem

Yuefeng Wu
University of New Orleans

CONSTRUCTION OF SOME UNBALANCED DESIGNS FOR THE PARTITION PROBLEM

A Thesis

Submitted to the Graduate Faculty of the
University of New Orleans
in partial fulfillment of the
requirements for the Degree of

Master of Science
in
Mathematics

by

Yuefeng Wu

B.S. Nanjing University, 2000
M.S. Florida State University, 2003

May 2005

To My Wife

Acknowledgments

I would like to thank Professor Solanky for his continual guidance during the course of my study program at the University of New Orleans.

Yuefeng Wu

University of New Orleans
May 2005

Contents

List of Tables

Abstract

In a pioneering work, Bechhofer (1954) introduced the concept of indifference-zone formulation and formulated some methodologies in the case of the problem of selecting the best normal population. In statistical literature, numerous $vector - at - a\ time$ and unbalanced methodologies are available for the selecting the best normal population. However, the literature is not that rich for the partition problem. In this thesis, an unbalanced methodology of sampling along the lines of Mukhopadhyay and Solanky (2002) is introduced for the partition problem. A two-stage and a purely sequential procedure are introduced which takes $c\,(\geq 1)$ observations from the control population for each observation from the non-control populations. The theoretical properties of the two introduced procedures are derived. Also the two proposed procedures are simulated via Monte Carlo simulations and then small to moderate sample size performances have been studied. The robustness of various already known procedures in the statistical literature and the ones proposed in this thesis are studied. An attempt has also been made to determine the optimal choice of the value of c.

Introduction

0.1 A Brief History

Since the first appearance in the early 1950's of the ranking and selection formulation of statistical inference problem, the literature in the area has grown enormously in all ramifications. There are numerous procedures along the lines of Bechhofer's (1954) indifference-zone formulation, and also along the lines of Gupta's subset selection, to carry out multiple comparisons.

The idea of sampling in two stages was first considered by Mahalanobis (1940), and later by Stein (1945, 1949) to construct a fixed-width confidence intervals for a normal mean problem. The purely sequential procedures have been considered by Chow and Robbins (1965) and Srivastava (1966) for some ranking problems. Tong (1969) formulated the partition problem using Bechhofer's (1954) indifference zone formulation and constructed two-stage and purely sequential procedures. Starr (1966) and Woodroofe (1977) have also done some ground breaking work to further the theory behind sequential and other multistage procedures. A brief history of these procedures is available in Mukhopadhyay and Solanky (1994), Ghosh, Mukhopadhyay and Sen (1997), and Ghosh and Sen (1991).

Finally, in Solanky and Wu (2004), the unbalanced two-stage procedure and the unbalanced purely sequential procedures are proposed. The details and properties of these two procedures will be discussed in the next Chapter.

0.2 Progress Made in This Thesis

In Chapter 1, the author proposed the unbalanced two-stage procedure and the unbalanced purely sequential procedure. These procedures take $c(>1)$ observations from the control population while taking 1 observation from each of the non-control populations. These procedures will reduce the average sample sizes from the non-control populations. When the price of sampling is under consideration, especially the case when the price for sampling from non-control populations is higher, using these two procedures will have tremendous advantage on the total cost of sampling. The theoretical first-order and second-order asymptotics of the purely sequential procedure are obtained. The performance of the two proposed procedures is studied via Monte Carlo simulations for small and moderately large sample sizes.

In Chapter 2, the robustness of the procedures against deviations from model assumptions is assessed. The description of the procedures under the investigation and the distributions used for these simulations are specified in Chapter 2. Briefly speaking, the sequential procedure with elimination is the best choice, with respect to robustness and the total sample size. If sequential sampling is not convenient, then the fine tuned three-stage procedure is the next best choice. The two-stage procedure with elimination tends to over sample under the LFC. Hence this procedure is not preferred, unless we have some prior knowledge about the location parameters of the treatment populations. Finally, we suggested that the direction for future research is to propose a unbalanced sequential procedure with elimination along the lines of Solanky (2001).

In Chapter 3, a rule for choosing the optimal value of c is given, where c is the number of observations we take from the control population while taking one from each of the treatment populations.

Chapter 1

Unbalanced Procedures

1.1 Introduction

Suppose that we have $\pi_0, \pi_1, \ldots, \pi_k$, independent and normally distributed populations, with unknown means μ_i, and, unknown but common variance σ^2, $i = 0, 1, \cdots, k$. We consider π_0 to be the control population. The goal is to partition the set of treatments $\Omega = (\pi_i : i = 1, 2, \cdots, k)$, into two disjoint and exhaustive subsets, corresponding to "Good" and "Bad" populations compared to the control population, as defined later, and also, with a pre specified probability of correct partition.

Given arbitrary but fixed constants δ_1 and δ_2, $\delta_1 < \delta_2$, we define three subsets of Ω along the lines of Bechhofer's (1954) indifference-zone formulation, as:

$$\begin{aligned} \Omega_L &= \{\pi_i : \mu_i \leq \mu_0 + \delta_1,\ i = 1, \cdots, k\}, \\ \Omega_M &= \{\pi_i : \mu_0 + \delta_1 < \mu_i < \mu_0 + \delta_2,\ i = 1, \cdots, k\}, \\ \Omega_R &= \{\pi_i : \mu_i \geq \mu_0 + \delta_2,\ i = 1, \cdots, k\}. \end{aligned} \tag{1.1.1}$$

We refer to Ω_R as the set of "good populations" and Ω_L as the set of "bad populations". The set Ω_M would be referred to as the set of "mediocre populations". Adopting the Bechhofer's indifference zone approach, we are interested in the correct population of the populations in Ω_R and Ω_L. And, we will be indifferent to correct partition of populations in Ω_M. That is, with high accuracy we want to partition the set Ω into two disjoint subsets P_L and P_R, such that, $\Omega_L \subseteq P_L$ and $\Omega_R \subseteq P_R$. Such a partition is known in the literature as a *correct decision* (CD). In other words, given a pre assigned number P^*, $2^{-k} < P^* < 1$, we seek statistical methodologies $\wp$ to determine P_L and P_R, such that

$$P\{CD|\boldsymbol{\mu}, \sigma^2, \wp\} \geq P^* \qquad \forall\ \boldsymbol{\mu} \in \boldsymbol{R}^{k+1},\ \sigma \in \boldsymbol{R}^{+}. \tag{1.1.2}$$

Also, we will use the following notation in the rest of this thesis for convenience:

$$\begin{aligned} &d = (\delta_1 + \delta_2)/2, \quad a = (-\delta_1 + \delta_2)/2, \quad \lambda = \sigma/a, \text{ and,} \\ &r = \begin{cases} k/2 & \text{if } k \text{ is even;} \\ (k+1)/2 & \text{if } k \text{ is odd.} \end{cases} \end{aligned} \tag{1.1.3}$$

Customarily, in many situations it is possible to collect a larger sample from the control population. We assume, in general, that we observe random variables $\boldsymbol{X}_{0i}, X_{1i}, \cdots, X_{ki}$ from $\pi_0, \pi_1, \cdots, \pi_k$, respectively, where $\boldsymbol{X}'_{0i} = (X_{0\,(i-1)c+1},\ X_{0\,(i-1)c+2}, \cdots, X_{0\,ic})$, in a sequential framework, $i = 1, 2, \cdots$, and, $c(\geq 1)$ being an integer. In other words, as needed, we take c observations form π_0 and one observation from $\pi_1, \cdots, \pi_k$.

Assuming that σ^2 is known, we observe the sequence $\boldsymbol{X}_{0i}, X_{1i}, \cdots, X_{ki}$ for $i = 1, 2, \cdots, n$, where n is to be determined below. We denote

$$\begin{aligned} &\bar{X}_{0cn} = (c\,n)^{-1} \textstyle\sum_{p=1}^{n} \sum_{q=1}^{c} X_{0\,(p-1)c+q}\,, \\ &\bar{X}_{jn} = n^{-1} \textstyle\sum_{p=1}^{n} X_{jp}\,, \quad j = 1, \cdots, k\,. \end{aligned} \tag{1.1.4}$$

Consider the decision rule $\wp$ defined as:

$$\begin{aligned} P_L &= \{\pi_i :\ \bar{X}_{in} - \bar{X}_{0cn} < d,\ i = 1, \cdots, k\}, \\ P_R &= \{\pi_i :\ \bar{X}_{in} - \bar{X}_{0cn} > d,\ i = 1, \cdots, k\}. \end{aligned} \tag{1.1.5}$$

Next, observe that for a mean vector $\boldsymbol{\mu}$ to be a least favorable configuration under $\wp$, the set Ω_M must be empty, and, all the populations in Ω_L and Ω_R must have common means $\mu_0 + \delta_1$ and $\mu_0 + \delta_2$, respectively. Let $\boldsymbol{\mu}^0(r')$ be the configuration such that $\mu_i = \mu_0 + \delta_2$ and $\mu_j = \mu_0 + \delta_1$, $0 < i \leq r'$, $r' < j \leq k$ for some r' such that $0 < r' \leq k$. Then, we have

$$P\Big[CD|\boldsymbol{\mu}^0(r'),\, \sigma^2,\, \wp\Big]$$

$$= P\Big[\bar{X}_{jn} - \bar{X}_{0cn} < d,\, \bar{X}_{in} - \bar{X}_{0cn} > d,\, 0 < i \le r',\, r' < j \le k|\boldsymbol{\mu}^0(r'),\, \sigma^2\Big],$$

$$= P\Big[Y_i \le a/\sqrt{\tfrac{\sigma^2}{n}(\tfrac{c+1}{c})}, \quad i = 1, \cdots, k\Big],$$

where, $Y_i = (\bar{X}_{0cn} - \bar{X}_{in} + \delta_2)/\sqrt{\frac{\sigma^2}{n}(\frac{c+1}{c})}$, for $0 < i \le r'$ and $Y_i = (\bar{X}_{in} - \bar{X}_{0cn} - \delta_1)/\sqrt{\frac{\sigma^2}{n}(\frac{c+1}{c})}$, for $r' < i \le k$. Note that under the parameter configuration $\boldsymbol{\mu}^0(r')$, Y_i has the standard normal distribution, $i = 1, \cdots, k$. Let the $(k \times k)$ covariance matrix $\Sigma_{r'} = (\sigma_{ij})$ be given by

$$\sigma_{ij} = \begin{cases} 1 & \text{for } i = j, \\ 1/(c+1) & \text{for } i \ne j, \text{ and, } 0 < i,\, j \le r' \text{ or } r' < i, j \le k, \\ -1/(c+1) & \text{for } 0 < i \le r', \text{ and, } r' < j \le k. \end{cases} \qquad (1.1.6)$$

Then, one can express

$$P\Big[CD|\boldsymbol{\mu}^0(r'), \sigma^2, \wp\Big]$$

$$= \int_{-\infty}^{a/\sqrt{\frac{\sigma^2}{n}(\frac{c+1}{c})}} \cdots \int_{-\infty}^{a/\sqrt{\frac{\sigma^2}{n}(\frac{c+1}{c})}} (2\pi)^{-\frac{k}{2}} |\Sigma_{r'}|^{-\frac{1}{2}} \exp\big(-\frac{1}{2}\boldsymbol{Y}'\Sigma_{r'}^{-1}\boldsymbol{Y}\big) \prod_{i=1}^{k} dy_i, \qquad (1.1.7)$$

where $\boldsymbol{Y}' = [Y_1, \cdots, Y_k]$. Note that (1.1.7) gives the infimum of the probability of correct decision under $\wp$ for the set of all configurations such that there are r' populations in Ω_R and $k - r'$ in Ω_L. Also, observe that (1.1.7) is similar to the equation (1.6) of Tong (1969). Next, using the theorem from the Appendix of Tong (1969), with $\rho = 1/(c+1)$ in the equation A.1 of Tong (1969), one obtains the Least Favorable Configuration (LFC) under the decision rule $\wp$ as: $\mu_1 = \cdots = \mu_r = \mu_0 + \delta_2$, and, $\mu_{r+1} = \cdots = \mu_k = \mu_0 + \delta_1$, where r is defined in (1.1.3). We will refer to the LFC as $\boldsymbol{\mu}^0$. Next, along the lines of (1.1.6) with r in place of r', we define the covariance matrix Σ as:

$$\Sigma = \begin{pmatrix} 1 & & \frac{1}{c+1} & -\frac{1}{c+1} & \cdots & -\frac{1}{c+1} \\ & \ddots & & \vdots & \ddots & \vdots \\ \frac{1}{c+1} & & 1 & -\frac{1}{c+1} & \cdots & -\frac{1}{c+1} \\ -\frac{1}{c+1} & \cdots & -\frac{1}{c+1} & 1 & & \frac{1}{c+1} \\ \vdots & \ddots & \vdots & & \ddots & \\ -\frac{1}{c+1} & \cdots & -\frac{1}{c+1} & \frac{1}{c+1} & & 1 \end{pmatrix}. \tag{1.1.8}$$

Next, as in Tong (1969), let $b = b(P^*, k, c)$ be the solution of the equation

$$P^* = \int_{-\infty}^{b} \cdots \int_{-\infty}^{b} (2\pi)^{-\frac{k}{2}} |\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\boldsymbol{Y}'\Sigma^{-1}\boldsymbol{Y}\right) \prod_{i=1}^{k} dy_i. \tag{1.1.9}$$

Then, one can immediately note that

$$P\Big[CD|\boldsymbol{\mu}, \sigma^2, \wp\Big] \geq P^*,$$

provided n satisfies

$$n \geq \frac{b^2\sigma^2}{a^2}\left(\frac{c+1}{c}\right) \quad (= n_c^*, \text{ say}). \tag{1.1.10}$$

In other words, if σ^2 is known, and one collects a sample of size n_c^* from each of $\pi_1, \ldots, \pi_k$ and a sample of size cn_c^* from π_0, and, uses the decision rule $\wp$ given by (1.1.5) to partition the k populations, then the probability requirement (1.1.2) is achieved.

For $c = 1$ case, Tong (1969) gave a single-stage procedure for the partition problem when the σ^2 known. The single-stage procedure provided above is a simple generalization of Tong's (1969) single-stage procedure. The values of the constant b, for selected values of P^*, k, and, c, have been tabulated in the Table 1.1, in section 4 of this chapter. Note that for $c = 1$, the values of b have been extensively tabulated in Tong (1969), and, as well in the chapter 10 of Gibbons, Olkin, and Sobel (1977).

For the case when σ^2 is unknown, it is known that there does not exist a single-stage procedure which can satisfy the probability requirement (1.1.2). So, for the σ^2 unknown case, Tong (1969) constructed a two-stage and a purely sequential procedure for $c = 1$. Datta and Mukhopadhyay (1998) studied this problem further for the $c = 1$ case and constructed a fine-tuned purely sequential procedure and some other multistage methodologies, emphasizing the second-order asymptotics. Solanky (2001) has constructed an elimination type procedure for the partition problem for the $c = 1$ case which takes samples of unequal sizes. The reader is also recommended to look at Aoshima and Takada (2000) and Solanky (2004), who have studied various aspects of the partition problem. Many other additional references to the partition problem are available in the articles mentioned in this paragraph.

In this chapter, we focus on the case when c can be any positive integer by constructing a two-stage and a purely sequential procedure for this problem, which are described in the sections 2 and 3 of this chapter, respectively. In section 4 of this chapter, we study the small and moderate sample size performance of these procedures via Monte Carlo Simulation studies and also provide relevant tables to facilitate practical usage of the two proposed procedures.

1.2 Two-Stage Procedure

Writing m ($\geq$2) for the starting sample size, one starts with mc observations from π_0 and m observations from each of $\pi_1, \cdots, \pi_k$, to obtain the stage I sample, of the two-stage sampling design, as $\boldsymbol{X}_{0i}, X_{1i}, \cdots, X_{ki}$, $i = 1, 2, \cdots, m$. Then, we define

$$\begin{aligned}
&\bar{X}_{0cm} = (cm)^{-1} \textstyle\sum_{p=1}^{m} \sum_{q=1}^{c} X_{0\,(p-1)c+q}\,, \\
&S_{0cm}^2 = (cm-1)^{-1} \textstyle\sum_{p=1}^{m} \sum_{q=1}^{c} (X_{0\,(p-1)c+q} - \bar{X}_{0\,m})^2\,, \\
&\bar{X}_{jm} = m^{-1} \textstyle\sum_{p=1}^{m} X_{j\,p}\,, \\
&S_{jm}^2 = (m-1)^{-1} \textstyle\sum_{p=1}^{m} (X_{j\,p} - \bar{X}_{j\,m})^2, \quad j = 1, \cdots, k\,.
\end{aligned}$$

Also, we define

$$S_\nu^2 = \{(cm-1)S_{0cm}^2 + (m-1)\sum_{j=1}^{k} S_{jm}^2\}/\{(cm-1) + k(m-1)\}, \qquad (1.2.1)$$

as the usual pooled estimator of the common unknown variance σ^2, with $\nu = (cm-1) + k(m-1)$ degree of freedom. Next, we define the two-stage procedure as:

$$N = max\{m, < \frac{h_\nu^2 S_\nu^2}{a^2}(\frac{c+1}{c}) >\}, \tag{1.2.2}$$

where, $< x >$ denotes the largest integer less than x, and $h_\nu = h_\nu(P^*, k, c)$ is a constant defined in (1.2.4).

Note that for the two-stage procedure, the sampling is carried out in two batches. We start with cm observations form π_0 and m observations from $\pi_1, \cdots, \pi_k$. Next, we determine the value of N using (1.2.2). If $N = m$, then no additional sampling is carried out. However, if $N > m$, the difference, that is, $Nc - mc$ observations form π_0, and, $N - m$ from $\pi_1, \cdots, \pi_k$, are sampled in one batch, known as the stage II of the two-stage procedure. Next, the sample mean $\bar{X}_{0cN}$ from π_0, and $\bar{X}_{iN}$ from π_i, $i = 1, \cdots, k$ are computed, as defined in (1.1.4) with N in place of n and the decision rule (1.1.5) is implemented accordingly.

Theorem 1.2.1 *If N is chosen according to (1.2.2) with $h_\nu = h_\nu(P^*, k, c)$ as defined in (1.2.4), then we have*

$$P\Big[CD|\boldsymbol{\mu}, \sigma^2, \wp\Big] \geq P^*,$$

provided the decision rule (1.1.5)is used to partition the populations based on N observations each from $\pi_1, \cdots, \pi_k$ and cN observations from π_0.

Proof: We consider without loss of generality, a parametric configuration $\boldsymbol{\mu}^0$ under the LFC given by $\mu_1 = \ldots = \mu_r = \mu_0 + \delta_2$ and $\mu_{r+1} = \cdots = \mu_k = \mu_0 + \delta_1$. Then, based on a sample of size cN form π_0 and N from $\pi_1, \cdots, \pi_k$, where N comes from (1.2.2), we have:

$$P\Big[CD|\boldsymbol{\mu}^0, \sigma^2, \wp\Big]$$

$$= P\Big[\bar{X}_{iN} - \bar{X}_{0cN} > d,\ \bar{X}_{jN} - \bar{X}_{0cN} < d,\ i = 1, \cdots, r,\ j = r+1, \cdots, k\Big].$$

Next, for $1 \leq i \leq r$, we write $t_i = \frac{\bar{X}_{0cN} - \bar{X}_{iN} + \delta_2}{\sqrt{\frac{\sigma^2}{N}(\frac{c+1}{c})}} / \sqrt{\frac{S_\nu^2}{\sigma^2}}$, and, for $r+1 \leq i \leq k$ we write $t_i =$

$\frac{\bar{X}_{iN}-\bar{X}_{0cN}-\delta_1}{\sqrt{\frac{\sigma^2}{N}(\frac{c+1}{c})}} / \sqrt{\frac{S_\nu^2}{\sigma^2}}$. Then, we can simplify the above expression as

$$P\Big[CD|\boldsymbol{\mu}^0, \sigma^2, \wp\Big] = P\Big[t_i < \frac{aN^{\frac{1}{2}}\sqrt{\frac{c}{c+1}}}{S_\nu}, \quad i = 1, \cdots, k\Big], \tag{1.2.3}$$

where, $(t_1, \cdots, t_k)$ follows a multivariate t distribution $f_{k,\nu,\Sigma}(\cdot)$ with $\nu = (mc-1)+k(m-1)$ degrees of freedom and correlation matrix Σ given by (1.1.8). Now, if $h_\nu = h_\nu(P^*, k, c)$ is chosen to satisfy

$$P^* = \int_{-\infty}^{h_\nu} \cdots \int_{-\infty}^{h_\nu} f_{k,\nu,\Sigma}(t_1, \cdots, t_k)\ dt_1 \cdots dt_k, \tag{1.2.4}$$

then using (1.2.3), one can immediately claim that $P\Big[CD|\boldsymbol{\mu}, \sigma^2, \wp\Big] \geq P^*$. The values of the constant $h_\nu = h_\nu(P^*, k, c)$ have been tabulated in the Table 1.2, in section 4 of this chapter.

Remark 1.2.1: Under some additional conditions one can also obtain the second-order properties for a two-stage procedure. The reader is referred to Mukhopadhyay and Duggan (1997, 1999) for details.

1.3 Purely Sequential Procedure

The purely sequential procedure starts with observations $\boldsymbol{X}_{0j}, X_{1j}, \cdots, X_{kj},\ j = 1, \cdots, m$, where m ($\geq$2) is the starting sample size from $\pi_1, \cdots, \pi_k$, and, cm is the starting sample size from π_0. After this, one takes c observations from π_0 and one observation from $\pi_1, \cdots, \pi_k$, at each step, according to the stopping rule

$$N = inf\{n \geq m :\ n \geq \frac{b^2 S_n^{*2}}{a^2}(\frac{c+1}{c})\}, \tag{1.3.1}$$

where S_n^{*2} is an estimator of σ^2 defined below. Note that in order to fully exploit the tools from Woodroofe (1977) to obtain the second-order expansions, one needs to express the estimator of σ^2 as a sum of $i.i.d.$ random variables. Based on a sample of size n from each of $\pi_1, \cdots, \pi_k$, and, cn from π_0, the following estimator S_n^{*2} is obtained along the lines of Mukhopadhyay and Solanky

(2002). We write

$$\bar{X}_{0n}^{(p)} = n^{-1}\sum_{j=1}^{n} X_{0\,(n-j)c+p}\,,\ S_{0n}^{2\ (p)} = \sum_{j=1}^{n}(X_{0\,(n-j)c+p} - \bar{X}_{0n}^{(p)})^2\,,\ p = 1,\cdots,c\,,$$

and, $\bar{X}_{jn}$, S_{jn}^2, $j = 1,\cdots,k$, are evaluated according to the expressions defined in the Section 2 of this chapter, for a sample size n. Then, we define S_n^{*2} as:

$$S_n^{*2} = \frac{\sum_{p=1}^{c} S_{0n}^{2\ (p)} + \sum_{j=1}^{k} S_{jn}^2}{c+k}.$$

Note that $(n-1)(c+k)S_n^{*2}/\sigma^2 \sim \chi^2_{(n-1)(c+k)}$, and, using the Helmert's orthogonal transformation, one can write $(n-1)(c+k)S_n^{*2}/\sigma^2 = \sum_{i=1}^{n-1} Y_i$, where $Y's$ are *i.i.d.* $\chi^2_{(k+c)}$ random variables.

Next, we put the unbalanced purely sequential procedure constructed here in a more general form and state two theorems to emphasize some important properties of the purely sequential procedure (1.3.1).

Consider a sequence $\{N_\nu : \nu \geq 1\}$ of positive integer valued random variables defined as follows:

$$N = n_\nu = \inf\{n \geq m : n \geq \psi_\nu T_n\} \tag{1.3.2}$$

where m is the starting sample size, ψ_ν is a sequence of positive constants, as $\nu \to \infty$, and $\{T-n : n \geq m\}$ are statistics such that $P(T_n) \leq 0 = 0$ for all $n \geq m$.

Lemma 1.3.1 *For the purely sequential procedure (1.3.2), if both*

$$N_\nu^{1/2}(T_{N_\nu} - a)/b \text{ and } N_\nu^{1/2}(T_{N_\nu - 1} - a)/b \tag{1.3.3}$$

converge to N(0,1) in distribution as $\nu \to \infty$, where $a(>0)$ and $b(>0)$ are constants, then we have:

$$a^{1/2}(N_\nu - a\psi_\nu)/(b\psi_\nu^{1/2}) \xrightarrow{L} N(0,1) \text{ as } \nu \to \infty.$$

This Lemma is Theorem 2.4.1 in Mukhopadhyay and Solanky (1994).

Theorem 1.3.1 *For the purely sequential procedure (1.3.1)and using the decision rule (1.1.5) based on a sample of size cN from* π_0 *and N from* $\pi_1, \cdots, \pi_k$*, we have as* $a \rightarrow 0$*:*

$$\begin{aligned} &(i) \quad N/n^* \rightarrow 1 \text{ with probability 1;} \\ &(ii) \quad E(N) \rightarrow n_c^*; \\ &(iii) \quad \lim P(CD) = P^* \text{ under the LFC;} \end{aligned}$$

where $n_c^* = \frac{b^2\sigma^2}{a^2}\frac{c+1}{c}$ *and b comes from (1.1.9).*

Proof: Using Lemma 1 of Chow and Robbins (1965), it follows that as a $\rightarrow 0$, we have $N \rightarrow \infty$ with probability 1, ${S_N^*}^2 \rightarrow \sigma^2$ with probability 1 and ${S_{N-1}^*}^2 \rightarrow \sigma^2$ with probability 1. Also, we have

$$\frac{b^2 {S_N^*}^2}{a^2}(\frac{c+1}{c}) \leq N \leq m + \frac{b^2 {S_{N-1}^*}^2}{a^2}(\frac{c+1}{c}) \tag{1.3.4}$$

Now divide throughout (1.3.4) by n_c^* and take limits as a $\rightarrow 0$. This leads to part (i).

From the right hand side of the inequality (1.3.4) it follows that

$$N \leq m + \frac{b^2}{a^2} W*$$

that is $N/n_c^* \leq m + \sigma^{-2} W*$ for sufficiently small a such that n_c^{*-1} becomes smaller than unity, where $W* = sup\{(n-1)^{-1}\sum_{i=1}^{n-1} Y_i\}$ where $Y\prime s$ are $i.i.d. \chi^2_{k+c}$ random variables, as we pointed out before. by Wiener's (1939) dominated erodic theorem one concludes that $E(W*) < \infty$. Now, the dominated converagence theorem and part (i) together imply part (ii).

From part(i), one gets $N^{1/2} a \sigma^{-1} \rightarrow h$ w.p.1 as $a \rightarrow 0$. Hence,

$$P(CD) = E[\int_{\infty}^{infty} \{Phi(y + N^{1/2} a \sigma^{-1})\}^{k-1} \phi(y/c) dy] \tag{1.3.5}$$

together with the dominated convergence theorem will lead to part(iii).

Theorem 1.3.2 *For the purely sequential procedure (1.3.1) and using the decision rule (1.1.5)* *based on a sample of size cN form* π_0 *and* N *from* $\pi_1, \cdots, \pi_k$*, we have as* $a \to 0$*:*

$$(i) \quad n_c^{*-\frac{1}{2}}(N - n_c^*) \xrightarrow{L} N(0, \tfrac{2}{k+c})\,;$$

$$(ii) \quad E(N) = n_c^* + (\nu^* - 2)(k+c)^{-1} + o(1)\,;$$

$$(iii) \quad P[CD|\mu, \sigma^2, \wp] = P^* + ((k+c)n_c^*)^{-1}\{(\nu^* - 2)g'(1) + g''(1)\} + o(n_c^{*-1})$$

$$\quad\quad if\ m > \tfrac{5}{k+c} + 1, \quad under\ the\ LFC;$$

where $n_c^* = \frac{b^2\sigma^2}{a^2}(\frac{c+1}{c})$*,* $g'(\cdot)$*, and* $g''(\cdot)$ *are defined in (1.3.7), and,* ν^* *comes from (1.3.11).*

Proof: Invoke helmert's orthogonal transformation to construct $(n-1)(c+k)S_n^{*2}/\sigma^2 = \sum_{i=1}^{n-1} Y_i$, where $Y's$ are *i.i.d.* $\chi^2_{(k+c)}$ random variables. Using Anscombe's(1952) results to claim that the sufficient conditions given in Lemma 1.3.1 hold with a=σ^2 and b=$(2/k)^{1/2}\sigma^2$. Now part (i) of this theorem follows from Lemma 1.3.1.

Next, observe that $N = Q + 1$, where

$$Q = inf\{n \geq m - 1 :\ \sum_{i=1}^{n} Y_i \leq \frac{1}{n_c^*}(c+k)n^2(1 + \frac{1}{n})\}. \tag{1.3.6}$$

Also, one can verify, $P[Y_1 < y] < By^{(k+c)/2}$, for some $B > 0$ and $\forall y > 0$. Let us define

$$\nu^* = \frac{1}{2}(k+c+2) - \sum_{n=1}^{\infty} \frac{1}{n} E[(\chi^2_{n(k+c)} - 2n(k+c))^+]\,. \tag{1.3.7}$$

Then, using the Theorem 2.4 of Woodroofe (1977), one will obtain

$$E(Q) = n_c^* + \nu^*(k+c)^{-1} - 1 - 2(k+c)^{-1} + o(1),$$

and, noting that $N = Q + 1$, the part *(ii)* of the theorem follows.

In order to verify part *(iii)*, note that for $i = 1, \cdots, r$ and $j = r+1, \cdots, k$, and, for the parametric configuration $\boldsymbol{\mu}^0$ under LFC, we have:

$$P\Big[CD|\boldsymbol{\mu}^0, \sigma^2, \wp\Big] = P\Big[\bar{X}_{iN} - \bar{X}_{0cN} > d,\ \bar{X}_{jN} - \bar{X}_{0cN} < d\Big]$$

$$= P\Big[\frac{\bar{X}_{iN} - \mu_i}{\sqrt{\sigma^2/N}} > \frac{\bar{X}_{0cN} - \mu_0}{\sqrt{\sigma^2/N}} - \frac{a}{\sqrt{\sigma^2/N}},$$

$$\frac{\bar{X}_{jN} - \mu_j}{\sqrt{\sigma^2/N}} < \frac{\bar{X}_{0cN} - \mu_0}{\sqrt{\sigma^2/N}} - \frac{a}{\sqrt{\sigma^2/N}}\Big]$$

$$= P\Big[Z_i > \frac{Z_0}{\sqrt{c}} - \frac{a\sqrt{N}}{\sigma},\ Z_j > \frac{Z_0}{\sqrt{c}} + \frac{a\sqrt{N}}{\sigma}\Big]$$

where $Z_i = \frac{\bar{X}_{iN}-\mu_i}{\sqrt{\sigma^2/N}}$, $i = 1, 2, \cdots, k$, $Z_0 = \sqrt{c}\frac{\bar{X}_{0cN}-\mu_0}{\sqrt{\sigma^2/N}}$, and $Z_0, Z_1, \cdots, Z_k$ are independent and have standard normal distributions. That is,

$$P\Big[CD|\boldsymbol{\mu}^0, \sigma^2, \wp\Big] = E\Big\{P\Big[-Z_i < \frac{a\sqrt{N}}{\sigma} - \frac{z}{\sqrt{c}},\ Z_j < \frac{a\sqrt{N}}{\sigma} + \frac{z}{\sqrt{c}}|Z_0 = z\Big]\Big\}.$$

The above expression can be expressed as

$$P\Big[CD|\boldsymbol{\mu}^0, \sigma^2, \wp\Big] = \int_{-\infty}^{\infty} \Phi^r(\frac{a\sqrt{N}}{\sigma} - \frac{z}{\sqrt{c}})\ \Phi^{k-r}(\frac{a\sqrt{N}}{\sigma} + \frac{z}{\sqrt{c}})\phi(z)dz\,, \tag{1.3.8}$$

where $\Phi(x)$ and $\phi(x)$ denotes the cdf and the pdf of the standard normal distribution, respectively. Let us write

$$\beta(x) = \int_{-\infty}^{\infty} \Phi^r\Big(\sqrt{\frac{c+1}{c}}x - \frac{z}{\sqrt{c}}\Big)\Phi^{k-r}\Big(\sqrt{\frac{c+1}{c}}x + \frac{z}{\sqrt{c}}\Big)\phi(z)dz. \tag{1.3.9}$$

Now, as in Mukhopadhyay and Solanky (1994), and, also in Datta and Mukhopadhyay (1998), we have

$$\begin{aligned}\beta'(x) \quad = \\ &\sqrt{\tfrac{c+1}{c}}\int_{-\infty}^{\infty} \Phi^{r-1}\Big(\sqrt{\frac{c+1}{c}}x - \frac{z}{\sqrt{c}}\Big)\Phi^{k-r-1}\Big(\sqrt{\frac{c+1}{c}}x + \frac{z}{\sqrt{c}}\Big) \\ &\Big\{r\phi\Big(\sqrt{\tfrac{c+1}{c}}x - \tfrac{z}{\sqrt{c}}\Big)\Phi\Big(\sqrt{\tfrac{c+1}{c}}x + \tfrac{z}{\sqrt{c}}\Big) \\ &+(k-r)\phi\Big(\sqrt{\tfrac{c+1}{c}}x + \tfrac{z}{\sqrt{c}}\Big)\Phi\Big(\sqrt{\tfrac{c+1}{c}}x - \tfrac{z}{\sqrt{c}}\Big)\Big\}\phi(z)dz,\end{aligned}$$

$$
\begin{aligned}
\beta''(x) \quad = & \\
& \tfrac{c+1}{c} \int_{-\infty}^{\infty} \Phi^{r-2}\Big(\sqrt{\frac{c+1}{c}}x - \frac{z}{\sqrt{c}}\Big)\Phi^{k-r-2}\Big(\sqrt{\frac{c+1}{c}}x + \frac{z}{\sqrt{c}}\Big) \\
& \qquad \Big[(r-1)\phi\Big(\sqrt{\tfrac{c+1}{c}}x - \tfrac{z}{\sqrt{c}}\Big)\Phi\Big(\sqrt{\tfrac{c+1}{c}}x + \tfrac{z}{\sqrt{c}}\Big) \\
& \qquad\quad +(k-r-1)\phi\Big(\sqrt{\tfrac{c+1}{c}}x + \tfrac{z}{\sqrt{c}}\Big)\Phi\Big(\sqrt{\tfrac{c+1}{c}}x - \tfrac{z}{\sqrt{c}}\Big)\Big]\phi(z)dz \\
& +\tfrac{c+1}{c} \int_{-\infty}^{\infty} \Phi^{r-1}\Big(\sqrt{\frac{c+1}{c}}x - \frac{z}{\sqrt{c}}\Big)\Phi^{k-r-1}\Big(\sqrt{\frac{c+1}{c}}x + \frac{z}{\sqrt{c}}\Big) \\
& \qquad \Big[r\phi\Big(\sqrt{\tfrac{c+1}{c}}x + \tfrac{z}{\sqrt{c}}\Big)\phi\Big(\sqrt{\tfrac{c+1}{c}}x - \tfrac{z}{\sqrt{c}}\Big) \\
& \qquad -r(\sqrt{\tfrac{c+1}{c}}x - \tfrac{z}{\sqrt{c}})\phi\Big(\sqrt{\tfrac{c+1}{c}}x - \tfrac{z}{\sqrt{c}}\Big)\Phi\Big(\sqrt{\tfrac{c+1}{c}}x + \tfrac{z}{\sqrt{c}}\Big) \\
& \qquad -(k-r)(\sqrt{\tfrac{c+1}{c}}x + \tfrac{z}{\sqrt{c}})\phi\Big(\sqrt{\tfrac{c+1}{c}}x + \tfrac{z}{\sqrt{c}}\Big)\Phi\Big(\sqrt{\tfrac{c+1}{c}}x - \tfrac{z}{\sqrt{c}}\Big) \\
& \qquad +(k-r)\phi\Big(\sqrt{\tfrac{c+1}{c}}x + \tfrac{z}{\sqrt{c}}\Big)\phi\Big(\sqrt{\tfrac{c+1}{c}}x - \tfrac{z}{\sqrt{c}}\Big)\Big]\phi(z)dz.
\end{aligned}
$$

Then, we define

$$g(x) = \beta(bx^{1/2}),\ x > 0. \tag{1.3.10}$$

It is easy to verify that

$$
\begin{aligned}
& g'(x) = \tfrac{1}{2}bx^{-1/2}\beta'(bx^{1/2}), \\
& g''(x) = \tfrac{1}{4}b[bx^{-1}\beta''(bx1/2) - x^{-3/2}\beta'(bx^{1/2})], \qquad (1.3.11) \\
& |g''(x)| \le a_1x^{-1/2} + a_2x^{-1} + a_3x^{-3/2}, a_1, a_2, a_3 \text{ being positive constants.}
\end{aligned}
$$

One may note that since $I(N = n)$ is independent of $(\bar{X}_{0cn},\ \bar{X}_{1n}, \cdots, \bar{X}_{kn})$ for all $n \ge m$, by using Theorem 3.2.1 of Mukhopadhyay and Solanky (1994), we have

$$Inf_{\boldsymbol{\mu}} P\Big[CD|\boldsymbol{\mu},\, \sigma^2,\, \wp\Big] = E\Big[g(N/n^*)\Big]. \tag{1.3.12}$$

Now, for $m > \frac{5}{k+c} + 1$, one will obtain

$$E[g(N/n^*)] \quad = \quad g(1) + n_c^{*-1}\Big[(\nu^* - 2)(k+c)^{-1}g'(1) + \frac{1}{2}\frac{2}{k+c}g''(1)\Big] + o(n_c^{*-1}),$$

which is part (iii) of the theorem.

1.4 Computations of the Design Constants and Simulations

We start this section by tabulating the values of some design constants which are needed in order to implement the procedures proposed in the sections 2 and 3 of this chapter. We also tabulate the value of constants $g'(\cdot)$ and $g''(\cdot)$ which are defined in (1.3.7), and, the constant ν^* which is defined (1.3.3). The computations of these constants will allow us to clearly explain the usage of second-order expansions obtained in the Theorem 1.3.1 (ii, iii) to the reader. We will conclude this section by simulating the two proposed procedures via Monte Carlo simulations in order to study the small and moderately large sample performances.

Table 1.1: Values of $b = b(P^*, k, c)$ as defined in (1.1.9)

	c				
k	1	2	3	5	10
1	1.64485	1.64485	1.64485	1.64485	1.64485
2	1.95993	1.95955	1.95902	1.95809	1.95680
3	2.10574	2.11592	2.11906	2.12099	2.12171
4	2.21212	2.22643	2.23074	2.23342	2.23450
5	2.28653	2.30633	2.31247	2.31651	2.31847
6	2.34897	2.37192	2.37898	2.38361	2.38587
7	2.39816	2.42483	2.43309	2.43860	2.44142
8	2.44177	2.47096	2.47995	2.48594	2.48900
9	2.47820	2.51021	2.52011	2.52674	2.53019
10	2.51146	2.54555	2.55607	2.56309	2.56675
15	2.63309	2.67635	2.68972	2.69868	2.70343
20	2.71629	2.76603	2.78141	2.79170	2.79715

In the Table 1.1, we provide the values of design constant $b = b(P^*, k, c)$ given by equation (1.1.9), for $P^* = 0.95$, $k = 1(1)10,\ 15,\ 20$, and $c = 1,\ 2,\ 3,\ 5,\ 10$, for the covariance matrix Σ defined in (1.1.8). As remarked earlier, for $c = 1$, the values of b have been also tabulated in Tong (1969) and Gibbons, Olkin, and Sobel (1977). For the sake of completeness, we have included the case $c = 1$, in the Table 1 as well, and, we must mention that the values provided in Table 1 for $c = 1$ matches with the other two sources described above. The value of b is needed in order to implement the purely sequential procedure (1.3.1) and also to compute the optimal sample size n_c^*.

Table 1.2: Values of $h_\nu = h_\nu(P^*, k, c)$ as defined in (1.2.4): $P^* = .95$

	c				
k	1	2	3	5	10
			m = 5		
1	1.85955	1.77093	1.73406	1.70113	1.67412
2	2.17806	2.10838	2.07146	2.03366	1.99829
3	2.29301	2.25802	2.23294	2.20230	2.16901
4	2.37855	2.36100	2.34268	2.31679	2.28519
5	2.43430	2.43161	2.41982	2.39935	2.37091
6	2.48299	2.48946	2.48190	2.46524	2.43927
7	2.52027	2.53487	2.53122	2.51830	2.49512
8	2.55449	2.57470	2.57381	2.56367	2.54270
9	2.58267	2.60809	2.60977	2.60225	2.58358
10	2.60913	2.63841	2.64200	2.63650	2.61971
15	2.70710	2.75024	2.76056	2.76235	2.75296
20	2.77646	2.82783	2.84189	2.84774	2.84296
			m = 10		
1	1.73427	1.70113	1.68595	1.67155	1.65909
2	2.05163	2.02532	2.01021	1.99353	1.97681
3	2.18542	2.17875	2.17063	2.15884	2.14439
4	2.28346	2.28592	2.28126	2.27202	2.25867
5	2.35022	2.36173	2.36080	2.35470	2.34336
6	2.40696	2.42391	2.42524	2.42112	2.41112
7	2.45116	2.47352	2.47714	2.47512	2.46672
8	2.49082	2.51687	2.52204	2.52146	2.51422
9	2.52375	2.55354	2.56028	2.56118	2.55520
10	2.55412	2.58667	2.59455	2.59652	2.59150
15	2.66559	2.70910	2.72137	2.72750	2.72636
20	2.74278	2.79343	2.80840	2.81698	2.81822
			m = 15		
1	1.70113	1.68107	1.67155	1.66235	1.65425
2	2.01796	2.00180	1.99215	1.98124	1.96999
3	2.15633	2.15625	2.15239	2.14566	2.13662
4	2.25751	2.26461	2.26336	2.25853	2.25037
5	2.32712	2.34189	2.34365	2.34132	2.33479
6	2.38597	2.40529	2.40881	2.40795	2.40241
7	2.43200	2.45608	2.46149	2.46228	2.45797
8	2.47311	2.50043	2.50708	2.50896	2.50548
9	2.50732	2.53803	2.54599	2.54905	2.54652
10	2.53875	2.57195	2.58086	2.58474	2.58290
15	2.65391	2.69738	2.71010	2.71731	2.71834
20	2.73328	2.78364	2.79878	2.80802	2.81083

Table 1.3: Value of $g'(1)$ as defined in (1.3.11)

	c				
k	1	2	3	5	10
1	0.08482	0.08482	0.08482	0.08482	0.08482
2	0.11452	0.11430	0.11405	0.11368	0.11323
3	0.12698	0.12869	0.12924	0.12954	0.12959
4	0.13739	0.14015	0.14106	0.14161	0.14175
5	0.14413	0.14812	0.14956	0.15055	0.15101
6	0.15030	0.15520	0.15697	0.15822	0.15881
7	0.15487	0.16070	0.16288	0.16446	0.16529
8	0.15921	0.16578	0.16827	0.17008	0.17104
9	0.16264	0.16997	0.17278	0.17488	0.17604
10	0.16595	0.17393	0.17701	0.17931	0.18060
15	0.17778	0.18847	0.19277	0.19609	0.19804
20	0.18596	0.19870	0.20393	0.20805	0.21053

Next, in the Table 1.2, we provide the values of design constant $h_\nu = h_\nu(P^*, k, c)$ which is defined in (1.2.4), for $P^* = 0.95$, $k = 1(1)10,\ 15,\ 20$, $c = 1,\ 2,\ 3,\ 5,\ 10$, and $m = 5,\ 10,\ 15$, for the covariance matrix Σ defined in (1.8). The values of b for $c = 1$ have been also tabulated in Tong (1969) and Gibbons, Olkin, and Sobel (1977). Again, we have included the case $c = 1$ in the Table 1.2 and the values provided in Table 1.2 for $c = 1$ matches with the other two sources described above. The value of h_ν is needed in order to implement the two-stage procedure (1.2.2).

In the Tables 1.3 and 1.4, we provide the values of constants $g'(1)$ and $g''(1)$, respectively for $k = 1(1)10,\ 15,\ 20$, and, $c = 1,\ 2,\ 3,\ 5,\ 10$. These constants, defined in (1.3.7), are needed to compute the asymptotic expansion provided in Theorem 1.3.1 (iii).

In the Table 1.5, we report the value of constant $\nu^* = \nu^*(k, c)$ as defined in (1.3.3). Note that since the constant ν^* depends on k and c only via $k + c$, we provide the values of ν^* for different values of $k + c$. Also, when $k + c > 60$, the second term on the right side of (1.3.3), $\sum_{n=1}^{\infty} \frac{1}{n} E[(\chi^2_{n(k+c)} - 2n(k+c))^+]$ is negligible ($\leq 2 \cdot 10^{-5}$). Therefore, in Table 5, we provide the value of ν^* for $k + c = 2(1)60$.

Table 1.4: Value of $g''(1)$ as defined in (1.3.11)

	c				
k	1	2	3	5	10
1	-0.34805	-0.27483	-0.25755	-0.24490	-0.23582
2	-0.27696	-0.27523	-0.27351	-0.27115	-0.26845
3	-0.28735	-0.29752	-0.30072	-0.30262	-0.30320
4	-0.30247	-0.32191	-0.32871	-0.33357	-0.33628
5	-0.32240	-0.34592	-0.35493	-0.36194	-0.36642
6	-0.34170	-0.36903	-0.37993	-0.38866	-0.39447
7	-0.36096	-0.39059	-0.40294	-0.41314	-0.42019
8	-0.37930	-0.41118	-0.42483	-0.43628	-0.44433
9	-0.39685	-0.43035	-0.44509	-0.45769	-0.46673
10	-0.41363	-0.44872	-0.46447	-0.47808	-0.48793
15	-0.48653	-0.52728	-0.54697	-0.56474	-0.57811
20	-0.54599	-0.59098	-0.61366	-0.63463	-0.65070

Next, in order to explain the role of second-order expansions to the reader, we look at the expansions provided in Theorem 1.3.1, parts (ii) and (iii). From Theorem 1.3.1(ii), note that as a$\rightarrow$ 0, $E(N) - n_c^* = (\nu^* - 2)(k+c)^{-1} + o(1)$. For example, for $k = 10$ the value of term $(\nu^* - 2)(k+c)^{-1}$ can be computed using the Table 1.5 as, .40187 for $c = 1$, .41880 for $c = 3$, and, .43072 for $c = 5$. Note that, these are the asymptotic values of the difference between $E(N)$ and n_c^* for the selected values of k, c, and P^*. Later in this section, we use these to evaluate the performance of the purely sequential procedure (1.3.1) for small or moderate sample sizes.

Now, we study the performance of the proposed procedures via Monte Carlo simulation studies and also compare the procedures with the balanced ones, which correspond to $c = 1$.

The two-stage procedure (1.2.2) and the purely sequential procedure (1.3.1) were simulated for $m = 10$, $c = 1, 3, 5$, $k = 10$ and $P^* = .95$, under a LFC. Without loss of generality we took $\sigma = 1$ for the purpose of generating populations. We took $\delta_1 = -\delta_2$, giving $a = \delta_2 (= \delta$, say). Next, using $n_c^* = \frac{b^2\sigma^2}{a^2}(\frac{c+1}{c})$, we computed the values of δ corresponding to $n_c^* = 25, 100, 200, 400$ and 800. Then, each procedure was independently repeated 1000 times. The performance of the two-stage procedure (1.2.2) is summarized in the Table 1.6 and that of the purely sequential procedure (1.3.1)

Table 1.5: Values of ν^* as defined in (1.3.7)

2	3	4	5	6	7
1.49000	2.10441	2.68634	3.24766	3.79489	4.33199
8	9	10	11	12	13
4.86155	5.38538	5.90474	6.42058	6.93362	7.44440
14	15	16	17	18	19
7.95334	8.46078	8.96699	9.47218	9.97653	10.48019
20	21	22	23	24	25
10.98326	11.48585	11.98803	12.48986	12.99142	13.49273
26	27	28	29	30	31
13.99383	14.49477	14.99556	15.49624	15.99680	16.49730
32	33	34	35	36	37
16.99770	17.49805	17.99834	18.49859	18.99880	19.49898
38	39	40	41	42	43
19.99913	20.49926	20.99938	21.49947	21.99955	22.49962
44	45	46	47	48	49
22.99967	23.49972	23.99976	24.49980	24.99983	25.49985
50	51	52	53	54	55
25.99988	26.49989	26.99991	27.49992	27.99993	28.49994
56	57	58	59	60	
28.99995	29.49996	29.99997	30.49997	30.99998	

(The value on top is $(k+c)$ and below it is ν^*)

in the Table 1.7. In the Tables 1.6 and 1.7, we report the values of n_c^*, $\delta, \bar{n}_\pi$: the average sample size from $\pi_1, \cdots, \pi_k$, and $\bar{n}_t$: the average sample size from $\pi_0, \pi_1, \cdots, \pi_k$, and, $\bar{P}$: the proportion of times all the k populations are partitioned correctly. We also report the standard errors of the reported estimates.

Note that as expected, using Theorems 1.2.1 and 1.3.1, the value of $\bar{P}$ is close to or above the target value of 0.95, for all the cases considered and for both the procedures. One should note that one of the inbuilt advantages of taking a larger sample from the control population is to compensate for a smaller sample sizes from the other populations $\pi_1, \cdots, \pi_k$. This is clearly evident for the cases $c = 3$ and $c = 5$ by comparing the values of $\bar{n}_\pi$ and $\bar{n}_t$, in the Tables 1.6 and 1.7.

From Table 1.6, it is evident that the two-stage procedure is oversampling compared to the

Table 1.6: Performance of the two-stage procedure (1.2.2)

c=1

n_c^*	δ	$\bar{n}_\pi$ $s(\bar{n}_\pi)$	$\bar{n}_t$ $s(\bar{n}_t)$	$\bar{p}$ $s(\bar{p})$
25	0.7104	25.402	25.402	0.948
		(0.1152)	(0.1152)	(0.0070)
100	0.3552	102.026	102.026	0.950
		(0.4391)	(0.4391)	(0.0069)
200	0.2512	207.277	207.277	0.941
		(0.9167)	(0.9167)	(0.0075)
400	0.1776	412.936	412.936	0.955
		(1.8806)	(1.8806)	(0.0066)
800	0.1256	828.887	828.887	0.951
		(3.7084)	(3.7084)	(0.0068)

c=3

25	0.5903	25.380	29.995	0.954
		(0.1044)	(0.1458)	(0.0066)
100	0.2952	102.785	121.473	0.954
		(0.4053)	(0.5661)	(0.0066)
200	0.2087	204.828	242.069	0.949
		(0.8041)	(1.1231)	(0.0070)
400	0.1476	411.029	485.762	0.955
		(1.6920)	(1.9996)	(0.0066)
800	0.1044	823.220	972.896	0.955
		(3.3405)	(4.6657)	(0.0066)

c=5

25	0.5616	24.997	34.087	0.949
		(0.0983)	(0.1828)	(0.0070)
100	0.2808	101.858	138.897	0.948
		(0.3840)	(0.7140)	(0.0070)
200	0.1986	206.100	281.045	0.958
		(0.8012)	(1.4898)	(0.0063)
400	0.1404	410.180	599.336	0.957
		(1.5499)	(2.8820)	(0.0064)
800	0.0993	818.437	1116.050	0.955
		(3.2087)	(5.9666)	(0.0066)

$k = 10$, $P^* = .95$, and $m = 10$

Table 1.7: Performance of the purely sequential procedure (1.3.1)

c=1

n_c^*	δ	$\bar{n}_\pi$ $s(\bar{n}_\pi)$	$\bar{n}_t$ $s(\bar{n}_t)$	$\bar{p}$ $s(\bar{p})$
25	0.7104	25.496	25.496	0.941
		(0.0687)	(0.0687)	(0.0075)
100	0.3552	100.287	100.287	0.949
		(0.1364)	(0.1364)	(0.0070)
200	0.2512	200.713	200.713	0.953
		(0.1904)	(0.1904)	(0.0067)
400	0.1776	400.354	400.354	0.947
		(0.2610)	(0.2610)	(0.0071)
800	0.1256	800.180	800.180	0.947
		(0.3826)	(0.3826)	(0.0071)
		$[n_c^*+.402]^a$		

c=3

n_c^*	δ	$\bar{n}_\pi$ $s(\bar{n}_\pi)$	$\bar{n}_t$ $s(\bar{n}_t)$	$\bar{p}$ $s(\bar{p})$
25	0.5903	25.485	30.119	0.941
		(0.0687)	(0.0960)	(0.0075)
100	0.2952	100.347	118.619	0.953
		(0.1264)	(0.1765)	(0.0067)
200	0.2087	200.654	237.137	0.958
		(0.1754)	(0.2450)	(0.0063)
400	0.1476	400.347	473.173	0.958
		(0.2408)	(0.3363)	(0.0063)
800	0.1044	800.360	945.88	0.947
		(0.3674)	(0.5131)	(0.0071)
		$[n_c^*+.419]^a$		

c=5

n_c^*	δ	$\bar{n}_\pi$ $s(\bar{n}_\pi)$	$\bar{n}_t$ $s(\bar{n}_t)$	$\bar{p}$ $s(\bar{p})$
25	0.5616	25.394	34.628	0.942
		(0.0651)	(0.1211)	(0.0074)
100	0.2808	100.410	136.923	0.956
		(0.1264)	(0.2350)	(0.0067)
200	0.1986	200.716	273.704	0.963
		(0.1754)	(0.3262)	(0.0063)
400	0.1404	400.311	545.879	0.966
		(0.2407)	(0.4476)	(0.0063)
800	0.0993	800.362	1091.402	0.941
		(0.3262)	(0.6066)	(0.0075)
		$[n_c^*+.431]^a$		

$k = 10$, $P^* = .95$, and $m = 10$

(^{a}: denotes the asymptotic value from Theorem 1.3.1(ii)) optimal sample size. For example, for $c = 1$ and $n_c^* = 800$, the two stage procedure oversamples by 29 or so observations. Such a behavior of the two-stage procedures is well documented in the statistical literature. One way to eliminate over-sampling is to adopt a purely sequential procedure. Note that in the Table 1.7, the values of n_c^* and $\bar{n}_\pi$ are quite close and there does not appear to be any oversampling. In the Table 1.7, we also provide the asymptotic value of $E(N) - n_c^*$. Note that even for small sample sizes, such as 25 or 100, the agreement between the asymptotic value and the observed values is remarkable for all the three cases. In addition, using the Theorem 1.3.1(iii) and the Tables provided in this section, one can easily verify that the observed $\bar{p}$ value is in agreement with the asymptotic value. For example, in Table 1.7, for $c = 5$, we expect $\bar{p} - P^*$ to be close to $((k+c)n_c^*)^{-1}\{(\nu^* - 2)g'(1) + g''(1)$ (=.0004536) and the observed difference is .006 with standard error of .0067.

Remark 1.4.1: It is important to note that within the Table 1.6 , and, also within the Table 1.7, one cannot compare the blocks corresponding to different values of c with one another. This is so because even though the n_c^* values are same in the three blocks, the value of δ is smaller for the larger c value. Also, note that in Tables 6 and 7, for $c = 3$ and $c = 5$, the value of $\bar{n}_t$ is significantly larger than that of $\bar{n}_\pi$, as we take more observations from π_0. In other words, since n_c^* denotes the optimal sample size from $\pi_1, \cdots, \pi_k$, it needs to be compared with $\bar{n}_\pi$. An alternative way to compute $\bar{n}_t$ would be to divide the number of samples collected from π_0 by c before computing the average.

Chapter 2

Assessing Robustness of Procedures

2.1 Introduction

In real world applications, the partition problem is a routine problem which gets applied in numerous different areas, such as, biological sciences, medical sciences, agricultural sciences, etc, to name a few, in order to compare newer treatments with a control. However, a large proportion of the statistical theory is developed for the normal distribution case and also under various assumptions.

In this chapter, we consider the robustness of various partition procedures known in the statistical literature, including the ones proposed in Chapter 1, from the point of view of mild to moderate departures from the assumptions. The goal of the study is to document the performance of the different procedures under such several mild/moderate departures.

2.2 Description of The Procedures

In this chapter, we have selected a few procedures to study the robustness issues. It should be noted that the literatures is quite rich and has many such procedures and inclusion of all such procedures in not practical. However, we have selected a few, to illustrate our point. The selected procedures are somewhat the standard procedures and have been cited regularly in the statistical literature.

Two Stage Procedure (DS): In a two-stage procedure, samples are collected in two batches.

The procedure described below was developed by Tong (1969).

Let $m > 1$ be a pre-assigned positive integer indicating the starting sample size. We collect m observations from each of the $k+1$ populations, and compute the estimator of σ^2 given by

$$S^2 = \nu^{-1} \sum_{i=0}^{k} \sum_{j=1}^{m} [X_{ij} - m^{-1}(\sum_{n=1}^{m} X_{ij})]^2$$

where $\nu = (k+1)(m-1)$. After this, in the second stage we collect $N - m$ additional samples from each population, where N is the smallest integer satisfying

$$N \geq max\{m, < 2h_\nu^2 S_\nu^2 / a^2 >\}. \tag{2.2.1}$$

Then we partition the k populations based on N samples using (1.1.5). Note that h_ν is available in the Table 2 in Tong (1969).

Three Stage Procedure (TS): In a three-stage procedure, the samples are collected in three batches. The procedure stated below and its fine tuned version were developed by Datta and Mukhopadhyay(1998).

Choose and fix $\rho \in (0, 1)$, collect m observations from each population as the starting sample size, and compute

$$T = max\{m, < 2\rho b^2 S_m^2 a^{-2} > +1\}.$$

Collect $T - m$ additional samples from each populations in the second batch and compute:

$$N = max\{T, < 2b^2 S_T^2 a^{-2} > +1\}.$$

In the third batch, we collect $N - T$ additional samples from each population, compute overall sample means and apply the same decision as described in (1.1.5), where $< x >$=largest integer $< x$ and b is available in the Table 1 in Tong (1969).

Fined Tuned Three Stage Procedure (TSR):

Choose and fix $\rho \in (0, 1)$, collect m observations from each population as the starting sample size, and compute

$$T = max\{m, < 2\rho b^2 S_m^2 a^{-2} > +1\}.$$

Collect $T - m$ additional samples from each populations in the second batch and compute:

$$N = max\{T, < 2b^2 S_T^2 a^{-2} + \epsilon > +1\}.$$

In the third batch, we collect $N - T$ additional samples from each population, compute overall sample means and apply the same decision as described in (1.1.5), where $< x >$=largest integer $< x$, b is available in table 1 in Tong (1969), and $\epsilon = \rho^{-1}(k+1)^{-1}[2 - \{g''(1)/g'(1)\}] - 1/2$. Here $g(\cdot)$ is a special case for $c = 1$ of the $g(\cdot)$, which is defined in the Chapter 1.

Purely Sequential Procedure (PS): This procedure and its fine tuned version were developed by Datta and Mukhopadhyay (1998).

Define $N = inf\{n \geq m : \; n \geq 2b^2 S_n^2/a^2\}$. Then apply the decision rule as described in (1.1.5) based on N samples, where b is available in the Table 1 in Tong (1969).

Fined Tuned Purely Sequential Procedure (PSR):

Define $N = inf\{n \geq m : \; n + \epsilon \geq 2b^2 S_n^2/a^2\}$. Then apply the decision rule as described in (1.1.5) based on N samples, where b is available in the Table 1 in Tong (1969). Where, the constant $\epsilon = (k+1)^{-1}[(\nu - 2) + g''(1)/g'(1)]$ and $g(\cdot)$ is same as the one introduced for the fine tuned three stage procedure.

Purely Sequential Procedure with Elimination (ES): This procedure can eliminate and partition "inferior" or "superior" populations using triangular boundaries. It was developed by

Solanky (2001). This procedure has the following steps.

(1) Start with the sample size $m\,(>1)$ samples from each population, compute:

$$\overline{X}_{i\ m} = \sum_{j-1}^{m} X_{i\ j}/m,\ S_{i\ m}^2 = \sum_{j=1}^{m}(X_{i\ j} - \overline{X}_{i\ m})^2/(m-1),$$

$$S^2 = \sum_{i=0}^{k} S_{i\ m}^2/k,\ a_\lambda = \eta f S^2/a,\ W_\lambda = [a_\lambda/\lambda].$$

(2) Draw one observation from those populations, which have not been eliminated, until

(i) $m \geq W_\lambda$, or

(ii) all the populations have been partitioned,

and then do step (4).

(3) Within each population that to be partitioned, partition any populations into S_B for which

$$\sum_{j=1}^{r} X_{i\ j} < \sum_{j=1}^{r} (X_{0\ j} + d - a_\lambda + r\lambda),$$

partition any populations into S_G for which

$$\sum_{j=1}^{r} X_{i\ j} > \sum_{j=1}^{r} (X_{0\ j} + d + a_\lambda - r\lambda).$$

(4) if $m = W_\lambda$, then get one more observation from the populations haven't been partitioned, and them apply decision rule (1.1.5) to those populations. Here η could be found in table 1 in Solanky (2001) and $\lambda = a/(2j)$.

Unbalanced Procedures

There are two kinds of unbalanced procedures, which are the two-stage unbalanced procedure **(UDS)** and purely sequential unbalanced procedure **(UPS)**. The details of these procedures are provided in the Chapter 1 of this thesis.

2.3 Performance of the Procedures

We start by simulating all the selected procedures when all of the assumptions are satisfied. We generated $k+1$ groups of samples, which are independent within each group and from independent and normally distributed populations. We choose $k = 10$, and the populations are assumed under the LFC and the variance of the populations was taken to be 1, without the loss of generality. We took $\mu_0 = 0$ and $\delta_1 = -\delta_2$, giving $a = \delta_2 (= \delta$, say). Next, using $n^* = \frac{2b^2\sigma^2}{a^2}$, we computed the values of δ corresponding to $n^* = 15, 20, 25, 30, 50, 100, 200, 400$ and 800. The values are in the following table:

Table 2.1: Values of δ for specified optional sample sizes

n^*	15	20	25	30	50	100	200	400	800
$\delta = -\delta_1 = \delta_2$	0.9171	0.7942	0.7103	0.6485	0.5023	0.3552	0.2511	0.1776	0.1256

We took the starting sample size to be $m = 10$, $P^* = .95$, $\rho = \frac{1}{4}$, for TS and TSR and $j = 2$ for the ES. For each value of δ, each procedure was independently repeated for 5000 times and we recorded the sample size as well as whether the partition is a CD or not, in each iteration. The average sample sizes and the actual percentages of CD for the procedures are displayed in the Table 2.2.

To summarize, all the procedures are working quite well with the estimated probability of CD ($= \bar{p}$) being close to its target value and the average sample size ($= \bar{n}$) being close to its optimal value n^*.

2.4 Departure from Normality

2.4.1 Symmetric Distributions Case

In this section, we restrict our attention to symmetric distributions only. For the underlying distributions, we chose a variety of non-normal distributions to represent a wide range of symmetric configuration with varying degrees of "heaviness" in the tails.

Table 2.2: Simulation Results under Normal Distribution Assumption

	DS		TS		TSR		UDS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	15.0562	0.9448	15.5416	0.9494	16.6230	0.9618	10.6448	0.9546
	0.03168	0.00323	0.03072	0.00310	0.03044	0.00271	0.01351	0.00294
20	20.1540	0.9438	20.4766	0.9450	21.6296	0.9550	13.7272	0.9410
	0.04193	0.00326	0.04054	0.00322	0.04039	0.00293	0.02599	0.00333
25	25.2774	0.9466	25.4224	0.9494	26.5636	0.9504	17.2460	0.9448
	0.05170	0.00318	0.04995	0.00310	0.05044	0.00307	0.03260	0.00323
30	30.6572	0.9510	30.1223	0.9512	31.5738	0.9542	21.0410	0.9518
	0.06630	0.00305	0.06231	0.03050	0.06011	0.00296	0.03920	0.00303
50	51.2702	0.9496	49.7364	0.9440	50.8938	0.9462	35.0980	0.9430
	0.10548	0.00309	0.08835	0.00325	0.08799	0.00319	0.06591	0.00328
100	102.8248	0.9508	99.6156	0.9490	100.7132	0.9436	70.6572	0.9446
	0.20593	0.00306	0.12375	0.00311	0.12254	0.00326	0.12997	0.00324
200	206.5910	0.9528	199.8712	0.0947	200.9702	0.9522	141.9520	0.9518
	0.42360	0.00300	0.17729	0.00316	0.17288	0.00302	0.26548	0.00303
400	412.0918	0.9472	399.6728	0.9484	400.7262	0.9544	283.2800	0.9466
	0.83922	0.00316	0.24720	0.00313	0.24978	0.00295	0.52596	0.00318
800	825.8844	0.9508	799.6728	0.9512	800.2626	0.9478	567.5690	0.9496
	1.65895	0.00306	0.34115	0.00305	0.34602	0.00315	1.04258	0.00309

	PS		PSR		ES		UPS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	15.3840	0.9484	15.2130	0.9466	10.7336	0.9815	11.0140	0.9576
	0.02520	0.00313	0.02530	0.00318	0.01093	0.00191	0.01464	0.00285
20	20.4332	0.9532	20.2496	0.9502	11.9840	0.9760	14.2140	0.9518
	0.02834	0.00299	0.02843	0.00308	0.01920	0.00216	0.02199	0.00303
25	25.3502	0.9522	25.1702	0.9510	13.6548	0.9752	17.7180	0.9466
	0.03162	0.00302	0.03179	0.00305	0.02713	0.00220	0.02449	0.00318
30	30.3360	0.9468	30.1704	0.9460	15.6448	0.9680	21.1066	0.9504
	0.03377	0.00317	0.03379	0.00320	0.03485	0.00249	0.02638	0.00307
50	50.2884	0.9468	50.1086	0.9452	24.6672	0.9682	34.8538	0.9478
	0.04348	0.00317	0.04355	0.00322	0.06377	0.00248	0.03340	0.00315
100	100.5278	0.9518	100.3614	0.9520	48.2293	0.9646	69.5042	0.9474
	0.06001	0.00303	0.06000	0.00302	0.13131	0.00261	0.04750	0.00316
200	200.5360	0.9492	200.3634	0.9494	94.8163	0.9610	138.5652	0.9516
	0.08566	0.00311	0.08582	0.00310	0.26151	0.00274	0.06457	0.00304
400	400.3552	0.9478	400.1814	0.9480	187.2401	0.9642	276.5646	0.9460
	0.12283	0.00315	0.12292	0.00314	0.50991	0.00263	0.09209	0.00320
800	800.1210	0.9484	799.9554	0.9486	372.4971	0.9646	553.0370	0.9516
	0.17098	0.00313	0.17093	0.00312	1.03014	0.00261	0.12900	0.00304

We included two *Student t* distributions with 10 and 20 *d.f.* respectively, a Laplace distribution and three mixture of normal distributions with 2, 4 and 6 squared Mahalanobis distances, respectively, to represent a family of distributions with varying kurtosis. The *cdf* of these mixture-normal distributions was

$$F_i(x) = \pi_{i1}N(x; \mu_{i1}, \sigma^2) + \pi_{i2}N(x; \mu_{i2}, \sigma^2), \tag{2.4.1}$$

where $N(x; \mu_{i1}, \sigma^2)$ and $N(x; \mu_{i2}, \sigma^2)$ are Gaussian random variables with locations (μ_{i1}, μ_{i2}) and a common covariance, σ^2 and $\pi_{i1} + \pi_{i2} = 1$ are the mixing proportions. To specify a mixture-normal distribution with a given mean and variance 1, we define

$$\Delta = (\mu_{i1} - \mu_{i2})^2\sigma^{-2} \tag{2.4.2}$$

as squared Mahalanobis distance associated with the distribution. We choose the mean and variance of the component normal distributions, such that the mixture-normal distribution will have squared Mahalanobis distances as 2, 4 and 6 respectively. The parameters for such distribution with mean $\mu = 0$ are in the Table 2.3:

Table 2.3: The Parameters for Symmetrical Mixture-Normal Distributions

	$\Delta = 2$	$\Delta = 4$	$\Delta = 6$
$\pi_1 = 0.5$	$\mu_1 = -\mu_2 = 0.70107$	$\mu_1 = -\mu_2 = 0.89443$	$\mu_1 = -\mu_2 = 0.94868$
	$\sigma = 0.70107$	$\sigma = 0.44721$	$\sigma = 0.31623$

As before, we choose $k = 10$, and assumed the k non-control populations have some non-normal distributions in the same location family and the control population still has the standard normal distribution. Also, we assumed that the populations are independent and they are under LFC and have common variance $\sigma^2 = 1$, without the loss of generality. Note that the mixture-normal distributions with the parameters given in Table 2.3 have variance $\sigma^2 = 1$. Hence, the distributions of the non-control populations are in the location families of these distributions. And for the *student t* and Laplace distributions, we chose the ones with variance $\sigma^2 = 1$ from their location-scale family to be the location families, which include the distributions of the non-control populations. The value of δ_1 and δ_2 were set in the same way as in the first paragraph in Section 2.3 to specify the distributions of the non-control populations from the location families described above. Then we generated $k+1$ groups of samples from such populations, each group is corresponding to one population, for our robustness study.

We took the starting sample size to be $m = 10$, $P^* = .95$, $\rho = \frac{1}{4}$, for TS and TSR and $j = 2$ for the ES. For each distribution we mentioned above and each value of δ, each procedure was repeated 5000 times independently, and the value of $\bar{n}$ and $s(\bar{n})$, $\bar{p}$ and $s(\bar{p})$ are recorded and displayed in Tables 2.4 - 2.9.

From the Tables 2.7, 2.8 and 2.9, we see that for the distributions with lighter tails than normal distributions, i.e., with smaller kurtosis, the procedures perform well. The values of $\bar{p}$ are increasing as the tails of the distributions become lighter.

The Table 2.4, 2.5 and 2.6, indicate that the performance of the procedures with heavy tails. Generally, the heavier the tail, the smaller the $\bar{p}$ for all the procedures and in all the cases. We also found that the ES is robust to the heavy-tailedness. Since the ES procedure is based on some inequalities and the $\bar{p}$ values are generally overshooting the target, the ES procedure works well even under such moderate violation. The validity of the UPS procedure is marginally affected by heavy-tailedness. The performance of the other six procedures is moderately affected by heavy-tailedness.

From the Tables, it is difficult to pin point the exact robustness of the procedures. However, it is important to note that the worst performance is for TSR with $n^* = 50$, giving $\bar{p} = 0.9364$. Note that this worst case is well within 2 standard errors of the target value. Hence, our conclusion is that the procedures considered are quite rubust to mild to moderate heavy/light tailedness violations.

Table 2.4: Simulation Results for t_{20} Distribution

	DS		TS		TSR		UDS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	15.0024	0.9418	15.4900	0.9498	16.6312	0.9644	10.6444	0.9514
	0.03329	0.00331	0.03213	0.00309	0.03321	0.00262	0.01379	0.00304
20	20.1856	0.9462	20.5038	0.9484	21.6098	0.9606	13.7262	0.9438
	0.04440	0.00319	0.04287	0.00313	0.04253	0.00275	0.02751	0.00326
25	25.2694	0.9490	25.4130	0.9500	26.6638	0.9538	17.2460	0.9458
	0.05635	0.00311	0.05449	0.00308	0.05371	0.00297	0.03523	0.00320
30	30.6144	0.9474	30.5438	0.9476	31.6022	0.9508	20.9094	0.9426
	0.06818	0.00316	0.06449	0.00315	0.06340	0.00306	0.04238	0.00329
50	51.0944	0.9436	49.5350	0.9406	50.7556	0.9510	35.0158	0.9436
	0.11229	0.00326	0.09412	0.00334	0.09291	0.00305	0.06971	0.00326
100	102.6926	0.9482	99.5318	0.9472	100.6496	0.9480	70.5460	0.9512
	0.21929	0.00313	0.13294	0.00316	0.13474	0.00314	0.13708	0.00305
200	205.8878	0.9478	199.6950	0.9482	200.9040	0.9518	141.5770	0.9446
	0.45609	0.00315	0.19233	0.00313	0.18828	0.00303	0.28342	0.00324
400	413.2788	0.9488	399.4182	0.9514	401.1600	0.9562	284.2158	0.9510
	0.90211	0.00312	0.26372	0.00304	0.26234	0.00289	0.55626	0.00305
800	827.1610	0.9420	799.3644	0.9452	800.9254	0.9510	568.7584	0.9486
	1.82051	0.00331	0.36993	0.00322	0.37371	0.00305	1.12719	0.00312

	PS		PSR		ES		UPS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	15.4140	0.9470	15.2336	0.9460	10.7394	0.9796	11.0422	0.9616
	0.02669	0.00317	0.02673	0.00320	0.01112	0.00200	0.01547	0.00272
20	20.3616	0.9498	20.1926	0.9484	11.9849	0.9778	14.2002	0.9528
	0.03043	0.00309	0.03069	0.00313	0.01961	0.00208	0.02315	0.00300
25	25.3738	0.9504	25.1986	0.9506	13.6879	0.9716	17.6798	0.9488
	0.03355	0.00307	0.03361	0.00306	0.02808	0.00235	0.02569	0.00312
30	30.3482	0.9540	30.1722	0.9540	15.5872	0.9710	21.0728	0.9538
	0.03702	0.00296	0.03723	0.00296	0.03566	0.00237	0.02827	0.00297
50	50.4142	0.9482	50.2378	0.9480	24.6208	0.9656	34.9134	0.9490
	0.04710	0.00313	0.04706	0.00314	0.06701	0.00258	0.03558	0.00311
100	100.3500	0.9524	100.1764	0.9518	47.9847	0.9664	69.4396	0.9492
	0.06624	0.00301	0.06623	0.00303	0.13321	0.00255	0.04993	0.00311
200	200.3658	0.9502	200.1868	0.9506	94.8547	0.9662	138.5380	0.9524
	0.09233	0.00308	0.09232	0.00306	0.27186	0.00256	0.07047	0.00301
400	400.2880	0.9480	400.1088	0.9480	186.3133	0.9564	276.5396	0.9478
	0.13098	0.00314	0.13102	0.00314	0.54313	0.00289	0.09826	0.00315
800	800.4784	0.9470	800.2932	0.9468	371.3969	0.9562	552.8602	0.9514
	0.18607	0.00317	0.18611	0.00317	1.06446	0.00289	0.13985	0.00304

Table 2.5: Simulation Results for t_{10} Distribution

	DS		TS		TSR		UDS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	15.0414	0.9462	15.5194	0.9504	16.6158	0.9580	10.7104	0.9534
	0.03716	0.00319	0.03608	0.00307	0.03640	0.00284	0.01570	0.00298
20	20.1950	0.9458	20.5152	0.9510	21.6144	0.9486	13.7524	0.9438
	0.04849	0.00320	0.04678	0.00305	0.04793	0.00312	0.02956	0.00326
25	25.4304	0.9472	25.5588	0.9474	26.7004	0.9598	17.3444	0.9442
	0.06260	0.00316	0.06031	0.00316	0.05915	0.00278	0.03850	0.00325
30	30.6018	0.9508	30.5184	0.9498	31.4970	0.9532	20.8838	0.9442
	0.07407	0.00306	0.06986	0.00309	0.06932	0.00299	0.04561	0.00325
50	51.2114	0.9462	49.5132	0.9406	50.6020	0.9506	35.1008	0.9460
	0.12259	0.00319	0.10240	0.00334	0.10268	0.00306	0.07537	0.00320
100	102.9150	0.9504	99.3822	0.9442	100.1190	0.9420	70.6942	0.9452
	0.24633	0.00307	0.15121	0.00325	0.14953	0.00331	0.15011	0.00322
200	205.6234	0.9444	199.1190	0.9438	200.3764	0.9530	141.3906	0.9508
	0.48899	0.00324	0.20941	0.00326	0.20819	0.00299	0.30073	0.00306
400	413.5736	0.9434	399.3724	0.9438	400.7564	0.9476	284.7942	0.9500
	1.03575	0.00327	0.29783	0.00326	0.29242	0.00315	0.62894	0.00308
800	826.0084	0.9436	799.5828	0.9506	801.4692	0.9486	567.8116	0.9428
	1.97272	0.00326	0.41629	0.00306	0.42254	0.00312	1.21229	0.00328

	PS		PSR		ES		UPS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	15.3318	0.9510	15.1532	0.9478	10.7509	0.9800	11.0424	0.9524
	0.02920	0.00305	0.02928	0.00315	0.01161	0.00198	0.01642	0.00301
20	20.2626	0.9460	20.0852	0.9446	11.9635	0.9742	14.2442	0.9496
	0.03350	0.00320	0.03383	0.00324	0.02041	0.00224	0.02496	0.00309
25	25.3076	0.9516	25.1230	0.9506	13.6690	0.9730	17.6604	0.9544
	0.03655	0.00304	0.03670	0.00306	0.03015	0.00229	0.02872	0.00295
30	30.3390	0.9544	30.1632	0.9520	15.7363	0.9644	21.0422	0.9506
	0.03983	0.00295	0.03998	0.00302	0.03917	0.00262	0.03015	0.00306
50	50.3626	0.9510	50.1842	0.9510	24.7543	0.9650	34.8968	0.9520
	0.05196	0.00305	0.05209	0.00305	0.07137	0.00260	0.03868	0.00302
100	100.3638	0.9576	100.1866	0.9570	47.9579	0.9614	69.3964	0.9464
	0.07302	0.00285	0.07305	0.00287	0.14168	0.00272	0.05480	0.00319
200	200.3668	0.9474	200.1964	0.9480	94.3464	0.9620	138.5110	0.9548
	0.10386	0.00316	0.10396	0.00314	0.28114	0.00270	0.07716	0.00294
400	400.4982	0.9490	400.3254	0.9490	186.7014	0.9578	276.6496	0.9522
	0.14545	0.00311	0.14537	0.00311	0.57102	0.00284	0.10927	0.00302
800	800.7582	0.9530	800.5872	0.9528	371.7556	0.9642	552.7822	0.9494
	0.20511	0.00299	0.20518	0.00300	1.12777	0.00263	0.15239	0.00310

Table 2.6: Simulation Results for Laplace Distribution

	DS		TS		TSR		UDS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	15.0680	0.9370	15.5354	0.9444	16.6516	0.9598	10.8632	0.9508
	0.04599	0.00344	0.04480	0.00324	0.04536	0.00278	0.01920	0.00306
20	20.1848	0.9428	20.4912	0.9470	21.6520	0.9526	13.7868	0.9396
	0.06267	0.00328	0.06050	0.00317	0.05928	0.00301	0.03629	0.00337
25	25.3860	0.9420	25.5094	0.9444	26.5384	0.9546	17.2860	0.9384
	0.07714	0.00331	0.07373	0.00324	0.07490	0.00294	0.04595	0.00340
30	30.4968	0.9418	30.3122	0.9410	31.4456	0.9504	20.8304	0.9394
	0.09307	0.00331	0.08494	0.00333	0.08593	0.00307	0.05563	0.00337
50	51.3894	0.9412	49.2110	0.9370	50.3252	0.9432	35.1426	0.9430
	0.15312	0.00333	0.12471	0.00344	0.12572	0.00327	0.09162	0.00328
100	103.1948	0.9436	98.7002	0.9396	99.7328	0.9452	70.8246	0.9450
	0.30878	0.00326	0.18795	0.00337	0.18853	0.00322	0.18572	0.00322
200	205.7662	0.9466	198.2938	0.9498	200.1640	0.9450	141.3786	0.9410
	0.62266	0.00318	0.26657	0.00309	0.26400	0.00322	0.37134	0.00333
400	411.7306	0.9430	398.4490	0.9524	399.0816	0.9550	283.0924	0.9484
	1.24781	0.00328	0.37917	0.00301	0.38075	0.00293	0.74708	0.00313
800	826.4782	0.9414	798.7574	0.9470	799.8952	0.9490	568.7938	0.9466
	2.46818	0.00332	0.54711	0.00317	0.53572	0.00311	1.46721	0.00318

	PS		PSR		ES		UPS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	15.2296	0.9458	15.0444	0.9428	10.7962	0.9766	11.1348	0.9548
	0.03596	0.00320	0.03594	0.00328	0.01282	0.00214	0.01885	0.00294
20	20.2624	0.9452	20.0806	0.9434	12.0138	0.9710	14.0720	0.9498
	0.04245	0.00322	0.04256	0.00327	0.02293	0.00237	0.03051	0.00309
25	25.2372	0.9434	25.0656	0.9420	13.7042	0.9682	17.5930	0.9418
	0.04812	0.00327	0.04839	0.00331	0.03385	0.00248	0.03421	0.00331
30	30.2208	0.9480	30.0352	0.9462	15.6556	0.9658	20.9754	0.9518
	0.05086	0.00314	0.05091	0.00319	0.04326	0.00257	0.03779	0.00303
50	50.2000	0.9486	50.0232	0.9484	24.6831	0.9648	34.8032	0.9472
	0.06578	0.00312	0.06584	0.00313	0.08269	0.00261	0.04778	0.00316
100	100.2728	0.9488	100.0888	0.9480	48.0851	0.9586	69.4130	0.9524
	0.09313	0.00312	0.09322	0.00314	0.16771	0.00282	0.06793	0.00301
200	200.1482	0.9540	199.9750	0.9546	94.3808	0.9558	138.4542	0.9456
	0.13059	0.00296	0.13069	0.00294	0.33727	0.00291	0.09708	0.00321
400	400.0282	0.9454	399.8514	0.9454	186.3922	0.9550	276.5936	0.9518
	0.18563	0.00321	0.18575	0.00321	0.65378	0.00293	0.13509	0.00303
800	800.0730	0.9552	799.9012	0.9560	369.6602	0.9554	552.6472	0.9490
	0.26146	0.00293	0.26153	0.00290	1.31944	0.00292	0.19226	0.00311

Table 2.7: Simulation Results for Mixture-Normal Distribution with $\Delta = 2$

	DS		TS		TSR		UDS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	14.7822	0.9424	15.3096	0.9530	16.3676	0.9578	10.5068	0.9570
	0.02802	0.00330	0.02669	0.00299	0.02622	0.00284	0.01132	0.00287
20	19.8894	0.9434	20.1820	0.9506	21.3308	0.9568	13.5282	0.9400
	0.03669	0.00327	0.03530	0.00306	0.03607	0.00288	0.02351	0.00336
25	24.9434	0.9410	25.0874	0.9484	26.2310	0.9526	17.0688	0.9500
	0.04551	0.00333	0.04483	0.00313	0.04494	0.00301	0.02979	0.00308
30	30.0744	0.9472	30.0270	0.9468	31.2068	0.9552	20.6028	0.9486
	0.05516	0.00316	0.05289	0.00317	0.05359	0.00293	0.03532	0.00312
50	50.4348	0.9502	49.1058	0.9430	50.2124	0.9484	34.6038	0.9502
	0.09208	0.00308	0.07757	0.00328	0.07812	0.00313	0.05957	0.00308
100	101.3880	0.9488	98.5354	0.9468	99.3878	0.9526	69.8182	0.9470
	0.18231	0.00312	0.10894	0.00317	0.10751	0.00301	0.11673	0.00317
200	203.4292	0.9530	196.9604	0.9510	198.2364	0.9492	140.4634	0.9498
	0.36160	0.00299	0.15243	0.00305	0.15005	0.00311	0.24029	0.00309
400	407.4462	0.9494	394.0478	0.9460	395.0836	0.9492	280.4296	0.9538
	0.74799	0.00310	0.21205	0.00320	0.21168	0.00311	0.47231	0.00297
800	812.9954	0.9488	787.6208	0.9472	788.9544	0.9476	561.8558	0.9496
	1.48233	0.00312	0.29854	0.00316	0.30620	0.00315	0.92060	0.00309

	PS		PSR		ES		UPS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	15.1884	0.9506	15.0110	0.9482	10.6690	0.9822	10.8560	0.9580
	0.02226	0.00306	0.02226	0.00313	0.01015	0.00187	0.01325	0.00284
20	20.0958	0.9572	19.9140	0.9542	11.8592	0.9756	14.0668	0.9488
	0.02504	0.00286	0.02507	0.00296	0.01802	0.00218	0.01995	0.00312
25	25.0576	0.9512	24.8796	0.9492	13.5127	0.9738	17.4618	0.9536
	0.02757	0.00305	0.02767	0.00311	0.02552	0.00226	0.02222	0.00298
30	29.9230	0.9516	29.7428	0.9508	15.4671	0.9714	20.8318	0.9474
	0.02953	0.00304	0.02963	0.00306	0.03252	0.00236	0.02350	0.00316
50	49.6596	0.9524	49.4946	0.9524	24.2416	0.9708	34.5154	0.9560
	0.03755	0.00301	0.03770	0.00301	0.05896	0.00238	0.03046	0.00290
100	98.7670	0.9504	98.5916	0.9500	47.1625	0.9660	68.5512	0.9534
	0.05337	0.00307	0.05331	0.00308	0.12056	0.00256	0.04203	0.00298
200	197.3852	0.9516	197.2054	0.9508	93.0877	0.9624	136.8506	0.9514
	0.07495	0.00304	0.07509	0.00306	0.24241	0.00269	0.05875	0.00304
400	394.3956	0.9512	394.2206	0.9506	184.1207	0.9608	273.0638	0.9498
	0.10402	0.00305	0.10385	0.00306	0.47372	0.00275	0.08292	0.00309
800	787.9080	0.9464	787.7244	0.9468	365.8572	0.9568	545.4098	0.9490
	0.14764	0.00319	0.14768	0.00317	0.94331	0.00288	0.11740	0.00311

Table 2.8: Simulation Results for mixture-normal Distribution with $\Delta = 4$

	DS		TS		TSR		UDS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	15.0396	0.9498	15.5150	0.9500	16.6018	0.9612	10.5046	0.9544
	0.02195	0.00309	0.02125	0.00308	0.02073	0.00273	0.01052	0.00295
20	20.1904	0.9480	20.5146	0.9520	21.6242	0.9624	13.7326	0.9416
	0.02902	0.00314	0.02857	0.00302	0.02846	0.00269	0.02004	0.00332
25	25.3562	0.9512	25.5182	0.9544	26.6184	0.9566	17.2740	0.9466
	0.03690	0.00305	0.03480	0.00295	0.03528	0.00288	0.02431	0.00318
30	30.5156	0.9484	30.4444	0.9470	31.6876	0.9570	20.8012	0.9466
	0.04274	0.00313	0.04202	0.00317	0.04218	0.00287	0.02935	0.00318
50	51.2414	0.9504	50.0292	0.9486	51.2224	0.9538	35.0024	0.9490
	0.07230	0.00307	0.06135	0.00312	0.06005	0.00297	0.04939	0.00311
100	103.0512	0.9512	100.2854	0.9478	101.2198	0.9536	70.8156	0.9484
	0.14282	0.00305	0.08014	0.00315	0.08101	0.00298	0.09557	0.00313
200	207.0092	0.9550	200.1274	0.9542	201.4768	0.9576	141.5366	0.9534
	0.29155	0.00293	0.11238	0.00296	0.11140	0.00285	0.19484	0.00298
400	413.8526	0.9542	400.1616	0.9510	401.5578	0.9522	284.2274	0.9544
	0.57800	0.00296	0.15946	0.00305	0.15764	0.00302	0.39690	0.00295
800	828.0466	0.9532	800.1750	0.9512	801.6500	0.9540	569.5322	0.9530
	1.13162	0.00299	0.22046	0.00305	0.22553	0.00296	0.78164	0.00299

	PS		PSR		ES		UPS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	15.4698	0.9524	15.2880	0.9500	10.7234	0.9828	10.9378	0.9560
	0.01696	0.00301	0.01704	0.00308	0.01036	0.00184	0.01228	0.00290
20	20.4664	0.9490	20.2974	0.9480	11.9036	0.9788	14.2362	0.9524
	0.01899	0.00311	0.01909	0.00314	0.01679	0.00204	0.01655	0.00301
25	25.4454	0.9594	25.2746	0.9576	13.6705	0.9720	17.6980	0.9418
	0.02094	0.00279	0.02106	0.00285	0.02405	0.00233	0.01782	0.00331
30	30.4350	0.9492	30.2508	0.9486	15.6417	0.9728	21.1694	0.9518
	0.02292	0.00311	0.02311	0.00312	0.03093	0.00230	0.01954	0.00303
50	50.4222	0.9514	50.2578	0.9514	24.5850	0.9720	34.9990	0.9466
	0.02884	0.00304	0.02883	0.00304	0.05569	0.00233	0.02439	0.00318
100	100.4730	0.9498	100.2974	0.9488	47.8925	0.9624	69.5152	0.9438
	0.03911	0.00309	0.03914	0.00312	0.11205	0.00269	0.03333	0.00326
200	200.4232	0.9498	200.2500	0.9504	94.4451	0.9660	138.5840	0.9508
	0.05595	0.00309	0.05597	0.00307	0.21734	0.00256	0.04708	0.00306
400	400.4350	0.9464	400.2584	0.9468	187.8742	0.9664	276.7572	0.9452
	0.07776	0.00319	0.07782	0.00317	0.44184	0.00255	0.06685	0.00322
800	800.4140	0.9444	800.2402	0.9448	372.2781	0.9600	552.8476	0.9472
	0.10904	0.00324	0.10888	0.00323	0.86274	0.00277	0.09236	0.00316

Table 2.9: Simulation Results for mixture-normal Distribution with $\Delta = 6$

	DS		TS		TSR		UDS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	15.0200	0.9460	15.5030	0.9494	16.6136	0.9634	10.4394	0.9548
	0.01869	0.00320	0.01799	0.00310	0.01758	0.00266	0.00949	0.00294
20	20.1962	0.9480	20.4936	0.9498	21.6266	0.9614	13.7304	0.9432
	0.02488	0.00314	0.02407	0.00309	0.02389	0.00272	0.01784	0.00327
25	25.3644	0.9546	25.5044	0.9556	26.6318	0.9558	17.2764	0.9474
	0.03038	0.00294	0.02942	0.00291	0.02974	0.00291	0.02174	0.00316
30	30.4864	0.9492	30.4554	0.9490	31.6822	0.9578	20.7960	0.9460
	0.03668	0.00311	0.03552	0.00311	0.03554	0.00284	0.02615	0.00320
50	51.1438	0.9512	50.1592	0.9492	51.3248	0.9542	35.0202	0.9492
	0.06120	0.00305	0.05123	0.00311	0.04970	0.00296	0.04414	0.00311
100	103.1460	0.9514	100.3686	0.9466	101.3582	0.9518	70.8074	0.9488
	0.11653	0.00304	0.06544	0.00318	0.06575	0.00303	0.08527	0.00312
200	206.3396	0.9564	200.2512	0.9550	201.5550	0.9590	141.5674	0.9550
	0.24184	0.00289	0.09006	0.00293	0.08959	0.00280	0.17340	0.00293
400	413.0340	0.9542	400.2968	0.9496	401.6666	0.9498	284.2840	0.9560
	0.49152	0.00296	0.12781	0.00309	0.12557	0.00309	0.35539	0.00290
800	829.3414	0.9526	800.3340	0.9510	801.6526	0.9550	569.2280	0.9548
	0.96623	0.00301	0.17547	0.00305	0.17871	0.00293	0.69797	0.00294

	PS		PSR		ES		UPS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	15.4556	0.9528	15.2770	0.9524	10.6908	0.9820	10.8684	0.9530
	0.01430	0.00300	0.01445	0.00301	0.00992	0.00188	0.01105	0.00299
20	20.4530	0.9522	20.2810	0.9520	11.9233	0.9794	14.2660	0.9528
	0.01601	0.00302	0.01601	0.00302	0.01668	0.00201	0.01454	0.00300
25	25.4798	0.9528	25.3016	0.9510	13.6366	0.9800	17.7192	0.9556
	0.01716	0.00300	0.01717	0.00305	0.02311	0.00198	0.01577	0.00291
30	30.4544	0.9546	30.2742	0.9526	15.6298	0.9728	21.1996	0.9510
	0.01826	0.00294	0.01840	0.00301	0.02925	0.00230	0.01697	0.00305
50	50.4944	0.9542	50.3224	0.9544	24.6040	0.9732	35.0016	0.9512
	0.02262	0.00296	0.02268	0.00295	0.05190	0.00228	0.02127	0.00305
100	100.4356	0.9520	100.2538	0.9510	47.9398	0.9690	69.5248	0.9516
	0.03202	0.00302	0.03203	0.00305	0.10325	0.00245	0.02891	0.00304
200	200.5076	0.9474	200.3368	0.9474	94.3115	0.9714	138.5306	0.9526
	0.04475	0.00316	0.04469	0.00316	0.20720	0.00236	0.04073	0.00301
400	400.5458	0.9512	400.3738	0.9514	187.0386	0.9668	276.6980	0.9524
	0.06165	0.00305	0.06149	0.00304	0.40854	0.00253	0.05752	0.00301
800	800.3734	0.9502	800.2028	0.9504	371.2251	0.9638	553.1274	0.9516
	0.08682	0.00308	0.08687	0.00307	0.80890	0.00264	0.07946	0.00304

2.4.2 Asymmetric Distributions Case

Also, we studied some asymmetric distributions. We chose that $k = 10$ and three mixture-normal distributions to represent the asymmetric distributions. The distributions we considered here is similar to those mixture-normal distributions we considered above. The only difference is that the mixing proportions are 0.9 and 0.1 here while they are 0.5 and 0.5 above. The parameters of the asymmetric mixture-normal distributions with the variance $\sigma^2 = 1$ and mean $\mu = 0$ are summarized in following Table 2.10.

Table 2.10: Parameters of Asymmetrical mixture-normal Distribution

	$\Delta = 2$	$\Delta = 4$	$\Delta = 6$
$\pi_1 = 0.9$	$\mu_1 = 0.17150$	$\mu_1 = 0.25607$	$\mu_1 = 0.29139$
	$\mu_2 = -1.54349$	$\mu_2 = -2.30467$	$\mu_2 = -2.62247$
	$\sigma = 0.85749$	$\sigma = 0.64018$	$\sigma = 0.48564$

The distributions of the populations were specified in the same way as described in Section 2.4.1 for our robustness study, except that the distributions for the non-control populations used here are the three asymmetric mixture-normal distributions. Also, samples were generated in the same way as described in Section 2.4.1 to obtain the results displayed in Table 2.11 - 2.13.

In addition, we not only studied the situation that all the non control populations have the same shape (distribution), but also studied when the skewness in these populations differs considerably. For such situation, we assumed the control population still has standard normal distribution. Also, we assumed the non-control populations $X_1, \cdots X_r \sim f(-\delta - x)$ and $X_{r+1}, \cdots X_k \sim f(-\delta + x)$, where the $f(\cdot)$ is the *pdf* of the asymmetric mixture-normal distribution with $\Delta = 6$ and other parameters as specified in Table 2.10. Note that such populations are also under LFC and have common variance $\sigma^2 = 1$. By generating samples from such populations, we obtained the probability of correct decision of the procedures for the most strongly skewed and most badly differed populations, in some degree.

We took the starting sample size to be $m = 10$, $P^* = .95$, $\rho = \frac{1}{4}$, for TS and TSR and $j = 2$

for the ES. For each set of distributions of the populations and each value of δ, each procedure was repeated 5000 times independently and the values of $\bar{n}$, $s(\bar{n})$, $\bar{p}$ and $s(\bar{p})$ are recorded and displayed in tables 2.11-2.14.

Based on the simulations, we conclude that, if the populations have the same shape, then the validity of all the procedures is marginally affected by skewness. Secondly, if the skewness in the populations differs considerably, then the validity of all the procedures, except the ES, is moderately affected. However, the averse effects diminish, with increasing the sample size. Thirdly, if the skewness in the populations differs considerably, then the validity of the ES procedure is marginally affected, due to its wider error margin compared with other procedures. Another difference between ES and the others is that, it performs relatively better for smaller sample sizes.

From the Table 2.14, it is important to note that one of the worst performance is for UDS with $n^* = 20$, giving $\bar{p} = 0.9202$. Note that in this case the $\bar{p}$ value is not within 2 standard errors of the target value.

2.5 Departure from Independence

For studying the performance of the procedures against the departure from independence, we assumed that the samples collected from each non-control population have a serial effect. The samples from each non-control population were generated from normally distributed AR(1) model. The first order correlation coefficients considered were 0.05, 0.1 and 0.2. The other settings of the populations and the generation of the samples are the same as described in the third paragraph in Section 2.4.1.

We took the starting sample size to be $m = 10$, $P^* = .95$, $\rho = \frac{1}{4}$, for TS and TSR and $j = 2$ for the ES. For each population setting and each value of δ, each procedure was repeated 5000 times independently and the values of $\bar{n}$, $s(\bar{n})$, $\bar{p}$ and $s(\bar{p})$ are recorded and displayed in Tables 2.15-2.17.

Table 2.11: Simulation Results for Asymmetric Mixture-Normal Distribution with $\Delta = 2$

	DS		TS		TSR		UDS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	14.9990	0.9456	15.5014	0.9520	16.6664	0.9596	10.6536	0.9564
	0.03297	0.00321	0.03250	0.00302	0.03312	0.00278	0.01385	0.00289
20	20.1948	0.9468	20.4870	0.9524	21.6740	0.9538	13.7402	0.9450
	0.04420	0.00317	0.04277	0.00301	0.04261	0.00297	0.02730	0.00322
25	25.4710	0.9480	25.5704	0.9460	26.6436	0.9538	17.3502	0.9462
	0.05516	0.00314	0.05413	0.00320	0.05415	0.00297	0.03438	0.00319
30	30.4466	0.9460	30.4036	0.9464	31.6656	0.9550	20.7866	0.9440
	0.06540	0.00320	0.06220	0.00319	0.06425	0.00293	0.04058	0.00325
50	51.3060	0.9472	49.6712	0.9416	50.6582	0.9412	35.1318	0.9518
	0.11038	0.00316	0.09394	0.00332	0.09253	0.00333	0.06787	0.00303
100	102.9954	0.9516	99.8588	0.9496	100.7194	0.9480	70.6456	0.9452
	0.21814	0.00304	0.13247	0.00309	0.13487	0.00314	0.13666	0.00322
200	206.2216	0.9486	199.4292	0.9494	200.6042	0.9446	141.8386	0.9486
	0.43955	0.00312	0.18633	0.00310	0.18791	0.00324	0.27459	0.00312
400	414.2140	0.9534	399.7994	0.9474	401.2096	0.9468	284.7490	0.9496
	0.89133	0.00298	0.26507	0.00316	0.26225	0.00317	0.55277	0.00309
800	825.7868	0.9532	799.1056	0.9460	801.0206	0.9488	568.1014	0.9514
	1.77177	0.00299	0.36822	0.00320	0.37410	0.00312	1.09229	0.00304

	PS		PSR		ES		UPS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	15.3318	0.9518	15.1548	0.9488	10.7107	0.9808	11.0082	0.9576
	0.02613	0.00303	0.02633	0.00312	0.01087	0.00194	0.01492	0.00285
20	20.3768	0.9498	20.1946	0.9506	12.0003	0.9782	14.1886	0.9486
	0.03016	0.00309	0.03027	0.00306	0.01982	0.00207	0.02334	0.00312
25	25.3660	0.9500	25.1822	0.9498	13.6673	0.9724	17.6534	0.9524
	0.03360	0.00308	0.03392	0.00309	0.02819	0.00232	0.02611	0.00301
30	30.3474	0.9536	30.1670	0.9512	15.6482	0.9758	21.1108	0.9520
	0.03655	0.00298	0.03651	0.00305	0.03626	0.00217	0.02816	0.00302
50	50.3408	0.9570	50.1584	0.9546	24.7065	0.9684	34.9490	0.9538
	0.04753	0.00287	0.04753	0.00294	0.06615	0.00247	0.03527	0.00297
100	100.2924	0.9522	100.1262	0.9520	48.1225	0.9604	69.5040	0.9492
	0.06409	0.00302	0.06407	0.00302	0.13318	0.00276	0.04956	0.00311
200	200.1708	0.9510	200.0028	0.9508	94.3494	0.9642	138.6014	0.9500
	0.09344	0.00305	0.09339	0.00306	0.26377	0.00263	0.07224	0.00308
400	400.4172	0.9562	400.2394	0.9564	187.4586	0.9558	276.4798	0.9492
	0.12987	0.00289	0.12991	0.00289	0.53050	0.00291	0.09951	0.00311
800	800.2326	0.9544	800.0652	0.9544	372.1724	0.9568	552.8752	0.9514
	0.18781	0.00295	0.18753	0.00295	1.07422	0.00288	0.13900	0.00304

Table 2.12: Simulation Results for Asymmetric Mixture-Normal Distribution with $\Delta = 4$

	DS		TS		TSR		UDS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	15.0584	0.9480	15.5240	0.9524	16.6618	0.9618	10.7760	0.9532
	0.03992	0.00314	0.03900	0.00301	0.03997	0.00271	0.01616	0.00299
20	20.1360	0.9466	20.4466	0.9522	21.6662	0.9520	13.7332	0.9496
	0.05456	0.00318	0.05281	0.00302	0.05263	0.00302	0.03241	0.00309
25	25.4654	0.9446	25.5932	0.9446	26.5956	0.9540	17.3372	0.9426
	0.06837	0.00324	0.06604	0.00324	0.06596	0.00296	0.04150	0.00329
30	30.4174	0.9426	30.3226	0.9424	31.5922	0.9552	20.7656	0.9466
	0.08043	0.00329	0.07545	0.00330	0.07787	0.00293	0.04822	0.00318
50	51.1738	0.9454	49.3676	0.9376	50.2938	0.9434	35.0508	0.9408
	0.13617	0.00321	0.11709	0.00342	0.11340	0.00327	0.08241	0.00334
100	102.8406	0.9432	99.3544	0.9406	100.0428	0.9456	70.5534	0.9428
	0.27067	0.00327	0.16624	0.00334	0.16970	0.00321	0.16459	0.00328
200	206.6810	0.9466	199.0426	0.9462	199.8490	0.9472	142.0886	0.9444
	0.54658	0.00318	0.23671	0.00319	0.23907	0.00316	0.33338	0.00324
400	413.9704	0.9494	399.4036	0.9516	400.9342	0.9478	284.6354	0.9484
	1.09602	0.00310	0.32847	0.00304	0.33439	0.00315	0.66944	0.00313
800	825.2236	0.9480	798.1866	0.9484	800.6584	0.9510	566.9474	0.9436
	2.20728	0.00314	0.46344	0.00313	0.47313	0.00305	1.34515	0.00326

	PS		PSR		ES		UPS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	15.2552	0.9444	15.0760	0.9416	10.7489	0.9792	11.1102	0.9608
	0.03289	0.00324	0.03282	0.00332	0.01153	0.00202	0.01719	0.00274
20	20.2482	0.9488	20.0674	0.9486	12.0293	0.9754	14.1694	0.9562
	0.03838	0.00312	0.03850	0.00312	0.02158	0.00219	0.02755	0.00289
25	25.3118	0.9446	25.1252	0.9448	13.7105	0.9718	17.5474	0.9498
	0.04191	0.00324	0.04210	0.00323	0.03080	0.00234	0.03079	0.00309
30	30.3240	0.9508	30.1460	0.9494	15.6616	0.9694	21.0206	0.9472
	0.04563	0.00306	0.04577	0.00310	0.03989	0.00244	0.03425	0.00316
50	50.3338	0.9528	50.1580	0.9524	24.6200	0.9706	34.9104	0.9512
	0.05814	0.00300	0.05852	0.00301	0.07488	0.00239	0.04328	0.00305
100	100.1100	0.9490	99.9344	0.9482	47.8133	0.9598	69.4076	0.9552
	0.08201	0.00311	0.08217	0.00313	0.15291	0.00278	0.06036	0.00293
200	200.4696	0.9468	200.2942	0.9474	94.8592	0.9568	138.4360	0.9540
	0.11616	0.00317	0.11615	0.00316	0.30665	0.00288	0.08577	0.00296
400	400.3050	0.9484	400.1274	0.9482	186.6472	0.9528	276.6862	0.9516
	0.16205	0.00313	0.16204	0.00313	0.60866	0.00300	0.12207	0.00304
800	800.2752	0.9524	800.1032	0.9520	370.3924	0.9546	552.9212	0.9508
	0.23219	0.00301	0.23227	0.00302	1.20475	0.00294	0.16864	0.00306

Table 2.13: Simulation Results for Asymmetric Mixture-Normal Distribution with $\Delta = 6$

	DS		TS		TSR		UDS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	15.0612	0.9464	15.5554	0.9528	16.6794	0.9572	10.8498	0.9548
	0.04406	0.00319	0.04357	0.00300	0.04486	0.00286	0.01759	0.00294
20	20.1798	0.9410	20.4312	0.9508	21.6518	0.9532	13.8012	0.9444
	0.06204	0.00333	0.05990	0.00306	0.06003	0.00299	0.03532	0.00324
25	25.4598	0.9398	25.6030	0.9422	26.5626	0.9516	17.3532	0.9494
	0.07618	0.00336	0.07431	0.00330	0.07487	0.00304	0.04546	0.00310
30	30.3822	0.9378	30.2814	0.9392	31.5328	0.9542	20.7486	0.9452
	0.09147	0.00342	0.08554	0.00338	0.08747	0.00296	0.05408	0.00322
50	51.1468	0.9406	49.1280	0.9372	50.0296	0.9430	35.0446	0.9452
	0.15369	0.00334	0.13294	0.00343	0.12880	0.00328	0.09155	0.00322
100	102.7134	0.9404	98.7680	0.9404	99.4736	0.9430	70.4878	0.9414
	0.30811	0.00335	0.19546	0.00335	0.19658	0.00328	0.18511	0.00332
200	206.7632	0.9424	198.5944	0.9432	199.2688	0.9502	142.1524	0.9428
	0.61957	0.00330	0.27487	0.00327	0.27766	0.00308	0.37349	0.00328
400	414.3772	0.9496	398.8766	0.9490	400.4740	0.9474	284.8458	0.9474
	1.22546	0.00309	0.37875	0.00311	0.38832	0.00316	0.73291	0.00316
800	823.3690	0.9440	797.4944	0.9482	800.3252	0.9528	566.7178	0.9514
	2.42116	0.00325	0.53727	0.00313	0.54135	0.00300	1.44602	0.00304

	PS		PSR		ES		UPS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	15.3076	0.9554	15.1356	0.9530	10.7759	0.9802	11.1798	0.9538
	0.03666	0.00292	0.03664	0.00299	0.01264	0.00197	0.01868	0.00297
20	20.1930	0.9500	20.0242	0.9478	12.0744	0.9742	14.1206	0.9518
	0.04459	0.00308	0.04476	0.00315	0.02227	0.00224	0.03039	0.00303
25	25.1176	0.9438	24.9322	0.9422	13.8075	0.9716	17.5514	0.9518
	0.04996	0.00326	0.05022	0.00330	0.03325	0.00235	0.03549	0.00303
30	30.2096	0.9484	30.0364	0.9476	15.7642	0.9672	20.9880	0.9526
	0.05166	0.00313	0.05168	0.00315	0.04412	0.00252	0.03976	0.00301
50	50.2706	0.9554	50.0822	0.9554	24.6522	0.9622	34.8400	0.9486
	0.06628	0.00292	0.06647	0.00292	0.08162	0.00270	0.04965	0.00312
100	100.1110	0.9520	99.9446	0.9508	47.9093	0.9600	69.3456	0.9524
	0.09457	0.00302	0.09461	0.00306	0.16595	0.00277	0.06857	0.00301
200	200.2054	0.9528	200.0302	0.9520	94.0554	0.9570	138.2852	0.9540
	0.13082	0.00300	0.13097	0.00302	0.32391	0.00287	0.09545	0.00296
400	399.9724	0.9498	399.7940	0.9494	187.9268	0.9540	276.4626	0.9478
	0.18468	0.00309	0.18483	0.00310	0.66273	0.00296	0.13645	0.00315
800	800.3956	0.9522	800.2188	0.9520	371.4461	0.9496	552.8152	0.9498
	0.25989	0.00302	0.25991	0.00302	1.32183	0.00309	0.19385	0.00309

Table 2.14: Simulation Results for Asymmetric Mixture-Normal Distribution with $\Delta = 6$ and Different Shapes

	DS		TS		TSR		UDS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	15.0660	0.9322	15.5220	0.9380	16.6734	0.9492	10.8728	0.9216
	0.04564	0.00356	0.04484	0.00341	0.04488	0.00311	0.01800	0.00380
20	20.1914	0.9302	20.4892	0.9332	21.6518	0.9430	13.7926	0.9202
	0.06138	0.00360	0.05955	0.00353	0.06003	0.00328	0.03520	0.00383
25	25.3212	0.9314	25.4654	0.9338	26.5626	0.9428	17.2530	0.9172
	0.07686	0.00358	0.07415	0.00352	0.07487	0.00328	0.04607	0.00390
30	30.6708	0.9290	30.5168	0.9292	31.5328	0.9472	20.9182	0.9192
	0.09407	0.00363	0.08786	0.00363	0.08747	0.00316	0.05647	0.00385
50	51.3566	0.9410	49.1114	0.9364	50.0296	0.9380	35.1454	0.9270
	0.15559	0.00333	0.13145	0.00345	0.12880	0.00341	0.09296	0.00368
100	102.9644	0.9360	98.5054	0.9374	99.4736	0.9406	70.7056	0.9308
	0.30669	0.00346	0.19225	0.00343	0.19658	0.00334	0.18381	0.00359
200	205.6000	0.9386	198.7860	0.9446	199.2688	0.9496	141.1410	0.9340
	0.61147	0.00340	0.26723	0.00324	0.27766	0.00309	0.36576	0.00351
400	413.9356	0.9418	398.7570	0.9490	400.4740	0.9450	284.7916	0.9422
	1.24298	0.00331	0.38307	0.00311	0.38832	0.00322	0.74708	0.00330
800	821.9058	0.9400	798.3420	0.9478	800.3252	0.9460	564.9754	0.9400
	2.48996	0.00336	0.54640	0.00315	0.54135	0.00320	1.49379	0.00336

	PS		PSR		ES		UPS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	15.1904	0.9364	15.0148	0.9338	10.8622	0.9702	11.1686	0.9316
	0.03669	0.00345	0.03666	0.00352	0.01430	0.00240	0.01837	0.00357
20	20.1610	0.9396	19.9682	0.9386	12.0838	0.9636	14.1128	0.9246
	0.04422	0.00337	0.04447	0.00340	0.02545	0.00265	0.03029	0.00373
25	25.2198	0.9358	25.0288	0.9346	13.7642	0.9620	17.5412	0.9224
	0.04840	0.00347	0.04859	0.00350	0.03801	0.00270	0.03557	0.00378
30	30.2824	0.9420	30.1016	0.9406	15.7479	0.9576	20.9560	0.9334
	0.05198	0.00331	0.05210	0.00334	0.05050	0.00285	0.03895	0.00353
50	50.2814	0.9444	50.1074	0.9438	24.3950	0.9546	34.6928	0.9374
	0.06648	0.00324	0.06651	0.00326	0.09163	0.00294	0.04910	0.00343
100	100.0222	0.9454	99.8388	0.9442	47.7278	0.9490	69.4418	0.9380
	0.09311	0.00321	0.09335	0.00325	0.18546	0.00311	0.06813	0.00341
200	200.4214	0.9390	200.2434	0.9396	94.1081	0.9498	138.3922	0.9440
	0.13275	0.00338	0.13271	0.00337	0.35916	0.00309	0.09704	0.00325
400	400.2520	0.9454	400.0878	0.9458	187.1722	0.9462	276.6716	0.9486
	0.18486	0.00321	0.18494	0.00320	0.70431	0.00319	0.13570	0.00312
800	800.1906	0.9520	800.0170	0.9518	372.9378	0.9442	552.7680	0.9458
	0.26542	0.00302	0.26560	0.00303	1.40287	0.00325	0.18876	0.00320

The seriousness of consequences depends on the magnitude of the autocorrelation. In Table 2.15, where the correlation coefficient is only 0.05, the effect is moderate, In Table 2.17, where the correlation coefficient is 0.2, the effect becomes obvious, with $\bar{p}$ values ranging from .85 to .90.

Note that The UDS, the UPS and the ES performance better for smaller sample size than for larger one; the PS and the PSR are on the contrary; and the performance the DS, the TS and the TSR are almost not related with the sample size. For the ES, such performance pattern is due to its error margin, which is larger while the sample size are small; for UDS and UPS, the extra information from the control population aggravate the adverse; for the DS, the TS and the TSR, they only use the standard error statistics 2 or 3 times, so the sample size will not affect them as much as it do the PS and the PSR.

The ES performance better than all the other procedures, however, it is still not good enough to be called robust when the independence is violated.

2.6 Discussion of Results and Conclusions

1. The sequential procedure with elimination is the best one for the partitioning problem, from the point of view of robustness and smaller total sample size.

2. If the sequential sampling is too inconvenient to carry out, the fine tuned three-stage procedure is the best one among the procedures we studied here.

3. If the price for sampling from control is not significant compared with the price for sampling from non-control populations, then the total cost of sampling by using unbalanced purely sequential procedure will be nearly the same as the sequential procedure with elimination.

4. If the sampling implementation is an issue, which means that two-stage procedures are the only feasible procedures in this case, then the unbalanced two-stage procedure are better than two-stage procedure in consideration of both sample size and robustness. However, when the independence is suspect, then the two-stage procedure is better. This is because, the larger the sample size the smaller the effect of auto correlation. The unbalanced two-stage procedure has

Table 2.15: Simulation Result for data with correlations, $AR(0.05)$

	DS		TS		TSR		UDS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	14.8198	0.9314	15.3132	0.9358	16.5218	0.9554	10.5722	0.9392
	0.03098	0.00358	0.03017	0.00347	0.03000	0.00292	0.01279	0.00338
20	20.0358	0.9304	20.3440	0.9360	21.4014	0.9484	13.6556	0.9272
	0.04108	0.00360	0.03984	0.00346	0.03981	0.00313	0.02571	0.00367
25	25.1168	0.9328	25.2744	0.9340	26.4844	0.9484	17.1556	0.9234
	0.05230	0.00354	0.05052	0.00351	0.05005	0.00313	0.03320	0.00376
30	30.1286	0.9326	30.1014	0.9332	31.3056	0.9480	20.6242	0.9294
	0.06142	0.00355	0.05887	0.00353	0.05992	0.00314	0.03832	0.00362
50	50.7210	0.9354	49.3342	0.9304	50.5170	0.9384	34.7970	0.9314
	0.10321	0.00348	0.08666	0.00360	0.08851	0.00340	0.06495	0.00358
100	101.9740	0.9364	99.3924	0.9340	100.4540	0.9420	70.2458	0.9360
	0.20752	0.00345	0.12217	0.00351	0.12580	0.00331	0.12962	0.00346
200	204.1586	0.9330	199.2846	0.9358	200.6922	0.9372	140.4372	0.9314
	0.41283	0.00354	0.17528	0.00347	0.17254	0.00343	0.25756	0.00358
400	409.4112	0.9362	399.4694	0.9390	400.3716	0.9406	281.9422	0.9316
	0.82977	0.00346	0.24967	0.00338	0.25094	0.00334	0.51634	0.00357
800	816.6754	0.9360	799.0676	0.9364	800.6816	0.9378	563.0594	0.9336
	1.64991	0.00346	0.34583	0.00345	0.34527	0.00342	1.04812	0.00352

	PS		PSR		ES		UPS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	15.2638	0.9420	15.0870	0.9394	10.7562	0.9738	10.9192	0.9352
	0.02463	0.00331	0.02494	0.00337	0.01100	0.00226	0.01422	0.00348
20	20.2532	0.9402	20.0724	0.9368	11.9661	0.9716	14.1316	0.9376
	0.02856	0.00335	0.02871	0.00344	0.01862	0.00235	0.02208	0.00342
25	25.3292	0.9368	25.1548	0.9360	13.6264	0.9640	17.5344	0.9382
	0.03165	0.00344	0.03161	0.00346	0.02739	0.00263	0.02488	0.00341
30	30.3080	0.9414	30.1330	0.9404	15.5562	0.9614	21.0410	0.9432
	0.03452	0.00332	0.03455	0.00335	0.03447	0.00272	0.02657	0.00327
50	50.2394	0.9404	50.0670	0.9392	24.4834	0.9584	34.8788	0.9448
	0.04432	0.00335	0.04440	0.00338	0.06436	0.00282	0.03375	0.00323
100	100.3396	0.9466	100.1616	0.9472	47.5340	0.9546	69.3912	0.9340
	0.06190	0.00318	0.06189	0.00316	0.13077	0.00294	0.04774	0.00351
200	200.2198	0.9434	200.0372	0.9428	93.3788	0.9524	138.3312	0.9366
	0.08699	0.00327	0.08693	0.00328	0.25559	0.00301	0.06611	0.00345
400	400.5484	0.9390	400.3710	0.9388	184.8790	0.9492	276.3754	0.9348
	0.12253	0.00338	0.12261	0.00339	0.51285	0.00311	0.09201	0.00349
800	800.0738	0.9410	799.8984	0.9410	368.1539	0.9486	552.8460	0.9364
	0.17051	0.00333	0.17048	0.00333	1.01666	0.00312	0.13134	0.00345

Table 2.16: Simulation Result for data with correlations, $AR(0.1)$

	DS		TS		TSR		UDS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	14.6634	0.9140	15.2146	0.9226	16.3924	0.9432	10.5412	0.9226
	0.03089	0.00397	0.03010	0.00378	0.03033	0.00327	0.01236	0.00378
20	19.8144	0.9164	20.0518	0.9158	21.2630	0.9350	13.4914	0.9062
	0.04083	0.00391	0.03965	0.00393	0.03977	0.00349	0.02583	0.00412
25	24.8466	0.9170	25.0486	0.9196	26.1660	0.9318	17.0264	0.9072
	0.05198	0.00390	0.04999	0.00385	0.04892	0.00357	0.03239	0.00410
30	29.8106	0.9176	29.8234	0.9170	30.9914	0.9322	20.4668	0.9044
	0.06094	0.00389	0.06001	0.00390	0.05830	0.00356	0.03923	0.00416
50	50.1878	0.9214	48.8672	0.9196	49.9824	0.9244	34.4576	0.9130
	0.10261	0.00381	0.08830	0.00385	0.08852	0.00374	0.06433	0.00399
100	100.9106	0.9200	98.9784	0.9166	99.9582	0.9256	69.5706	0.9088
	0.20618	0.00384	0.12622	0.00391	0.12358	0.00371	0.12857	0.00407
200	202.0468	0.9166	198.8186	0.9214	200.1706	0.9308	139.1198	0.9134
	0.41137	0.00391	0.17816	0.00381	0.17946	0.00359	0.26094	0.00398
400	405.1124	0.9228	399.0918	0.9272	399.8942	0.9336	278.8278	0.9094
	0.82529	0.00378	0.25283	0.00367	0.25059	0.00352	0.50009	0.00406
800	807.8902	0.9220	798.6962	0.9300	800.7510	0.9304	559.8158	0.9136
	1.63932	0.00379	0.35452	0.00361	0.35583	0.00360	1.05497	0.00397

	PS		PSR		ES		UPS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	15.1562	0.9272	14.9702	0.9240	10.7755	0.9676	10.8878	0.9338
	0.02531	0.00367	0.02541	0.00375	0.01094	0.00250	0.01407	0.00352
20	20.2304	0.9298	20.0514	0.9278	11.9544	0.9602	14.0354	0.9250
	0.02829	0.00361	0.02838	0.00366	0.01885	0.00276	0.02223	0.00373
25	25.1954	0.9256	25.0186	0.9244	13.6195	0.9554	17.4954	0.9138
	0.03141	0.00371	0.03160	0.00374	0.02711	0.00292	0.02426	0.00397
30	30.1668	0.9280	29.9802	0.9258	15.4877	0.9530	20.9708	0.9174
	0.03405	0.00366	0.03420	0.00371	0.03467	0.00299	0.02661	0.00389
50	50.1568	0.9274	49.9826	0.9268	24.1905	0.9418	34.7198	0.9144
	0.04464	0.00367	0.04477	0.00368	0.06443	0.00331	0.03360	0.00396
100	100.2350	0.9298	100.0678	0.9286	47.0807	0.9358	69.1912	0.9140
	0.06011	0.00361	0.06029	0.00364	0.13201	0.00347	0.04737	0.00397
200	200.0638	0.9274	199.8910	0.9264	92.5431	0.9342	138.3886	0.9114
	0.08706	0.00367	0.08713	0.00369	0.25751	0.00351	0.06536	0.00402
400	400.1608	0.9274	399.9840	0.9270	183.6118	0.9358	276.5896	0.9098
	0.12041	0.00367	0.12041	0.00368	0.51031	0.00347	0.09363	0.00405
800	800.3764	0.9296	800.1906	0.9288	361.8973	0.9216	552.7104	0.9190
	0.17419	0.00362	0.17410	0.00364	1.03628	0.00380	0.13133	0.00386

Table 2.17: Simulation Result for data with correlations, $AR(0.2)$

	DS		TS		TSR		UDS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	14.3732	0.8862	14.8766	0.8986	16.0040	0.9144	10.4436	0.8838
	0.03107	0.00449	0.03015	0.00427	0.03003	0.00396	0.01134	0.00453
20	19.2744	0.8812	19.6246	0.8904	20.7254	0.8982	13.2364	0.8584
	0.04174	0.00458	0.04034	0.00442	0.03951	0.00428	0.02600	0.00493
25	24.1738	0.8792	24.3620	0.8844	25.4800	0.8922	16.6260	0.8578
	0.05148	0.00461	0.04967	0.00452	0.04908	0.00439	0.03260	0.00494
30	29.1298	0.8746	29.1376	0.8756	30.3090	0.8958	20.0468	0.8582
	0.06184	0.00468	0.05959	0.00467	0.05871	0.00432	0.03899	0.00493
50	48.9138	0.8888	47.8716	0.8868	49.0262	0.8904	33.7912	0.8572
	0.10239	0.00445	0.09076	0.00448	0.09073	0.00442	0.06425	0.00495
100	98.4660	0.8800	97.6806	0.8828	98.9256	0.8966	68.1388	0.8514
	0.20871	0.00460	0.13070	0.00455	0.13224	0.00431	0.13055	0.00503
200	196.8512	0.8846	197.6866	0.8882	198.9316	0.9006	136.3426	0.8590
	0.41298	0.00452	0.18153	0.00446	0.18241	0.00423	0.25884	0.00492
400	394.6762	0.8806	398.1012	0.8918	399.0012	0.8986	273.3500	0.8478
	0.82346	0.00459	0.26102	0.00439	0.25775	0.00427	0.51564	0.00508
800	790.5134	0.8852	797.5044	0.8934	799.2918	0.8966	548.1428	0.8494
	1.62006	0.00451	0.36747	0.00436	0.36032	0.00431	1.02630	0.00506

	PS		PSR		ES		UPS	
n^*	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$	$\bar{n}$	$\bar{p}$
	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$	$s(\bar{n})$	$s(\bar{p})$
15	14.9384	0.9018	14.7520	0.8972	10.7779	0.9538	10.7708	0.8882
	0.02547	0.00421	0.02571	0.00430	0.01086	0.00297	0.01361	0.00446
20	19.8586	0.9004	19.6830	0.8972	11.9673	0.9322	13.7884	0.8694
	0.03002	0.00424	0.03015	0.00430	0.01910	0.00356	0.02277	0.00477
25	24.9748	0.8938	24.7942	0.8892	13.5369	0.9204	17.2404	0.8614
	0.03262	0.00436	0.03276	0.00444	0.02681	0.00383	0.02511	0.00489
30	29.9430	0.8980	29.7612	0.8958	15.3719	0.9176	20.7150	0.8660
	0.03599	0.00428	0.03609	0.00432	0.03452	0.00389	0.02751	0.00482
50	49.9086	0.8916	49.7268	0.8920	23.7007	0.9090	34.5172	0.8680
	0.04559	0.00440	0.04566	0.00439	0.06414	0.00407	0.03408	0.00479
100	99.9410	0.8980	99.7646	0.8976	45.8801	0.8976	69.0146	0.8708
	0.06417	0.00428	0.06423	0.00429	0.13036	0.00429	0.04824	0.00474
200	199.8624	0.9016	199.6814	0.9014	89.8575	0.8864	138.0524	0.8660
	0.08977	0.00421	0.08972	0.00422	0.26616	0.00449	0.06839	0.00482
400	400.1784	0.9030	400.0030	0.9022	177.6256	0.8840	276.1792	0.8724
	0.12652	0.00419	0.12655	0.00420	0.52476	0.00453	0.09502	0.00472
800	799.7134	0.8972	799.5372	0.8972	350.7136	0.8792	552.6698	0.8730
	0.17614	0.00430	0.17612	0.00430	1.05054	0.00461	0.13354	0.00471

relatively smaller sample size on non-control populations, where the auto correlation exists, hence, it will performs worse than two-stage procedure, while the independence is violated.

5. Based on the conclusions discussed above, I believe that the unbalanced sequential procedure with elimination may be the most valuable procedures to be proposed. it will have many advantages such as:

    - robustness property along the lines of the sequential procedure with elimination procedure.
    - smaller total cost of sampling.

Chapter 3

Optimal Choice of c

3.1 Introduction

This chapter is devoted to the optimal choice of c for the unbalanced purely sequential procedure discussed earlier. The "optimal choice of c" means choosing c such that the expectation of total cost for partitioning treatment populations is minimized. Suppose, the cost for collecting a single sample form each population is known, denoted as $p_0, p_1, \cdots, p_k$, corresponding to $\pi_0, \pi_1, \cdots, \pi_k$. Then the expected total cost for the unbalanced purely sequential procedure, which collects c observations from π_0, and one from $\pi_1, \cdots, \pi_k$, at each step, is

$$C = E[N \cdot c \cdot p_0 + N \cdot p_1 + N \cdot p_2 + \cdots + N \cdot p_k]. \tag{3.1.1}$$

This equals

$$E[N][c \cdot p_0 + \sum_{i=1}^{k} p_i]$$

$$= E[N][c + k \cdot (\sum_{i=1}^{k} p_i)/(k \cdot p_0)] \cdot p_0.$$

Since p_0 is a known constant under our setting, it does not affect the choice of c. We define the average relative price p as

$$p = \frac{\sum_{i=1}^{k} p_i}{k \cdot p_0}. \tag{3.1.2}$$

Now, (3.1.1) can be simplified as

$$C = E[N] \cdot [c + kp] \tag{3.1.3}$$

Using Theorem 1.3.1, $E[N] = n_c^* + (\nu^* - 2)(k+c)^{-1} + o(1)$. So

$$\begin{aligned} C &= [n_c^* + (\nu^* - 2)(k+c)^{-1} + o(1)] \cdot (c + kp) \\ &\simeq (n_c^* + (\nu^* - 2)(k+c)^{-1}) \cdot (c + kp) \\ &= [\tfrac{b^2\sigma^2}{a^2} + \tfrac{(\nu^2-2)}{k+c}](c+kp) \\ &= \tfrac{\sigma^2}{a^2}[b^2\tfrac{c+1}{c} + \tfrac{a^2}{\sigma^2}\tfrac{\nu^2-2}{k+c}](c+kp). \end{aligned}$$

Note that for given populations and given indifference-zone, $\frac{\sigma^2}{a^2}$ is constant although unknown.

Theoretically, the value of σ^2/a^2 affects the optimal choice of c, unless C can be expressed as $\frac{\sigma^2}{a^2} \cdot f(c,k,p)$, where $f(\cdot)$ is a function that does not depend on σ or a. Hence, it seems that we have to fix $\frac{\sigma^2}{a^2}$ on some values and to get the optimal choice for c under these different values. Fortunately, in practice, we do not have to do so. We only assume that $\frac{\sigma^2}{a^2}$ is very large and so is n_c^*, such that we can ignore $(\nu^* - 2)/(k+c)$ compared with n_c^*. In practice, ignoring $(\nu^* - 2)/(k+c)$ is reasonable. Intuitively, it is more important to minimize the total cost when the required sample sizes get large. That is, when we do not know how large the sample will turn out to be , it is advisable to find an optimal choice of c under the condition that the sample size is large. Also, $-0.25 \leq \frac{\nu^*-2}{k+c} \leq 0.5$ for all possible combination of k and c. Therefore, under the condition that n_c^* is large, ignoring the term $\frac{\nu^*-2}{k+c}$ does not affect our optimal choice for c much. Under these definition and assumption, the goal is:

$$\text{minimize } C = \frac{\sigma^2}{a^2}(b^2\frac{c+1}{c})(c+kp).$$

That is, minimize:

$$C^* = b^2\frac{c+1}{c}(c+kp) \qquad for\, given\, values\, of\;\; k, p. \tag{3.1.4}$$

3.2 Critical Values for choosing c

Now, we consider the problem with in the range that $0 < p \leq 10$. If necessary, 10 could be as large as 100. if larger, the numerical method may not give the reliable result. Suppose c_0 is the optimal choice for c for given k and p, then the inequality,

$$b_{k,\,c_0}^2 \frac{c_0+1}{c_0}(c_0+kp) \leq b_{k,\,c_0+1}^2 \frac{c_0+2}{c_0+1}(c_0+1+kp), \tag{3.2.5}$$

holds. Fix c_0 and k, the value of b will be fixed. Now the inequality can be solved for p, denote the result as $p \leq p_{\,k,\,c_0}$. Do such calculations for k from 1 to 10 and some c's that could be optimal choices for p in (0,10]. Organize the results in the table below:

It is easy to show that for given k and p, if $p_{c_0-1} < p \leq p_{c_0}$, then c_0 is the optimal choice of c for such given k and p, as long as p_{c_0} is monotone with c_0, which is the case here. From (5), when $p > p_{\,k,\,c_0-1}$, $C_{\,k,\,c_0} < C_{k,c_0-1}$. And because p_{c_0} is increasing with c_0, $p > p_{\,k,\,c'}$ for all $0 < c' < c_0 - 1$. Hence, $C_{k,\,c_0} < C_{k,\,c}$ for all $0 < c < c_0$. Similarly, $C_{k,\,c_0} < C_{k,\,c_0+1}$ because $p \leq p_{c_0}$ and $C_{k,\,c_0} < C_{k,\,c}$ for all $c > c_0$ because P_{c_0} is increasing with c_0.

	c								
k	1	2	3	4	5	6	7	8	9
1	2.00	6.00							
2	1.00	2.98	5.94	9.87					
3		2.07	4.10	6.79					
4		1.57	3.10	5.12	7.64				
5		1.28	2.51	4.15	6.18	8.60			
6		1.08	2.11	3.47	5.16	7.19	9.54		
7			1.82	3.00	4.45	6.20	8.22		
8			1.60	2.63	3.91	5.43	7.21	9.23	
9			1.43	2.35	3.49	4.85	6.43	8.23	
10			1.29	2.12	3.15	4.37	5.80	7.42	9.24

Table 3.1: Critical value for optimal choice of c

In Table 3.1, the entry corresponding to c is the critical average relative price of the treatment populations, at which, the expectation of the total cost for the purely sequential procedures with c and $c+1$ observations from π_0 and one from $\pi_1, \cdots, \pi_k$ at each step, are the same. Under these conditions, if the tabulated entry $p_{\,k,\,c}$ is the first entry that greater or equal to the given average relative price of p in the line that corresponds to k, then the corresponding c is the optimal choice for the unbalanced purely sequential procedure.

3.3 Examples for Choosing the Optimal c

Now, we consider some examples to illustrate the possible use of the Table 3.1. Note that, for k=3 p=5, the optimal choice of c is 4, and for k=7 p=9, the optimal choice of c is 9.

Also when k=5, c=2, and σ^2/a^2, which equal to λ=6, $p_{k,c}$ is the solution of the equation below:

$$[\frac{4.33199-2}{7}+2.30633^2\frac{3}{2}](2+5p)=[\frac{4.86115-2}{8}+2.31247^2\frac{4}{3}](3+5p)$$

The parameters in this equation are obtained from the tables in Chapter 1 and the solution of this equation is 1.28, the same as the corresponding in Table 3.1. Note that, for larger k, the σ^2/a^2 could be even much smaller than 36 and still $p_{k,c}$ be the same as the obtained from the Table 3.1

In order to obtain a more precise optimal c, the term $\frac{\nu^*-2}{k+c}$ should be considered. Also, note that if the prices for collecting samples form the non-control populations differ from each other, then the optimal choice given here is still quite reasonable when the unbalanced purely sequential procedure are used.

Bibliography

[1] Aoshima, M. and Takada, Y. Second order properties of a two stage procedure for comparing several treatments with a control. J. Japan Statist. Soc. 2000, 30(1), 27-41.

[2] Bechhofer, R. E. A single-sample multiple decision procedure for ranking means of normal populations with known variances. Ann. Math. Statist. 1954, 25, 16-39.

[3] Chow, Y. S. and Robbins, H. On the asymptotic theory of fixed width sequential confidence intervals for the mean. Ann. Math. Statist. 1965, 36, 457-462.

[4] Datta, S. and Mukhopadhyay, N. Second-order asymptotics for multistage methodologies in partitioning a set of normal populations having a common unknown variance. Statist. Decisions 1998, 16, No. 2, 1991-205.

[5] Ghosh, M., Mukhopadhyay, N. and Sen, P. K. Sequential Estimation; John Wiley: New York, 1997.

[6] Ghosh, B. K. and Sen, P. K. Handbook of Sequential Analysis; Marcel Deker, Inc., New York, 1991.

[7] Gibbons, J. D., Olkin, I. and Sobel, M. Selecting and ordering populations: A new statistical methodology; John Wiley: New York, 1977. Republished by SIAM: Philadelphia, 1999.

[8] Mahalanobis, P. C. A sample survey of acreage under jute in Bengal, with discussion on planning of experiments. Proc. 2nd Indian Statist. Conf. 1940, statistical Publishing society, Calcutta.

[9] Mukhopadhyay, N. and Duggan, W. T. Can a two-stage procedure enjoy second-order properties? Sankhya, Ser. A 1997, 59, 435-448.

[10] Mukhopadhyay, N. and Duggan, W. T. On a two-stage procedure having second-order properties with applications. Ann. Inst. Statist. Math. 1999, 51, 621-636.

[11] Mukhopadhyay, N. and Solanky, T. K. S. Multistage selection and ranking procedures; Marcel Dekker: New York, 1994.

[12] Mukhopadhyay, N. and Solanky, T. K. S. Multistage methodologies for comparing several treatments with a control. J. Statist. Plann. Inf. 2002, 100(2), 209-220.

[13] Solanky, T. K. S. A sequential procedure with elimination for partitioning a set of normal populations having a common unknown variance. Seq. Anal. 2001, 20(4), 279-292.

[14] Sonlanky, T. S. K. and Wu, Y. On Unbalanced Multistage Methodologies for Partition Problem, Proceedings of the International Sri Lankan Statistical Conference: Visions of Futuristic Methodologies, Dec. 2004

[15] Solanky, T. K. S. A Two-stage procedure with elimination for partitioning a set of normal populations with respect to a control, submitted to Seq. Anal., 2005.

[16] Srivastava, M. S. Some asymptotically efficient sequential procedures for ranking and slippage problems. J. roy. Statist. Soc., B 1996, 28, 370-380.

[17] Starr, N. The performance of a sequential procedure for the fixed-width interval estimation of the mean. Ann. Math. Statist. 1966, 37, 36-50.

[18] Stein, C. A two sample test for a liear hypothesis whose power is independent of the variance. Ann. Math. Statist. 1945, 16, 243-258.

[19] Stein, C. Some problems in sequential estimation (abstract). Econometrica 1949, 17, 77-78.

[20] Tong, Y. L. On partitioning a set of normal populations by their locations with respect to a control. Ann. Math. Statist. 1969, 40,4, 1300-1324.

[21] Woodroofe, M. Second order approximations for sequential point and interval estimation. Ann. Statist. 1977, 5, 984-995.

Vita

Yuefeng was born in Nantong Jiangsu, the People's Republic of China, on January 22, 1978. He attended elementary schools in the Chengzhong Xiaoxue and graduated from Jiangsu Sheng Nantong Middle School with honors in July 1996. The following September he entered Nanjing University and in July 2000 received the degree of Bachelor of Science in Mathematics. In June 2001, he attended the Florida State University and in August 2003 received the degree of Master of Science in Applied Mathematics. He entered the University of New Orleans in August 2003 and is a candidate for the Master of Science Degree in Mathematics.

Permanent Address: 2000 Lakeshore Dr. 323#
New Orleans, LA 70148

This thesis was typeset with LaTeX 2ε[1] by the author.

[1] LaTeX 2ε is an extension of LaTeX. LaTeX is a collection of macros for TeX. TeX is a trademark of the American Mathematical Society. The macros used in formatting this thesis were originally written by Dinesh Das, Department of Computer Sciences, The University of Texas at Austin, and extended by Bert Kay, James A. Bednar, and Ayman El-Khashab. The author of this thesis modified it to satisfy the requirement of University of New Orleans.