Summer 8-6-2018

Generalizing Multistage Partition Procedures for Two-parameter Exponential Populations

Rui Wang
University of New Orleans, rwang5@uno.edu

Generalizing Multistage Partition Procedures for Two-parameter Exponential Populations

A Dissertation

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

Doctor of Philosophy
in
Engineering and Applied Science
Mathematics

by

Rui Wang

M.S. University of New Orleans, 2014
B.S. Huazhong University of Science and Technology, 2010

August, 2018

Acknowledgments

I would like to give my sincere gratitude toward Professor Tumulesh Solanky, who generously spent considerable time in the past few years on discussing lots of details presented in this dissertation. Professor Solanky also read through the earlier draft of this dissertation with great patience and presented many valuable suggestions, without which, I believe, there is still a long way for it to take the present form. I would like to thank other four committee members,

Professor Linxiong Li Department of Mathematics,

Professor Jairo Santanilla Department of Mathematics,

Professor Vesselin P Jilkov Department of Electrical Engineering,

Professor Huimin Chen Department of Electrical Engineering,

for their advice and suggestions where contributed in several aspects to this dissertation. Also I would like to thank Department of Mathematics for support of this research.

Contents

List of Tables

Abstract

ANOVA analysis is a classic tool for multiple comparisons and has been widely used in numerous disciplines due to its simplicity and convenience. The ANOVA procedure is designed to test if a number of different populations are *all* different. This is followed by usual multiple comparison tests to rank the populations. However, the probability of selecting the best population via ANOVA procedure does not guarantee the probability to be larger than some desired prespecified level. This lack of desirability of the ANOVA procedure was overcome by researchers in early 1950's by designing experiments with the goal of selecting the best population. In this dissertation, a single-stage procedure is introduced to partition k treatments into "good" and "bad" groups with respect to a control population assuming some key parameters are known. Next, the proposed partition procedure is genaralized for the case when the parameters are unknown and a purely-sequential procedure and a two-stage procedure are derived. Theoretical asymptotic properties, such as first order and second order properties, of the proposed procedures are derived to document the efficiency of the proposed procedures. These theoretical properties are studied via Monte Carlo simulations to document the performance of the procedures for small and moderate sample sizes.

Key words: Two-parameter Exponential Distribution; Sequential Procedure; Probability of Correct Decision; Indifference Zone; Monte Carlo Simulation

1

Introduction

1.1 Introduction of Exponential Distribution

How long will a piece of machinery work without breaking down? How much time will elapse before an earthquake occurs in a given region? How long do we need to wait before a customer enters a shop? How long will it take before a call center receives the next phone call? All these questions above concern the time we need to wait before a given event occurs. If this waiting time is unknown, it is often appropriate to think it of as a random variable having an exponential distribution. The time we need to wait before a certain event occurs follows exponential distribution if the probability that the event occurs during a certain time interval is proportional of the length of that time interval. The probability density function (PDF) of an exponential distribution is

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0, \\ 0 & x < 0. \end{cases} \tag{1.1}$$

The parameter λ is called rate parameter. It is the inverse of the expected duration μ. A random variable X has this distribution can be expressed as $X \sim Exp(\lambda)$. The

cumulative distribution function (CDF) of an exponential distribution is given as

$$F(x;\lambda) \quad = \quad \begin{cases} 1-\lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0. \end{cases} \tag{1.2}$$

Alternatively, this can be defined using the Heaviside step function $H(x)$ as

$$F(x;\lambda) = (1-e^{-\lambda x})H(x). \tag{1.3}$$

The cumulative density function (CDF) can be written as the probability of the lifetime being less than some value x, which is given as

$$P(X \leq x) = 1 - e^{-\lambda x}. \tag{1.4}$$

A commonly used alternative parametrization is to define the probability density function (pdf) of an exponential distribution as

$$f(x;\beta) = \begin{cases} \frac{1}{\beta}e^{-\frac{x}{\beta}} & x \geq 0, \\ 0 & x < 0. \end{cases} \tag{1.5}$$

where $\beta > 0$ is mean, standard deviation, and is also known as the scale parameter of the distribution, the reciprocal of the rate parameter, λ, defined above. In this specification, β is a survival parameter in the sense that if a random variable X is the duration of time that a given biological or mechanical system manages to survive and $X \sim Exp(\beta)$ then $E[X] = \beta$. That is to say, the expected duration of survival of the system is β units of time. The parametrization involving the "rate" parameter arises in the context of events arriving at a rate λ, when the time between events (which might be modeled using an exponential distribution) has a mean of $\beta = \lambda^{-1}$.

The alternative specification is sometimes more convenient than the one given above, and some authors will use it as a standard definition. This alternative specification is not used in this chapter.

The expected value of an exponential variable X is :

$$E[X] = \frac{1}{\lambda}.$$

The variance of an exponential variable X is :

$$Var[X] = \frac{1}{\lambda^2},$$

so one properties of the exponential distribution is that the standard deviation and the mean of the distribution are equal. The moments of X, for n = 1, 2, ..., are given by

$$E\left[X^n\right] = \frac{n!}{\lambda^n}.$$

The median of X is given by

$$m[X] = \frac{\ln(2)}{\lambda} < E[X],$$

where ln refers to the natural logarithm. Thus the absolute difference between the mean and median is

$$|E[X] - m[X]| = \frac{1 - \ln(2)}{\lambda} < \frac{1}{\lambda} = \text{standard deviation},$$

in accordance with the median-mean inequality.

An exponentially distributed random variable T obeys the relation

$$P\left(T > s + t | T > s\right) = P(T > t), \qquad \forall s, t \geq 0.$$

When T is interpreted as the waiting time for an event to occur relative to some initial time, this relation implies that, if T is conditioned on a failure to observe the event over some initial period of time s, the distribution of the remaining waiting time is the same as the original unconditional distribution. For example, if an machine has not broken after 10 years, the conditional probability that fail will take place after at least 10 more years is equal to the unconditional probability of observing the fail after more than 10 years relative to the initial time.

The exponential distribution and the geometric distribution are the only memoryless probability distributions. The exponential distribution is consequently also necessarily the only continuous probability distribution that has a constant failure rate. That's why the exponential distribution the very commonly used in reliability engineering.

The mathematics associated with the exponential distribution is often of a simple nature, and so it is possible to obtain explicit formulas in terms of elementary functions, without to obtain troublesome quadrature. For this reason models constructed from exponential variables are sometimes used as an approximate representation of other models that are more appropriate for a particular application. Currently among the most prominent applications are in the field of life-testing. The lifetime(or life characteristic, as it is often called) can be usefully represented by an exponential random variable, with a relatively simple associated theory. Sometimes the representation is not adequate; in such cases a modification of the exponential distribution (often a Weibull distribution) is used.

Another application is producing usable approximate solutions to difficult distributional problems. An ingenious application of the exponential distribution to approximate

a sequential procedure is due to Ray (1957). He wished to calculate the distribution of the smallest n for which $\sum_{i=1}^{n} U_i^2 < K_n$, where $U_1, U_2, \ldots$ are independent unit normal variables and $K_1, K_2, \ldots$ are specified positive constants. By replacing this by the distribution of the smallest even n, he obtained a problem in which the sums $\sum_{i=1}^{n} U_i^2$ are replaced by sums of independent exponential variables(actually $\chi^{2,}s$ with two degrees of freedom each).

1.2 Parameter Estimation

Before 1959 a considerable amount of work had been done on inference procedures for the exponential distribution with both censored and uncensored data. It was realized, in the 1960s and 1970s, that although the exponential distribution can be handled rather easily, the consequent analysis is often poorly robust [see Zelen and Dannemiller (1961)]. Nevertheless, the study of properties of this distribution, and especially construction of estimation and testing procedures has continued steadily, during the last 30 years, with some emphasis on Bayesian analysis and order statistics methodology, and an explosion and results on characterizations.

Maximum likelihood estimation is one of the most useful technique which derives estimates of the unknown parameters by maximizing a likelihood function constructed through the available data. Suppose, a given variable X is exponentially distributed and the rate parameter λ is to be estimated. Then, the likelihood function, given by an independent and identically distributed sample $X = (x_1, \cdots, x_n)$, is given by

$$\begin{aligned} L(\lambda) &= \prod_{i=1}^{n} \lambda e^{-x_i \lambda} = \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i} \\ &= \lambda^n e^{-\lambda n \bar{x}}, \end{aligned}$$

where $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ is the sample mean.

The first derivative of the likelihood function's logarithm is:

$$\begin{aligned}\tfrac{d}{d\lambda}ln(L(\lambda)) &= \tfrac{d}{d\lambda}(nln(\lambda)-\lambda n\bar{x}) \\ &= \tfrac{n}{\lambda}-n\bar{x}\begin{cases} >0 & \text{if} \quad 0<\lambda<\frac{1}{\bar{x}}, \\ =0 & \text{if} \quad \lambda=\bar{x}, \\ <0 & \text{if} \quad \lambda>\frac{1}{\bar{x}}. \end{cases}\end{aligned}$$

The second derivative is easily obtained as

$$\frac{d^2}{d\lambda^2}ln(L(\lambda)) = -\frac{1}{\lambda^2} < 0.$$

Consequently the maximum likelihood estimate for the rate parameter is

$$\hat{\lambda} = \frac{1}{\bar{x}}.$$

Another classic method of parameter estimation is Moment generation method, which uses the moments of the distribution to estimate the parameter. Since the density function of exponential distribution only contains one single parameter λ, only first moment is needed.

$$\begin{aligned}E[x] &= \int_0^\infty xf(x)dx = \int_0^\infty x\lambda e^{-\lambda x}dx \\ &= -\int_0^\infty xde^{-\lambda x} \\ &= [-xe^{-\lambda x}]_0^\infty + \int_0^\infty e^{-\lambda x}dx \\ &= \Big[-\frac{1}{\lambda}e^{-\lambda x}\Big]_0^\infty = 0+\frac{1}{\lambda} = \frac{1}{\lambda}.\end{aligned}$$

Then let $E[x] = \bar{x}$, we have

$$\frac{1}{\lambda} = \tilde{x}.$$

So the moment generating estimator for the rate parameter is

$$\tilde{x} = \frac{1}{\bar{x}}.$$

1.3 Bayesian Inference

In Bayesian probability theory, if the posterior distribution $p(\theta|x)$ and the prior probability distribution $p(\theta)$ are in the same family, then the posterior and prior are called conjugate distribution, and the prior is called a conjugate prior for the likelihood function. For instance, the Gaussian family is self-conjugate to Gaussian likelihood function if the likelihood function is Gaussian. In order to have the posterior distribution to be Gaussian, one needs to choose a Gaussian prior over the mean. This means that Gaussian distribution is a conjugate prior for its likelihood function which is also Gaussian. Similarly, the conjugate prior for the exponential distribution is the gamma distribution(of which the exponential distribution is a special case). The following parametrization of the gamma probability density function is useful:

$$\text{Gamma}(\lambda; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} \exp(-\lambda\beta). \tag{1.6}$$

The posterior distribution p can then be expressed in terms of the likelihood function defined above and a gamma prior:

$$\begin{aligned} p(\lambda) &\propto L(\lambda) \times \text{Gamma}(\lambda; \alpha, \beta) \\ &= \lambda^n \exp\left(-\lambda n \overline{x}\right) \times \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} \exp(-\lambda\beta) \\ &\propto \lambda^{(\alpha+n)-1} \exp(-\lambda\left(\beta + n\overline{x}\right)). \end{aligned}$$

Now the posterior density p has been specified up to a missing normalizing constant.

Since it has the form of a gamma pdf, this can easily be filled in, and one obtains:

$$p(\lambda) = \text{Gamma}(\lambda; \alpha + n, \beta + n\overline{x}).$$

Here the hyperparameter α can be interpreted as the number of prior observations, and as the sum of the prior observations. The posterior mean here is:

$$\frac{\alpha + n}{\beta + n\overline{x}}.$$

1.4 Two-parameter Exponential Distribution

One commonly used generalization of the exponential distribution is the two-parameter exponential distribution. The density function is given as:

$$f_X(x) = \sigma^{-1} exp\{-(x-\theta)/\sigma\} I(x > \theta), \tag{1.7}$$

where σ is the scale parameter, which is equal to $\frac{1}{\lambda}$ shown in preview sections. And θ is the location parameter.For variables following two-parameter exponential density function, the possible values varies from θ to ∞. θ can be considered equal to 0 respect to one-parameter exponential distribution.

Similar to the properties with classic one-parameter exponential distributions, some useful results are shown as below; The expected value of a two-parameter exponential variable X is :

$$E[X] = \theta + \sigma.$$

The variance of an exponential variable X is :

$$Var[X] = \sigma^2.$$

So the standard deviation of the distribution is equal to σ. And the maximum likelihood estimator and moment generating estimator of σ is $\frac{1}{n}\sum_{i=1}^{n}(x_i - \theta)$. If $X_1, X_2, \ldots, X_n$ are independent random variables each following two-parameter exponential distribution, then the maximum likelihood estimators of θ and σ are

$$\hat{\theta} = min(X_1, X_2, \ldots, X_n),$$
$$\hat{\sigma} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \hat{\theta}) = \bar{X} - \hat{\theta}.$$

If θ is known, the maximum likelihood estimator of σ is $(\bar{X} - \theta)$. Even with σ known, $\hat{\theta}$ above is still the maximum likelihood estimator of θ. The probability density function of $\hat{\theta}$ is

$$f_{\hat{\theta}}(x) = (n/\theta)exp\{-n(x-\theta)/\sigma\}I(x > \theta),$$

which is of the same form as (1.7) but with σ replaced by σ/n. The variance of $\hat{\theta}$ is therefore σ^2/n^2, and its expected value is $\theta + \sigma/n$. It is interesting to note that the variance is proportional to n^{-2} and not to n^{-1}. The expected value of $\hat{\sigma}[= \bar{X} - \hat{\theta}]$ is $\sigma(1 - n^{-1})$, and its variance is $\sigma^2[n^{-1} + n^{-2} - 2n^{-3}]$. And the expected value of $(\bar{X} - \theta)$ is σ and its variance is $\sigma^2 n^{-1}$.

Moment estimators $(\tilde{\theta}, \tilde{\sigma})$ of (θ, σ) can be obtained by equating sample and population values of the mean and variance. They are

$$\tilde{\theta} = \bar{X} - \tilde{\sigma},$$
$$\tilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2.$$

Cohen and Helm (1973) discuss modified moment estimators obtained by replacing the second equation above by an equation that puts the first-order statistic X_1' equal to

its expected value. This gives

$$\tilde{\theta^*} + n^{-1}\tilde{\sigma^*} = X_1',$$

which leads to

$$\tilde{\theta^*} = \frac{nX_1' - \bar{X}}{n-1},$$
$$\tilde{\sigma^*} = \frac{n(\bar{X} - X_1')}{n-1}.$$

They show that these are minimum variance unbiased estimators (and a fortiori BLUEs). Also

$$Var(\tilde{\theta^*}) = \frac{\sigma^2}{n(n-1)}, \tag{1.8}$$

$$Var(\tilde{\sigma^*}) = \frac{\sigma^2}{n-1}, \tag{1.9}$$

$$Cov(\tilde{\theta^*}, \tilde{\sigma^*}) = \frac{\sigma^2}{n(n-1)}, \tag{1.10}$$

so that $Corr(\tilde{\theta^*}, \tilde{\sigma^*}) = 1/\sqrt{n}$. Further, since $\tilde{\sigma^*}$ is distributed as $\frac{1}{2}(n-1)^{-1}\chi^2_{2(n-1)}$, a $100(1-\alpha)\%$ confidence interval for σ is

$$\Big(\frac{2(n-1)}{\chi^2_{2(n-1),1-\frac{\alpha}{2}}}\tilde{\sigma^*}, \frac{2(n-1)}{\chi^2_{2(n-1),\frac{\alpha}{2}}}\tilde{\sigma^*}\Big). \tag{1.11}$$

Two-parameter exponential distribution has been used extensively in many reliability and life testing experiments for describing the failure rates of complex equipment, vacuum tubes and so on. It has also been recommended as a statistical model in clinical trials, such as the studies of behavior of tumor systems in animals and analysis of survival data in cancel research. The relative applications can be found in Johnson and Kotz (1994), Bain and Engelhardt (1991), Lawless and Singhal (1980), Zelen (1966).

1.5 Ranking and Partition Problems

In everyday life, one decides on the best medicine, best machine, best strategy or the best route for a destination, among a number of available options. In the statistical literature, such selections have been routinely carried out under the area of multiple comparisons. A commonly used statistical tool called Analysis of Variance (ANOVA) has been used extensively by practitioners to test whether or not the given treatments under consideration are all same or not. Generally, the ANOVA test is followed by some multiple comparisons tests, such as, LSD, Tukey Method, Scheffe Method to name a few, to decide which treatments are different from one another. However, those methods perform shortage in reality when implementation.

For example in clinical trials, a usual concern is comparing efficacy of the several essentially different varieties of drugs. A conclusion that whether those different drugs have the same efficacy or not can be easily obtained by setting that as the null hypotheses using above methods, which is meaningless. One main reason is that those different varieties of drugs tend to perform differently in most cases. But that is far away from enough to make any market value. Because rather than detecting the difference of the efficacy among those varieties of drugs, the experimenters are more willing to explore the one or several drugs that show better efficacy. So detecting the best or worst drugs is the need, which helps with business strategy making, which ANOVA and those comparison tests can not help achieve.

Thus, the experimenter's problem should not be only testing the equality of efficacy of these drugs, but rather to select the best one. The definition of the best would vary from situation to situation and it is generally for the experts in the area to dictate what best means in a given situation. For example, in some clinical trials. Sometimes, practitioners have even incorrectly used the ANOVA tests to even select the best treatment based on

the ranking of the means without realizing that the ANOVA test is designed to test if the given treatments are all same or not. The ANOVA test is not designed to select the best treatment and one cannot associate a probability statement with the selected treatment as being the best via the ANOVA approach.

In a pioneer work, Bechhofer (1954) introduced the concept of indifference-zone formulation and formulated some methodologies for the problem of selecting the best treatment from a set of several treatments. The formulation by Bechhofer had the desired property of selecting the best treatment with the pre-specified probability of correct selection. The formulation proposed by Bechhofer (1954) is referred to as the indifference-zone formulation in the statistical literature. Around the same time, Gupta (1956) formulated a strategy which controls the probability of correct selection in the whole parameter space, as opposed to the preferencezone which was the case under Bechhofers approach. The formulation of Gupta (1956), selects a subset of random size which includes the best treatment with some pre-specified probability. The formulation proposed by Gupta (1956) is referred to as the subset-selection formulation in the statistical literature.

1.6 Formulation

In this section, the idea of two most fundamental approaches in the area of selection and ranking are introduced under a classic case of population partition problem. Suppose we have π_i, $i = 1, \ldots, k(\geq 2)$, independent normally distributed populations, having unknown means μ_i and common unknown variance σ^2. We assume that $\mu_i \in R$ and $\sigma \in R^+$, i =1,...k. Let $\mu_{[1]}, \ldots, \mu_{[k]}$ be the ordered μ-values, Since the variance is same among all populations, we sample equally from each population and at any point of time, whenever we need new" samples from $\pi's$, we take a certain equal number, to be made specific later, of samples from each π at that step. our aim is to select the population

associated with $\mu_{[k]}$, and such a population is called the best population. We assume that there is just one population associated with $\mu_{[k]}$. We do not, however assume any knowledge about the association of the $\mu_i^{,}s$ with $\mu_{[i]}^{,}s$.

1.6.1 Indifference Zone Approach

Next, Bechhofer's (1954)indifference zone approach is introduced. Given $\delta^* > 0$, we define

$$\begin{aligned} \Omega &= \{\underline{\mu} = (\mu_1, ..., \mu_k) : \mu_i \in R, i = 1, ..., k\}, \\ \Omega(\delta^*) &= \{\underline{\mu} = (\mu_1, ..., \mu_k) : \mu_{[k]} - \mu_{[k-1]} \geq \delta^*\}, \end{aligned}$$

where Ω is the whole parameter space for $\underline{\mu}$ and $\Omega(\delta^*)$ is called the preference zone. For a given $P^* \in (k^{-1}, 1)$, we are interested in selection procedures such that the probability of correct selection (CS) of the population associated with $\mu_{[k]}$ is at least P^* or asymptotically (as $\delta^* \to 0$) at least P^* whenever the true parameter $\underline{\mu} \in \Omega(\delta^*)$. In other words, we are interested to identify the best population having certain minimum or approximately minimum probability P^* that our final decision is the right one when the best population's mean is at least δ^*-unit ahead of the mean of the second best. The parameter space $\Omega(\delta^*)$ is called, the preference zone, while $\Omega^C(\delta^*)(= \Omega - \Omega(\delta^*))$ is called the indifference zone in the sense that the experimenter is not willing to pick the best population when $\mu_{[k]} - \mu_{[k-1]} < \delta^*$, that is the experimenter expresses indifference in the parameter space $\Omega^C(\delta^*)$. Because for the indifference zone $\Omega^C(\delta^*)$ the best population is apparently not that much better than the second best.

1.6.2 Random Subset Approach

Gupta (1956) assigned a procedure for selecting a subset such that the probability that all the populations better than the standard are included in the subset is equal to or

greater than a predetermined number P^*. His goal is to separate those treatments which are better than the control from those that are worse(or not better). Considering the problem of selecting the best one of k categories when comparing k-1 categories with a control or standard, he controls the probability of selecting the standard as the best when the categories are equal to(or worse than) the control. Retain in the selected subset those and only those populations $\Omega_{[i]}(i = 1, 2, ..., k)$ for which

$$\bar{x}_{[i]} \geq \mu_0 - d\sigma/\sqrt{n_i}.$$

To determine the value of d let k_1,k_2 denote the true number of populations with $\mu \geq \mu_0$ and $\mu < \mu_0$, respectively, so that $k_1 + k_2 = k$. The the probability P of retaining all the k_1 populations with $\mu \geq \mu_0$ is given by

$$\begin{aligned} P &= \prod_{i=1}^{k_1} P\{\bar{x}_i' \geq \mu_0 - d\sigma\sqrt{n_i'}\} \\ &= \prod_{i=1}^{k_1} P\{\sqrt{n_i'}(\bar{x}' - \mu_i')/\sigma \geq -d + \sqrt{n_i'}(\mu_0 - \mu_i'/\sigma\}. \end{aligned} \tag{1.12}$$

where primes refer to values associated with the k_1 populations for which $\mu \geq \mu_0$. Hence

$$P = \prod_{i=1}^{k1}\{1 - F(-d + \sqrt{n_i'}(\mu_0 - \mu_i')/\sigma\},$$

where F(x) refers to the standard normal cumulative distribution function. The μ_i' above are restricted by the condition $\mu_i' \geq \mu_0$ and minimum of equation above is attained by setting $\mu_i' = \mu_0(i = 1, 2, ..., k_1)$. Now since result depends on the unknown integer k1, we can obtain a lower bound by setting $k_1 = k$. Then using the symmetry of F we have

$$P \geq [F(d)]^p.$$

The equation determining d is obtained by setting the right-hand member of equation above equal to P^* and is given by

$$F(d) = (P^*)^{1/p}.$$

1.7 Selecting Best Normal Populations

As mentioned in preview sections, under some cases checking whether the means of several populations are equal or not is far from enough. People need to find the "best" one which have the greatest location parameter. In this section we derive this problem under the setting of Indifference-Zone from Bechhofer. Mukhopadhyay, N. and Solanky, T.K.S. (1994) also summarize the relative results.

Suppose that we have $k(\geq)$ independent, normally distributed populations $\pi_1, \ldots, \pi_k$, with unknown mean μ_i and common unknown variance σ^2, with the density function given by

$$f(x_{ij}) = 1/\sqrt{2\pi\sigma^2} exp\{-(x_{ij} - \mu_i)^2/2\sigma^2\}. \tag{1.13}$$

Let us define the following:

$$\begin{aligned} \bar{X}_i &= n^{-1} \textstyle\sum_{j=1}^{n} X_{ij}, \\ V_i &= (n-1)^{-1} \textstyle\sum_{j=1}^{n} (X_{ij} - \bar{X}_i)^2, \\ \hat{\sigma}^2 &= \textstyle\sum_{i=0}^{k} \sum_{j=1}^{n} (X_{ij} - \bar{X}_i)^2/((k+1)(n-1)), \end{aligned} \tag{1.14}$$

where $\bar{X}_i$ is the point estimate of μ_i and $\hat{\sigma}$ is the pooled estimate of σ if a common unknown standard deviation is assumed. And $\mu_{[1]}, \ldots, \mu_{[k]}$ are defined as the ordered μ-values. Right now it is assumed that σ is known and $X_{i1}, \ldots, X_{in}$ from π_i have been recorded. The natural selection rule is simply to pick the population associated with the largest sample mean $\bar{X}_i$, $i = 1, \ldots, k$, which is procedure below.

R^*: Select π_j as the best population if $\bar{X_{jn}} = Max_{1\leq i\leq k}\bar{X}_{in}$.

It is denoted that $\phi(y) = (2\pi)^{-\frac{1}{2}}exp(-y^2/2)$ and $\Phi(y) = \int_{-\infty}^{y}\phi(x)dx$, for all $y \in R$. Let $\pi_{(i)}$ stand for the population associated with the location parameter $\mu_{[i]}$ and let us write $\bar{X}_{(in)}$ for the sample mean arising from the population $\pi_{(i)}, i = 1, \ldots, k$. Also, write $\delta_i = \mu_{[k]} - \mu_{[i]}$ and note that $\delta_i \geq 0$ for arbitrary $\underline{\mu} \in \Omega(\delta^*) \cup \Omega^c(\delta^*)$ and $\delta_i \geq \delta^*$ for arbitrary $\underline{\mu} \in \Omega(\delta^*)$. Then, the probability of correct selection is given by

$$\begin{aligned} P_n(CS) &= P\{\bar{X}_{kn} > \bar{X}_{in}, i = 1, \ldots, k-1\} \\ &= P\{n^{\frac{1}{2}}\sigma^{-1}(\bar{X}_{kn} - \mu_{[k]}) + n^{\frac{1}{2}}\sigma^{-1}(\mu_{[k]} - \mu_{[i]}) > n^{\frac{1}{2}}\sigma^{-1}(\bar{X}_{in} - \mu_{[i]}), i = 1, \ldots, k-1\} \\ &= P\{Y_{kn} + n^{\frac{1}{2}}\delta_i\sigma^{-1} > Y_{in}, i = 1, ..., k-1\}, \end{aligned}$$

where $Y_{in} = n^{\frac{1}{2}}(\bar{X}_{in} - \mu_{[i]})/\sigma, i = 1, ..., k$, which is following a standard normal distribution. From equation above, for all $\underline{\mu} \in \Omega(\delta^*)$,

$$\begin{aligned} P_n(CS) &= \int_{-\infty}^{\infty}\prod_{i=1}^{k-1}\Phi(y + n^{\frac{1}{2}}\delta_i\sigma^{-1})\phi(y)dy \\ &\geq \int_{-\infty}^{\infty}\Phi(y + n^{\frac{1}{2}}\delta^*\sigma^{-1})\phi(y)dy, \end{aligned}$$

where the equality holds when $\delta_i = \delta^*$ for all $i = 1, ..., k-1$. Therefore $\mu_{[1]} = \cdots = \mu_{[k-1]} = \mu_{[k]} - \delta^*$, a parameter configuration is referred to as the least favorable configuration(LFC). Hence,

$$Inf_{\mu\in\Omega(\delta^*)}P_n(CS) = \int_{-\infty}^{\infty}\Phi(y + n^{\frac{1}{2}}\delta^*\sigma^{-1})\phi(y)dy, \tag{1.15}$$

while this infimum is attained at the LFC. Let $h = h(k, P^*)$ be such that

$$\int_{-\infty}^{\infty}[\Phi(y + h)]^{k-1}\phi(y)dy = P^*, \tag{1.16}$$

then it follows that $P_n(CS) \geq P^*$ for all $\underline{\mu} \in \Omega(\delta^*)$ when $n \geq h^2\sigma^2/\delta^{*2} = C$. In other words, if σ was given, the experimenter should take $[C]^* + 1$ samples from each π and and

implement the corresponding selection rule above to pick the best normal distribution in order to meet the P^* requirement for P(CS) in $\Omega(\delta^*)$. However C is unknown. Moreover, one can show that there exists no fixed-sample size procedure that will meet the P^* requirement for $P(CS)$ in $\Omega(\delta^*)$, uniformly in $\sigma \in R^+$. The values of "h" was tabulated in Bechhofer (1954). For the case of σ being unknown, a two-stage procedure had been proposed by Bechhofer et al. (1954) and a purely sequential procedure had been proposed by Robbins et al. (1968).

1.8 Selecting Best Two-parameter Exponential Populations

Two-parameter exponential model has applications in many fields. In reliability and engineering, the location parameter is guarantee life of a component and scale parameter is the average life. When distribution is used to model the life lengths in dose-response experiments, location and scale parameters are referred as guaranteed and average effective duration of a drug respectively. For some applications of two-parameter exponential model one may refer to Johnson and Kotz (1994), Bain and Engelhardt (1991), Lawless and Singhal (1980), Zelen (1966).

Many researchers have discussed the problem of simultaneous comparisons of exponential location parameters when scale parameters are equal or unequal under heteroscedasticity. Ng et.al (1993) proposed multiple comparison procedures with control when scale parameters are equal. Wu et al. (2010) developed one-stage procedure of multiple comparisons with the control for exponential location parameters under heteroscedasticity. The procedures proposed by these researchers make use of Lam (1987,1988) technique to construct SCIs(Simultaneous Confidence Intervals).

While another approach of comparing multiple exponential distributions respect to

their location parameters come from the setting of Indifference-Zone as mentioned in previews sections which originated in comparing several normal populations.

Suppose that we have $k(\geq 2)$ independent and exponentially population $\pi_1, ..., \pi_k$, with density function of $\pi_i : i = 1, 2, \ldots, k$ given by

$$f_X(x) = \sigma^{-1} exp\{-(x_{ij} - \theta_i)/\sigma\} I(x_{ij} > \theta_i). \tag{1.17}$$

Let us define the following:

$$\begin{aligned} T_i &= Min_{1\leq j\leq n}\{X_{i1}, ..., X_{in}\}, \\ V_i &= (n-1)^{-1} \textstyle\sum_{j=1}^{n}(X_{ij} - T_i), \\ \hat{\sigma} &= \textstyle\sum_{i=0}^{k}\sum_{j=1}^{n}(X_{ij} - T_i)/((k+1)(n-1)), \end{aligned} \tag{1.18}$$

where T_i are the order statistic or point estimate of θ_i and $\hat{\sigma}$ is the pooled estimate of σ if a common unknown scale parameter is assumed. The density function of T is derived as

$$\begin{aligned} F(T) \quad &= P(T \leq t) = 1 - P(T > t) = 1 - P(X_1 > t, X_2 > t, \ldots, X(n) > t) \\ &= 1 - \textstyle\prod_i^n P(X_i > t) = 1 - \prod_i^n [1 - F(X_i)] \\ &= 1 - [1 - F(t)]^n. \end{aligned}$$

Since $F(X) = 1 - e^{-\frac{x-\theta}{\sigma}}, X > \theta$, then the cumulative function and density function of T are obtained as

$$\begin{aligned} F_T(t) &= 1 - e^{-\frac{n(t-\theta)}{\sigma}} \quad t > \theta, \\ f_T(t) &= [F(T)]' = \tfrac{n}{\sigma} e^{-\frac{n(t-\theta)}{\sigma}} \quad t > \theta. \end{aligned} \tag{1.19}$$

It is clear that T_is are following exponential distribution with location parameter of θ and scale parameter of $\frac{\sigma}{n}$. Here $\theta_{[1]}, \ldots, \theta_{[k]}$ are defined as the ordered θ-values. The partition problem was originated from the purpose of selecting the best two-parameter exponential

distribution, which has the largest or smallest location parameter. It is assumed that σ is known and the observations $X_{i1}, ..., X_{in}$ from $\pi_i, i = 1, \ldots, k$ have been recorded. The natural selection rule is simply to pick the population associated with the largest T_i value, i = 1,..., k, which is procedure

R^*: Select π_j as the best population if $T_j = Max_{1 \le i \le k} T_{in}$.

Let $\pi_{(i)}$ stand for the population associated with the location parameter $\theta_{[i]}$ and let us write T_{in} for the sample minimum arising from the population $\pi_{(i)}, i = 1, ..., k$. Also, write $\delta_i = \theta_{[k]} - \theta_{[i]}$ and note that $\delta_i \ge 0$ for arbitrary $\underline{\theta} \in \Omega(\delta^*) \cup \Omega^c(\delta^*)$ and $\delta_i \ge \delta^*$ for arbitrary $\underline{\theta} \in \Omega(\delta^*)$. Then, the probability of correct selection is given by

$$\begin{aligned} P_n(CS) &= P\{T_{kn} > T_{in}, i = 1, ..., k-1\} \\ &= P\{n\sigma^{-1}(T_{kn} - \theta_{[k]}) + n\sigma^{-1}(\theta_{[k]} - \theta_{[i]}) > n\sigma^{-1}(T_{in} - \theta_{[i]}), i = 1, ..., k-1\} \\ &= P\{Z_{kn} + n\delta_i\sigma^{-1} > Z_{in}, i = 1, ..., k-1\}, \end{aligned}$$

where $Z_{in} = n\{T_{in} - \theta_{[i]}\}/\sigma, i = 1, ..., k$, which are all following standard exponential distribution. From equation above, for all $\underline{\theta} \in \Omega(\delta^*)$,

$$\begin{aligned} P_n(CS) &= \int_0^\infty \prod_{i=1}^{k-1} [1 - exp(-z - n\delta_i\sigma^{-1})] exp(-z) dz \\ &\ge \int_0^\infty [1 - exp(-z - n\delta^*\sigma - 1)]^{k-1} exp(-z) dz, \end{aligned}$$

where the equality holds when $\delta_i = \delta^*$ for all $i = 1, ..., k-1$. Therefore $\theta_{[1]} = \cdots = \theta_{[k-1]} = \theta_{[k]} - \delta^*$, a parameter configuration is referred to as the least favorable configuration(LFC). Hence,

$$Inf_{\theta \in \Omega(\delta^*)} P_n(CS) = \int_0^\infty [1 - exp(-z - n\delta^*\sigma^{-1})]^{k-1} exp(-z) dz, \tag{1.20}$$

while this infimum is attained at the LFC. Let $b = b(k, P^*)$ be such that

$$\int_0^\infty \{1 - exp(-z - b))^{k-1} exp(-z) dz = P^*, \tag{1.21}$$

then it follows that $P_n(CS) \geq P^*$ for all $\underline{\theta} \in \Omega(\delta^*)$ when $n \geq b\sigma/\delta^* = C$. The values of "b" was tabulated in Raghavachari and Starr(1970). In other words, if σ was given, the experimenter should take $[C]^*+1$ samples from each π and and implement the corresponding selection rule above to pick the best two-parameter negative exponential distribution in order to meet the P^* requirement for P(CS) in $\Omega(\delta^*)$. For the case of σ being unknown, a two-stage procedure had been proposed by Desu et al.(1977) and a purely sequential procedure had been proposed by Mukhopadhyay(1986).

1.9 Partition Problem

However, in many cases the need for experimenters has been beyond selecting the best treatment. The experimenter may want the best to be some "specified amount better than what is already in use (known as Control or Standard). This requirement forced the researchers to seek out alternative formulations and thus the problem of comparisons with a control originated. The problem of comparisons with a control has been investigated by many researchers under different types of formulations, and under different criteria to be satisfied by an acceptable procedure. Among the early investigations some research related to comparisons with respect to a control population had been done. In 1969, Tong published his paper of partitioning a set of normal distributions by their locations with respect to a control through the formulation of Bechhofer's indifference zone.

Assume that there are $(k+1)$ independent populations, $\pi_0, \pi_1, ..., \pi_k$, with unknown location parameters $\mu_i, i = 0, 1, ..., k$, but common scale parameter σ^2. π_0 is denoted as the standard or control populations. Given arbitrary but fixed constants δ_1 and δ_2, $\delta_1 < \delta_2$, define three subsets along the lines of the Bechhofer's (1954) indifference zone formulation as

$$\begin{aligned}\Omega_B &= \{\pi_i : \mu_i \leq \mu_0 + \delta_1, i = 1, ..., k\},\\ \Omega_I &= \{\pi_i : \mu_0 + \delta_1 < \mu_i < \mu_0 + \delta_2, i = 1, ..., k\},\\ \Omega_G &= \{\pi_i : \mu_i \geq \mu_0 + \delta_2, i = 1, ..., k\}.\end{aligned} \tag{1.22}$$

We refer to Ω_G as the set of "good populations" and Ω_B as the set of "bad populations". It is important to note that the choice of the constants δ_1 and δ_2 is generally provided by the experimenters. Our aim is to correctly partition the populations into that two sets. The Ω_I is considered as the indifference zone set and a correct decision puts no restrictions on the partition of the populations belonging to this set. Next, with high accuracy, we want to partition the set Ω into two disjoint subsets S_B and S_G, such that, $\Omega_B \subseteq S_B$ and $\Omega_G \subseteq S_G$. Such a partition is know in the literature as a correct decision(CD). In other words, given a preassigned number P^*,$2^{-k} < P^* < 1$, we seek statistical methodologies $\wp$ to determine S_B and S_G, such that

$$P\{CD|\boldsymbol{\mu}, \sigma^2, \wp\} \geq P^* \qquad \forall\ \boldsymbol{\mu} \in \boldsymbol{R}^{k+1},\ \sigma \in \boldsymbol{R}^+, \tag{1.23}$$

where $\boldsymbol{\mu} = (\mu_0, \mu_1, \cdots, \mu_k)'$.

Tong (1969) considered the following decision rule to partition the set of treatments Ω, based on some appropriately N observations from each of the k treatments and the control population:

$$\begin{aligned}S_B &= \{\pi_i : \bar{X}_{iN} - \bar{X}_{0N} < d,\ i = 1, \cdots, k\},\\ S_G &= \{\pi_i : \bar{X}_{iN} - \bar{X}_{0N} > d,\ i = 1, \cdots, k\},\end{aligned} \tag{1.24}$$

where $\bar{X}_{iN}$ is the sample mean from $\pi_i, i = 0, 1, \cdots, k$. The following is defined for later

use:

$$d = (\delta_1 + \delta_2)/2, \quad a = (-\delta_1 + \delta_2)/2, \quad \lambda = \sigma/a, \text{ and}$$
$$m = \begin{cases} k/2 & \text{if } k \text{ is even,} \\ (k+1)/2 & \text{if } k \text{ is odd.} \end{cases} \tag{1.25}$$

Next, using the partition rule (1.24), Tong (1969) showed that the probability of correct decision for the normally distributed populations can be expressed as

$$\underset{\mu \in R^{k+1}}{Inf} P[CD] = \int_{-\infty}^{\left(\frac{1}{2}N\right)^{\frac{1}{2}}/\lambda} \cdots \int_{-\infty}^{\left(\frac{1}{2}N\right)^{\frac{1}{2}}/\lambda} \frac{|\Sigma|^{\frac{1}{2}}}{(2\pi)^{\frac{k}{2}}} \exp\left(-\frac{y'\Sigma^{-1}y}{2}\right) dy_1 \cdots dy_k, \tag{1.26}$$

where $y' = (y_1, \cdots, y_k)$ has a multivariate normal distribution with mean vector of zero, the covariance matrix Σ is given by

$$\Sigma = \begin{pmatrix} 1 & & \frac{1}{2} & -\frac{1}{2} & \cdots & -\frac{1}{2} \\ & \ddots & & \vdots & \ddots & \vdots \\ \frac{1}{2} & & 1 & -\frac{1}{2} & \cdots & -\frac{1}{2} \\ -\frac{1}{2} & \cdots & -\frac{1}{2} & 1 & & \frac{1}{2} \\ \vdots & \ddots & \vdots & & \ddots & \\ -\frac{1}{2} & \cdots & -\frac{1}{2} & \frac{1}{2} & & 1 \end{pmatrix},$$

and the infimum for $P(CD)$ is attained if $\mu_1 = \mu_2 = \cdots = \mu_m = \mu_0 + \delta_1$ and $\mu_{m+1} = \mu_{m+2} = \cdots = \mu_k = \mu_0 + \delta_2$. In the statistical literature, this parameter configuration is known as the *least favorable configuration* (LFC). Next, suppose b is a constant satisfying

$$P^* = \int_{-\infty}^{b} \cdots \int_{-\infty}^{b} \frac{|\Sigma|^{\frac{1}{2}}}{(2\pi)^{\frac{k}{2}}} \exp\left(-\frac{y'\Sigma^{-1}y}{2}\right) dy_1 \cdots dy_k. \tag{1.27}$$

Then, if we take a sample of size N, where $N \geqslant 2\lambda^2 b^2$, and partition the k treatments

according to the partition rule (1.24), we have

$$P\left[CD\right] \geqslant P^*, \qquad \forall\ \boldsymbol{\mu} \in \boldsymbol{R}^{k+1},\ \sigma \in \boldsymbol{R}^{+}.$$

The values of design constant b have been tabulated in the Table 1 of Tong (1969). For the unknown σ^2 case, Tong (1969) also constructed a two-stage and a purely sequential procedure. For the unknown σ^2 case, Datta and Mukhopadhyay (1998) had constructed a fine-tuned purely sequential procedure and some other multistage methodologies, emphasizing the second-order asymptotic properties. However, the sequential procedures are known to be operationally inconvenient and rather cumbersome to use, as decisions and computations need to be carried out after each stage of the sampling process. With that as the motivation, Solanky (2006) constructed a two-stage procedure with elimination which eliminates too inferior or too superior populations after the stage one of the sampling process and in the stage two sampling is carried out only from the competing treatments.

2

Partition Problem for Exponential Distributions

In the chapter 1, the method of partitioning a set of normal populations respect to the location parameter was introduced. However, what about the case when the populations needed partitioning are not following a normal distribution? For example, suppose several new developed medicine needs to be classified as "Good" and "Bad" based on the effectiveness compared a control medicine. Or, several different brand light bulbs need partitioning according to their lifetime compared to a standard bulb. In these cases, the populations will probably be distributed as two-parameter exponential distribution, and will require methodology to partition two-parameter exponential distributions with regarding to a control. In this chapter, a procedure of partitioning several two-parameter exponential distributions respect to their location parameter is introduced. The first few sections show the logical thinking and derivation of design parameters for this problem, including a detailed discussion on existence and derivation of the Least Favorable configuration. The later part of the chapter is devoted to developing first and second order asymptotic properties of the proposed procedure.

2.1 Formulation for Partitioning Two-parameter Exponential Distributions

In this section, we will introduce the formulation of partition problem for two-parameter exponential distributions respect to the location parameter along with the setting of Bechhofer's (1954) indifference zone idea and give some necessary notations of some statistical symbols for later sections' use.

It is assume that there are $(k+1)$ independent negative exponential populations $\pi_0, \pi_1, \ldots, \pi_k$, with unknown location parameters $\theta, i = 0, 1, \ldots, k$, and common scale parameter σ. $\pi_1, \ldots, \pi_k$ are denoted as the populations to be partitioned and π_0 is denoted as the standard or control populations. For $i = 0, \ldots, k$, the density function of the population is expressed as

$$f_X(x_{ij}) = \sigma^{-1} exp\{-(x_{ij} - \theta_i)/\sigma\} I(x_{ij} > \theta_i). \tag{2.1}$$

Given arbitrary but fixed constants δ_1 and δ_2, let us define three subsets along the lines of the Bechhofer's (1954) indifference zone formulation as

$$\begin{aligned} \Omega_B &= \{\pi_i : \theta_i \leq \theta_0 + \delta_1, i = 1, ..., k\}, \\ \Omega_I &= \{\pi_i : \theta_0 + \delta_1 < \theta_i < \theta_0 + \delta_2, i = 1, ..., k\}, \\ \Omega_G &= \{\pi_i : \theta_i \geq \theta_0 + \delta_2, i = 1, ..., k\}. \end{aligned} \tag{2.2}$$

We refer to Ω_G as the set of "good populations" and Ω_B as the set of "bad populations". It is important to note that the choice of the constants δ_1 and δ_2 is generally determined by the experimenters based on the prior knowledge of the practical situation. Our aim is to correctly partition the populations into two groups. The subset Ω_I is considered as the indifference zone set and a correct decision puts no restrictions on the partition

of the populations belonging to this set. Next, with high accuracy, we want to partition the set Ω into two disjoint subsets S_B and S_G, such that, $\Omega_B \subseteq S_B$ and $\Omega_G \subseteq S_G$. Such a partition can guarantee that all true "good population" and "bad population" are correctly labeled, which is known in the literature as a correct decision(CD). In other words, given a preassigned value P^*, $2^{-k} < P^* < 1$, we seek statistical methodologies $\wp$ to determine S_B and S_G, such that

$$P\{CD|\boldsymbol{\theta}, \sigma, \wp\} \geq P^* \qquad \forall\ \boldsymbol{\theta} \in \boldsymbol{R}^{k+1},\ \sigma \in \boldsymbol{R}^+, \tag{2.3}$$

where $\boldsymbol{\theta} = (\theta_0, \theta_1, \cdots, \theta_k)'$.

Alone with the setting originated in Tong (1969), with N observations from each of the k populations to be partitioned and the control population, we define a partition rule $\wp$ was defined as:

$$\begin{aligned} S_B &= \{\pi_i : T_i - T_0 < \ d,\ i = 1, \cdots, k\}, \\ S_G &= \{\pi_i : T_i - T_0 > \ d,\ i = 1, \cdots, k\}, \end{aligned} \tag{2.4}$$

where $T_i = Min_{1 \leq j \leq N}\{X_{i1}, \ldots, X_{iN}\}$ is a point estimator of $\theta_i, i = 0, 1, \cdots, k$. Through this decision rule, the probability of correct partition is controlled by the selection size. Or equivalently, the power of grouping procedure according the the rule $\wp$ is controlled by the setting of collection size of the $k+1$ populations. Next we will show how this relationship takes effect. In order to make the derivation clear, the following are defined for later use:

$$\begin{aligned} &d = (\delta_1 + \delta_2)/2, \quad a = (-\delta_1 + \delta_2)/2, \quad \lambda = \sigma/a, \text{ and} \\ &m = \begin{cases} k/2 & \text{if } k \text{ is even}, \\ (k+1)/2 & \text{if } k \text{ is odd}. \end{cases} \end{aligned} \tag{2.5}$$

For simplicity let us first assume that the scale parameter σ is known and a set of

exponential populations are partitioned based on rule $\wp$. Then the probability of making a correct decision($\Omega_B \subseteq S_B, \Omega_G \subseteq S_G$) can be obtained as

$$P(CD|\sigma, \wp) = P[T_i - T_0 < d, T_j - T_0 > d, 0 \leq i \leq N_1, 0 \leq j \leq N_2|\sigma, \wp]. \tag{2.6}$$

Where N_1 is the number of populations in Ω_B and N_2 is the number of populations in Ω_G. Since the population location parameters are unknown, in order to guarantee (2.1.3) to hold, the infimum of the left hand side should be controlled greater or equal to P^*. Holding this purpose, the least favorable configuration of the population location parameters(LFC) $\boldsymbol{\theta}^0$ among $\boldsymbol{R}^{k+1}$ minimizing the probability of correct partition is being analyzed here. In other words, we are interested in finding the worst or hardest situation to do the partition process.

The above states the importance of detecting the LFC to the partition process. Next we try to tackle the task of finding the LFC.

Theorem 2.1.1 *If a vector $\boldsymbol{\theta}^0$ is a least favorable configuration under under procedure $\wp$, it has to satisfy the two conditions below:*
(i) Ω_I is empty.
(ii) All populations that needs to be partitioned have a location parameter lying on the boundaries which are either $\theta_0 + \delta_1$ or $\theta_0 + \delta_2$.

Proof: Suppose there is a parameter configuration $\boldsymbol{\theta}^1$ which satisfies that Ω_I is not empty and the PCD of it can be expressed as (2.6). Then $k - N_1 - N_2 > 0$ and there always exits another parameter configuration which is more difficult to be partitioned. For example, let us assume a configuration $\boldsymbol{\theta}^2$ that $k - N_1 - N_2$ populations are in Ω_B while the rest $N_1 + N_2$ populations have the same location parameter with $\boldsymbol{\theta}^1$, then the PCD of under

configuration $\boldsymbol{\theta}^2$ can be expressed as

$$
\begin{aligned}
& P(CD|\boldsymbol{\theta}^2, \sigma, \wp) \\
= \; & P[T_i - T_0 < d, T_j - T_0 > d, T_l - T_0 < d, 0 \le i \le N_1, 0 \le j \le N_2, \\
& 0 \le l \le k - N_1 - N_2 | \sigma, \wp] \\
= \; & P[T_i - T_0 < d, T_j - T_0 > d \cap T_l - T_0 < d] \\
= \; & P[T_i - T_0 < d, T_j - T_0 > d] \cdot P[T_l - T_0 < d | T_i - T_0 < d, T_j - T_0 > d] \\
< \; & P[T_i - T_0 < d, T_j - T_0 > d] = P[CD|\boldsymbol{\theta}^1, \sigma, \wp].
\end{aligned}
$$

Compared to the case above, next a case that not all populations' location parameter are on the boundary of the indifference set is discussed.

Under the setting of (i) suppose there is a parameter configuration $\boldsymbol{\theta}^3$ which satisfies that Ω_I is empty and the PCD of it can be expressed as (2.6). So now $k - N_1 - N_2 = 0$. Let $\theta_i = \theta_0 + \delta_1 + \tau_i,\ \theta_j = \theta_0 + \delta_2 + \tau_j,\ \tau_i, \tau_j \ge 0$ for $i = 0, \ldots, N_1$ and $j = 0, \ldots, N_2$ be assumed as the real setting of components. Then the PCD under configuration $\boldsymbol{\theta}^3$ can be expressed as

$$
\begin{aligned}
& P\Big[CD|\boldsymbol{\theta}^3,\ \sigma;\ \wp\Big] \\
= \; & P\Big[T_i - T_0 < d,\ T_j - T_0 > d,\ \ 0 < i \le N_1,\ \ 0 < j \le N_2 | \boldsymbol{\theta}^3,\ \sigma\Big] \\
= \; & P\Big[\frac{T_i - \theta_i}{\sigma/n} - \frac{T_0 - \theta_0}{\sigma/n} < \frac{d - \theta_i + \theta_0}{\sigma/n},\ \frac{T_j - \theta_j}{\sigma/n} - \frac{T_0 - \theta_0}{\sigma/n} > \frac{d - \theta_j + \theta_0}{\sigma/n}, \\
& 0 \le i \le N_1,\ \ 0 \le j \le N_2\Big] \\
= \; & P\Big[Z_i - Z_0 < \frac{d - (\delta_1 - \tau_i)}{\sigma/n},\ Z_j - Z_0 > \frac{d - (\delta_2 + \tau_j)}{\sigma/n},\ \ 0 \le i \le N_1,\ \ 0 \le j \le N_2\Big] \\
= \; & P\Big[Z_i - Z_0 < \frac{a}{\sigma/n} + \frac{\tau_i}{\sigma/n},\ Z_j - Z_0 > -\frac{a}{\sigma/n} - \frac{\tau_j}{\sigma/n},\ \ 0 \le i \le N_1,\ \ 0 \le j \le N_2\Big] \\
= \; & P\Big[Z_i - Z_0 < \frac{an}{\sigma} + \frac{\tau_i n}{\sigma},\ Z_0 - Z_j < \frac{an}{\sigma} + \frac{\tau_j n}{\sigma},\ \ 0 \le i \le N_1,\ \ 0 \le j \le N_2\Big] \\
\le \; & P\Big[Z_i - Z_0 < \frac{an}{\sigma},\ Z_0 - Z_j < \frac{an}{\sigma},\ \ 0 \le i \le N_1,\ \ 0 \le j \le N_2\Big],
\end{aligned}
$$

It can be shown that the infimum of PCD under configuration $\boldsymbol{\theta}^3$ is attached when $\tau_i, \tau_j = 0$ for $i = 0, \ldots, N_1$ and $j = 0, \ldots, N_2$. Equivalently speaking, the infimum is attached when all populations have a location parameter on the boundary of the indifference zone. ■

Next, under the setting of conditions from Theorem 2.1.1, suppose out of k populations there are r exponential populations having a location parameter at $\theta_i = \theta_0 + \delta_1, 0 \leq i \leq r$ and $k - r$ exponential populations having a location parameter at $\theta_j = \theta_0 + \delta_2, r > j \leq k$. Here comes a goal that we need to find the value of $r(0 \leq r \leq k)$ which leads the probability of the correct decision to attach the infimum.

Without the loss of generality assume that the first r populations have the location parameter $\theta_0+\delta_1$ and the rest $k-r$ populations have the location parameter $\theta_0+\delta_2$. Denoting this configuration as $\boldsymbol{\theta}^0(r)$, then the probability of *correct decision* can be expressed as

$$
\begin{aligned}
& P\Big[CD|\boldsymbol{\theta}^0(r),\ \sigma;\ \wp\Big] \\
= \ & P\Big[T_i - T_0 < d,\ T_j - T_0 > d,\ \ 0 < i \leq r,\ \ r < j \leq k|\boldsymbol{\theta}^0(r),\ \sigma\Big] \\
= \ & P\Big[\frac{T_i - \theta_i}{\sigma/n} - \frac{T_0 - \theta_0}{\sigma/n} < \frac{d - \theta_i + \theta_0}{\sigma/n},\ \frac{T_j - \theta_j}{\sigma/n} - \frac{T_0 - \theta_0}{\sigma/n} > \frac{d - \theta_j + \theta_0}{\sigma/n}, \\
& 1 \leq i \leq r,\ \ r+1 \leq j \leq k\Big] \\
= \ & P\Big[Z_i - Z_0 < \frac{d - \delta_1}{\sigma/n},\ Z_j - Z_0 > \frac{d - \delta_2}{\sigma/n},\ \ 1 \leq i \leq r,\ \ r+1 \leq j \leq k\Big] \\
= \ & P\Big[Z_i - Z_0 < \frac{an}{\sigma},\ Z_0 - Z_j < \frac{an}{\sigma},\ \ 1 \leq i \leq r,\ \ r+1 \leq j \leq k\Big], \qquad (2.7)
\end{aligned}
$$

where $Z_i = \frac{T_i - \theta_i}{\sigma/n}$ for $i = 0, \ldots, k$, are all following standard exponential distributions. The density function of Z is derived below:

According the (1.19),

$$\begin{aligned} F(Z) \quad &= P(Z \leq z) = P(\tfrac{T-\theta}{\sigma/n} \leq z) \\ &= P(T \leq \tfrac{z\sigma}{n} + \theta) = 1 - e^{-\frac{n(\frac{z\sigma}{n}+\theta-\theta)}{\sigma}} \\ &= 1 - e^{-z}. \end{aligned}$$

Then $f_Z(z) = [F(Z)]' = e^{-z}$, which is the density function of standard exponential distribution.

Let us define $Y_i = Z_i - Z_0$ for $0 < i \leq r$ and $Y_i = Z_0 - Z_i$ for $r < i \leq k$, then Y_i for $i = 1, \ldots, k$ are all following Laplace distribution. Then it can be obtained that $Y = (Y_1, Y_2, \ldots, Y_k)'$ is a joint distribution of k Laplace distributions. And the $(k \times k)$ covariance matrix $\Sigma_Y = (\sigma_{ij})$ as

$$\sigma_{ij} \quad = \quad \begin{cases} 2 & \text{for} \quad i = j, \\ 1 & \text{for} \quad i \neq j, \text{and } i, j \in [1, r] \text{ or } i, j \in \ [r+1, k], \\ -1 & \text{for} \quad i \neq j, \text{and } i \in [1, r] \text{ and } j \in \ [r+1, k]. \end{cases} \tag{2.8}$$

Then *probability of of correct decision* (2.7) can be considered as a probability related to a multivariate distribution. Let us denote the density function of Y as a multivariate distribution expressed as $f_Y(\cdot)$.

Next, we will investigate if the multivariate distribution $f_Y(\cdot)$ is a Multivariate Laplace distribution. And it is important to note that a multivariate Laplace distribution is not a unique distribution. Also, using (2.7), we will next investigate the value of r which will minimize the probability requirement from (2.7). This choice of r, known as LFC in the statistical literature, will have that at a given fixed significance level α

$$P\Big[CD|\boldsymbol{\theta}^0(r), \ \ \sigma\Big] \geq P^* = 1 - \alpha. \tag{2.9}$$

After the value of r is identified, some critical constant related to n can be computed correspondingly. That will help statisticians with determining the sample size needed if the power of partition procedure is required. These two different directions are detailed discussed in the later sections.

2.2 Multivariate Distribution Identification

In the last section, a multivariate distribution is derived based on the form of the probability expression of the correct decision. In this section, the statement that distribution of Y is a Multivariate Laplace distribution is verified. In case the statement is true, as was the case in Tong (1969), formula of probability of *correct decision* (2.7) can be evaluated as a integral of a Multivariate Laplace distribution, which is somewhat easy to solve with or without standard software packages. In addition, some good properties of Multivariate Laplace distribution can be developed related to our problem and then used to identify the LFC (Least Favorable Configuration).

It is known that the joint distribution of k normal variables is not guaranteed to follow a Multivariate Normal distribution. The Multivariate distribution requires that any linear combination of its components to follow the same type of marginal distribution. For example, a multivariate density is known as a multivariate normal density if any linear combination of its components also follows a normal density. The realization of this condition is the reason why Tong could derive a multivariate normal distribution for the partition problem respect a set of normal populations in Tong (1969). However, the multivariate Laplace distribution doesn't follow the same properties. Next, we show how a multivariate normal variable can be derived from normal partition problem and why the multivariate Laplace variable can not, under the partition problem considered in this thesis.

Theorem 2.2.1 *Suppose Z_0, Z_1, Z_2 are independent standard normal variables and $Y_1 = Z_1 - Z_0$ and $Y_2 = Z_2 - Z_0$, then the joint distribution of Y_1 and Y_2 is a bivariate normal distribution*

Proof: Note that in probability theory and statistics, the multivariate normal distribution is a generalization of the one-dimensional normal distribution to higher dimensions. One definition is that a random vector is said to be k-variate normally distributed if every linear combination of its k components has a univariate normal distribution. Therefore, to prove the joint distribution of Y_1 and Y_2 is not a bivariate normal distribution is equivalent to show that any linear combination of Y_1 and Y_2 is following a normal distribution is wrong. Now we try to find whether the distribution of $aY_1 + bY_2$ is normal distribution when a and b are not both 0. Let us define

$$X = aY_1 + bY_2 = aZ_1 + bZ_2 - (a+b)Z_0.$$

So the expectation and variance of X can be obtained

$$E[X] = aE[Z_1] + bE[Z_2] - (a+b)E[Z_0] = 0$$
$$Var[X] = a^2Var[Z_1] + b^2Var[Z_2] + (a+b)^2Var[Z_0] = 2(a^2 + ab + b^2),$$

To find the distribution of X, we first find the joint distribution of $(X_1\ X_2\ X)'$ by the method of transformation of variables. Then, derive the distribution of X by obtaining the marginal distribution. Here we let $X_1 = Z_1$ and $X_2 = Z_2$. Since Z_0, Z_1, Z_2 are independent normal variables. then the joint density of Z_0, Z_1, Z_2 can be expressed as

$$f(Z_1, Z_2, Z_0) = (2\pi)^{-\frac{3}{2}} e^{-\frac{1}{2}(z_1^2 + z_2^2 + z_0^2)}.$$

And the relationship of Z_0, Z_1, Z_2 and X, X_1, X_2 is

$$\begin{aligned} Z_1 &= X_1 \\ Z_2 &= X_2 \\ Z_0 &= \tfrac{a}{a+b}X_1 + \tfrac{b}{a+b}X_2 - \tfrac{1}{a+b}X \end{aligned} \qquad J = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{a}{a+b} & \frac{b}{a+b} & -\frac{1}{a+b} \end{vmatrix} = -\tfrac{1}{a+b}.$$

Then the joint density function of $(X_1\ X_2\ X)'$ can be obtained as

$$\begin{aligned} g(x_1,x_2,x) &= (2\pi)^{-\frac{3}{2}} e^{-\frac{1}{2}\left[x_1^2+x_2^2+(\frac{a}{a+b}x_1+\frac{b}{a+b}x_2-\frac{1}{a+b}x)^2\right]} \cdot |J| \\ &= (2\pi)^{-\frac{3}{2}} e^{-\frac{1}{2}\left[x_1^2+x_2^2+\frac{a^2}{(a+b)^2}x_1^2+\frac{b^2}{(a+b)^2}x_2^2+\frac{1}{(a+b)^2}x^2+\frac{2ab}{(a+b)^2}x_1x-\frac{2a}{(a+b)^2}x_1x-\frac{2b}{(a+b)^2}x_2x\right]} \cdot |\frac{1}{a+b}| \\ &= \frac{(2\pi)^{-\frac{3}{2}}}{|a+b|} e^{-\frac{1}{2}\left[\frac{2a^2+2ab+b^2}{(a+b)^2}x_1^2+\frac{2ab}{(a+b)^2)}x_1x_2-\frac{2a}{(a+b)^2}x_1x\right]} e^{-\frac{1}{2}\left[\frac{a^2+2ab+2b^2}{(a+b)^2}x_2^2-\frac{2b}{(a+b)^2}x_2x\right]} e^{-\frac{1}{2(a+b)^2}x^2}. \end{aligned}$$

Then the marginal density $g(x_2, x)$ can be solved as

$$\begin{aligned} g(x_2,x) &= \int_{-\infty}^{\infty} g(x_1,x_2,x)dx_1 \\ &= \frac{(2\pi)^{-\frac{3}{2}}}{|a+b|} e^{-\frac{1}{2}\left[\frac{a^2+2ab+b^2}{(a+b)^2}x_2^2-\frac{2b}{(a+b)^2}x_2x\right]} e^{-\frac{1}{2(a+b)^2}x^2} \\ &\quad \cdot \int_{-\infty}^{\infty} e^{-\frac{2a^2+2ab+2b^2}{2(a+b)^2}\left[(x_1-\frac{ax-abx_2}{2a^2+2ab+b^2})^2-(\frac{ax-abx_2}{2a^2+2ab+b^2})^2\right]} dx_1. \end{aligned}$$

Let us make the part of $g(x_2, x)$ as

$$\begin{aligned} A &= \int_{-\infty}^{\infty} e^{-\frac{2a^2+2ab+2b^2}{2(a+b)^2}\left[(x_1-\frac{ax-abx_2}{2a^2+2ab+b^2})^2-(\frac{ax-abx_2}{2a^2+2ab+b^2})^2\right]} dx_1 \\ &= e^{\frac{(ax-abx_2)^2}{2(a+b)^2(2a^2+2ab+b^2)}} \int_{-\infty}^{\infty} e^{-\frac{(x_1-\frac{ax-abx_2}{2a^2+2ab+b^2})^2}{2(a+b)^2/(2a^2+2ab+b^2)}} dx_1 \\ &= e^{\frac{(ax-abx_2)^2}{2(a+b)^2(2a^2+2ab+b^2)}} \cdot \Big[\frac{2\pi(a+b)^2}{2a^2+2ab+b^2}\Big]^{(-\frac{1}{2})(-1)} \\ &= \Big[\frac{2\pi(a+b)^2}{2a^2+2ab+b^2}\Big]^{\frac{1}{2}} e^{\frac{a^2b^2x_2^2-2a^2bx_2x+a^2x^2}{2(a+b)^2(2a^2+2ab+b^2)}}. \end{aligned}$$

Then the density of (X_2, X) can be simplified as

$$
\begin{aligned}
g(x_2,x) &= \frac{(2\pi)^{-\frac{3}{2}}}{|a+b|} e^{-\frac{1}{2}\left[\frac{a^2+2ab+b^2}{(a+b)^2}x_2^2 - \frac{2b}{(a+b)^2}x_2 x\right]} e^{-\frac{1}{2(a+b)^2}x^2} \cdot \left[\frac{2\pi(a+b)^2}{2a^2+2ab+b^2}\right]^{\frac{1}{2}} e^{\frac{a^2b^2x_2^2 - 2a^2bx_2x + a^2x^2}{2(a+b)^2(2a^2+2ab+b^2)}} \\
&= \frac{(2\pi)^{-1}}{(2a^2+2ab+b^2)^{\frac{1}{2}}} e^{-\frac{1}{2(a+b)^2}x^2} \\
&\quad \cdot e^{-\frac{1}{2(a+b)^2(2a^2+2ab+b^2)}\left[(a^2+2ab+2b^2)(2a^2+2ab+b^2)x_2^2 - 2b(2a^2+2ab+b^2)x_2x - a^2b^2x2_2 + 2a^2bx_2x - a^2x^2\right]} \\
&= \frac{(2\pi)^{-1}}{(2a^2+2ab+b^2)^{\frac{1}{2}}} e^{-\frac{1}{2(a+b)^2}x^2} e^{\frac{a^2}{2(a+b)^2(2a^2+2ab+b^2)}x^2} \\
&\quad \cdot e^{-\frac{1}{2(a+b)^2(2a^2+2ab+b^2)}\left[2(a+b)^2(a^2+ab+b^2)x_2^2 - 2b(a+b)^2x_2x\right]} \\
&= \frac{(2\pi)^{-1}}{(2a^2+2ab+b^2)^{\frac{1}{2}}} e^{-\frac{1}{2(a+b)^2}x^2} e^{\frac{a^2}{2(a+b)^2(2a^2+2ab+b^2)}x^2} e^{-\frac{a^2+ab+b^2}{2a^2+2ab+b^2}\left[x_2^2 - \frac{bx}{a^2+ab+b^2}x_2\right]},
\end{aligned}
$$

then solve $g(x)$ by solving

$$
\begin{aligned}
g(x) &= \int_{-\infty}^{\infty} g(x_2,x)dx_2 \\
&= \frac{(2\pi)^{-1}}{(2a^2+2ab+b^2)^{\frac{1}{2}}} e^{-\frac{1}{2(a+b)^2}x^2} e^{\frac{a^2}{2(a+b)^2(2a^2+2ab+b^2)}x^2} \int_{-\infty}^{\infty} e^{-\frac{a^2+ab+b^2}{2a^2+2ab+b^2}\left[x_2^2 - \frac{bx}{a^2+ab+b^2}x_2\right]} dx_2.
\end{aligned}
$$

Similarly, we can let part of $g(x)$ be

$$
\begin{aligned}
B &= \int_{-\infty}^{\infty} e^{-\frac{a^2+ab+b^2}{2a^2+2ab+b^2}\left[x_2^2 - \frac{bx}{a^2+ab+b^2}x_2\right]} dx_2 \\
&= \int_{-\infty}^{\infty} e^{-\frac{a^2+ab+b^2}{2a^2+2ab+b^2}\left[(x_2 - \frac{bx}{2(a^2+ab+b^2)})^2 - (\frac{bx}{2(a^2+ab+b^2)})^2\right]} dx_2 \\
&= e^{\frac{b^2x^2}{4(2a^2+2ab+b^2)(a^2+ab+b^2)}} \int_{-\infty}^{\infty} e^{-\frac{(x_2 - \frac{\frac{1}{2}bx}{a^2+ab+b^2})^2}{(2a^2+2ab+b^2)/(a^2+ab+b^2)}} dx_2 \\
&= e^{\frac{b^2x^2}{4(2a^2+2ab+b^2)(a^2+ab+b^2)}} \cdot \left[2\pi \cdot 2(2a^2+2ab+b^2)/(a^2+ab+b^2)\right]^{(-\frac{1}{2})(-1)} \\
&= \left[\frac{4\pi(2a^2+2ab+b^2)}{a^2+ab+b^2}\right]^{\frac{1}{2}} e^{\frac{b^2x^2}{4(2a^2+2ab+b^2)(a^2+ab+b^2)}},
\end{aligned}
$$

then the density function of variable X can be obtained as

$$\begin{aligned}
g(x) &= \frac{(2\pi)^{-1}}{(2a^2+2ab+b^2)^{\frac{1}{2}}} e^{-\frac{1}{2(a+b)^2}x^2} e^{\frac{a^2}{2(a+b)^2(2a^2+2ab+b^2)}x^2} \cdot \left[\frac{4\pi(2a^2+2ab+b^2)}{a^2+ab+b^2}\right]^{\frac{1}{2}} \\
&\quad \cdot e^{\frac{b^2x^2}{4(2a^2+2ab+b^2)(a^2+ab+b^2)}} \\
&= \left[4\pi(a^2+ab+b^2)\right]^{-\frac{1}{2}} \\
&\quad \cdot e^{-\frac{1}{4(a+b)^2(2a^2+2ab+b^2)(a^2+ab+b^2)}\left[2(2a^2+2ab+b^2)(a^2+ab+b^2)x^2-2a^2(a^2+ab+b^2)x^2-b^2(a+b)^2x^2\right]} \\
&= \left[4\pi(a^2+ab+b^2)\right]^{-\frac{1}{2}} e^{-\frac{1}{4(a+b)^2(2a^2+2ab+b^2)(a^2+ab+b^2)}\left[(2a^2+2ab+b^2)(a+b)^2x^2\right]} \\
&= \left[4\pi(a^2+ab+b^2)\right]^{-\frac{1}{2}} e^{-\frac{x^2}{4(a^2+ab+b^2)}}.
\end{aligned}$$

It is shown that $g(x)$ is density function of a normal distribution with $\mu_x = 0$ and $\sigma_x^2 = 2(a^2+ab+b^2)$. It implies $X = aY_1 + bY_2$ follows normal distribution when a and b are not both 0. That is, the joint distribution of Y_1 and Y_2 is a multivariate normal distribution. ■

Above, we have shown that for the normal partition problem considered in Tong (1969), a Multivariate Normal distribution can be derived to calculate the probability of correct decision. Next, we show some results for the case of partitioning two-parameter exponential distributions.

Theorem 2.2.2 *Suppose Z_0, Z_1, Z_2 are independent standard exponential variables and $Y_1 = Z_1 - Z_0$ and $Y_2 = Z_2 - Z_0$, then the joint distribution of Y_1 and Y_2 is not a bivariate Laplace distribution.*

Proof: Since Z_0, Z_1, Z_2 are independent standard exponential variables, Y_1, Y_2 are both following Laplace distribution with location parameter of zero and scale parameter of 1. Then note that to prove the joint distribution of Y_1 and Y_2 is not a bivariate Laplace distribution, we will show that not all linear combinations of Y_1 and Y_2 does follow a

Laplace distribution. That is, we will prove that the distribution of $aY_1 + bY_2$ is not Laplace distribution. Let us define

$$X = aY_1 + bY_2 = aZ_1 + bZ_2 - (a+b)Z_0 \tag{2.10}$$

So $E[X] = aE[Z_1] + bE[X_2] - (a+b)E[Z_0] = 0$. And then denote $X = A - B$, where

$$\begin{aligned} A &= aZ_1 + bZ_2 \\ B &= (a+b)Z_0. \end{aligned}$$

Let us consider the case that a and b are both positive but not equal for simplicity. First, since Z_0, Z_1, Z_2 are independent exponential variables, then the characteristic function of A is

$$\varphi(A) = \varphi(aZ_1) \cdot \varphi(bZ_2) = \frac{1}{1-iat} \cdot \frac{1}{a-ibt}.$$

Since $a \neq b$, then the characteristic function can be expressed as

$$\varphi(A) = \frac{a}{a-b} \cdot \frac{1}{1-iat} + \frac{b}{b-a} \cdot \frac{1}{1-ibt}.$$

By inverse Fourier transform, it is easy to obtain the density function of A and B, which are expressed as

$$\begin{aligned} f(A) &= \tfrac{a}{a-b} \cdot \tfrac{1}{a} e^{-\frac{A}{a}} + \tfrac{b}{b-a} \cdot \tfrac{1}{b} e^{-\frac{A}{b}} = \tfrac{1}{a-b} e^{-\frac{A}{a}} + \tfrac{1}{b-a} e^{-\frac{A}{b}}, \quad 0 \leq A < \infty \\ f(B) &= \tfrac{1}{a+b} e^{-\frac{B}{a+b}}, \quad 0 \leq B < \infty. \end{aligned}$$

Then to solve then cumulative function of X,

$$\begin{aligned} P(X \leq x) &= P(A - B \leq x) \\ &= P(A \leq B + x). \end{aligned}$$

For $x \geq 0$,

$$
\begin{aligned}
P(X \leq x) &= \int_0^\infty \int_0^{B+x} \frac{1}{a-b} e^{-\frac{A}{a}} + \frac{1}{b-a} e^{-\frac{A}{b}} dA \frac{1}{a+b} e^{-\frac{B}{a+b}} dB \\
&= \int_0^\infty \Big[(-\frac{a}{a-b}) e^{-\frac{A}{a}} - \frac{b}{b-a} e^{-\frac{A}{b}} \Big]_0^{B+x} \frac{1}{a+b} e^{-\frac{B}{a+b}} dB \\
&= \int_0^\infty \Big[(-\frac{a}{a-b}) e^{-\frac{B+x}{a}} - \frac{b}{b-a} e^{-\frac{B+x}{b}} + \frac{a}{a-b} + \frac{b}{b-a} \Big] \cdot \frac{1}{a+b} e^{-\frac{B}{a+b}} dB \\
&= \int_0^\infty -\frac{a}{(a+b)(a-b)} e^{-\frac{x}{a}} e^{-\frac{2a+b}{a(a+b)}B} - \frac{b}{(a+b)(b-a)} e^{-\frac{x}{b}} e^{-\frac{a+2b}{b(a+b)}B} + \frac{1}{a+b} e^{-\frac{B}{a+b}} dB \\
&= \Big[\frac{a}{(a+b)(a-b)} \cdot \frac{a(a+b))}{2a+b} e^{-\frac{x}{a}} e^{-\frac{2a+b}{a(a+b)}B} + \frac{b}{(a+b)(b-a)} \\
&\cdot \frac{b(a+b)}{a+2b} e^{-\frac{x}{b}} e^{-\frac{a+2b}{b(a+b)}B} - e^{-\frac{B}{a+b}} \Big]_0^\infty \\
&= \Big[\frac{a^2}{(a-b)(2a+b)} e^{-\frac{x}{a}} e^{-\frac{2a+b}{a(a+b)}B} + \frac{b^2}{(b-a)(a+2b)} e^{-\frac{x}{b}} e^{-\frac{a+2b}{b(a+b)}B} - e^{\frac{B}{a+b}} \Big]_0^\infty \\
&= \frac{-a^2}{(a-b)(2a+b)} e^{-\frac{x}{a}} + \frac{-b^2}{(b-a)(a+2b)} e^{-\frac{x}{b}} + 1.
\end{aligned}
$$

For $x < 0$,

$$
\begin{aligned}
P(X \leq x) &= \int_{-x}^\infty \int_0^{B+x} \frac{1}{a-b} e^{-\frac{A}{a}} + \frac{1}{b-a} e^{-\frac{A}{b}} dA \frac{1}{a+b} e^{-\frac{B}{a+b}} dB \\
&= \Big[\frac{a^2}{(a-b)(2a+b)} e^{-\frac{x}{a}} e^{-\frac{2a+b}{a(a+b)}B} + \frac{b^2}{(b-a)(a+2b)} e^{-\frac{x}{b}} e^{-\frac{a+2b}{b(a+b)}B} - e^{\frac{B}{a+b}} \Big]_{-x}^\infty \\
&= \frac{-a^2}{(a-b)(2a+b)} e^{-\frac{x}{a}} e^{\frac{2a+b}{a(a+b)}x} + \frac{-b^2}{(b-a)(a+2b)} e^{-\frac{x}{b}} e^{\frac{a+2b}{b(a+b)}x} + e^{\frac{x}{a+b}} \\
&= \frac{-a^2}{(a-b)(2a+b)} e^{\frac{x}{a+b}} + \frac{-b^2}{(b-a)(a+2b)} e^{\frac{x}{a+b}} + e^{\frac{x}{a+b}} \\
&= \frac{-a^2(a+2b) + b^2(2a+b) + (a-b)(2a+b)(a+2b)}{(a-b)(2a+b)(a+2b)} e^{\frac{x}{a+b}} \\
&= \frac{(a+b)^2}{(2a+b)(a+2b)} e^{\frac{x}{a+b}}.
\end{aligned}
$$

Combining the results above the cumulative density function of X can be expressed as

$$F(X) = \begin{cases} \frac{-a^2}{(a-b)(2a+b)} e^{-\frac{x}{a}} + \frac{-b^2}{(b-a)(a+2b)} e^{-\frac{x}{b}} + 1 & \text{if} \quad x \geq 0 \\ \frac{(a+b)^2}{(2a+b)(a+2b)} e^{\frac{x}{a+b}} & \text{if} \quad x < 0. \end{cases}$$

Taking the derivative of the cumulative function, the probability density function of X for the case a and b are unequal positive can be obtained as

$$f(X) = \begin{cases} \frac{a}{(a-b)(2a+b)} e^{-\frac{x}{a}} + \frac{b}{(b-a)(a+2b)} e^{-\frac{x}{b}} & \text{if} \quad x \geq 0 \\ \frac{a+b}{(2a+b)(a+2b)} e^{\frac{x}{a+b}} & \text{if} \quad x < 0. \end{cases}$$

To summarize, we have shown that the probability density function of X is not symmetric about 0, which implies $X = aY_1 + bY_2$ does not follow Laplace distribution when a and b are unequal positive. That is, the joint distribution of Y_1 and Y_2 is not a multivariate Laplace distribution. For a, b are negative or equal values, the same conclusion can also be obtained. But here we don't provide the proof for brevity.

■

Remark: One may also consider there are some other multivariate Laplace distributions or define a new multivariate Laplace related to our problem. The author did some work and found out that there are no other Multivariate Laplace densities matching our density; in addition, since our distribution doesn't contain very classic properties and doesn't seem common in practice, the author doesn't try to define as a new joint multivariate Laplace.

2.3 LFC Validation

In the last section it is shown that $f_Y(\cdot)$ is not a multivariate Laplace distribution. In this section, we focus on finding the LFC from symmetric multivariate Laplace distribution. To detect the parametric configuration index r which minimizes (2.7), Monte Carol method is

used to simulate the performance of the partition for various type of setting of parametric configurations. In Tables (2.1)-(2.7), we have summarized the performance of the selection procedure under various parametric configurations for the case when $k = 3, 4, 6, 8$ for different values of σ, a and n. Recall that σ is the common scale parameter of those $k+1$ exponential populations, a is a half length of the indifference zone defined in (2.5) and n is the selection sample size. And we use $(r, k-r)$ to denote that there are r and $k-r$ populations need partitioning which have a location parameter at left boundary and right boundary of indifference zone respectively. For example, when $k = 3$, (3,0) indicates all populations' location parameter are at $\mu_0 + \delta_1$; (2,1) indicates that two populations have a location parameter at $\mu_0 + \delta_1$ and one population have a location parameter at $\mu_0 + \delta_2$.

The simulation studies summarized in the Tables (2.1) and (2.2) conclude that there is no significant difference between the probability of making a correct partition for the cases of $\mu_0 = 0$ and $\mu_0 = 1$ when $k = 3$, and no significant difference for the cases of $\mu_0 = 0$ and $\mu_0 = 5$ when $k = 4$, it is obvious that PCD through partition rule of 2.4 isn't affected by the value of the location parameter of the control population. While it only depends on the length of indifferent zone if the common standard deviation σ is fixed. The the proof is shown in the section (2.5) of simulation. Also, the results presented in the tables (2.1) and (2.2), the PCD is not symmetric about $r = k/2$. For example, (0,3) is very different with (3,1) and (1,2) is very different with (2,1) when $k = 3$; for $k = 4$, (0,4) is very different with (4,0) and (1,3) is very different with (3,1); and PCD does not occur at (2,2).

In addition, the Table (2.3) and (2.4) is consistent with the assumption that all the location parameters should be located on the boundary of the indifference zone to diminish the probability of making correct partition.

And Table (2.1) and (2.2) shows that for $n = 25, 50, 75, 100$, the PCD of configuration (2,1) and (3,1) turns out to be the smallest among all different cases for $k = 3$ and $k = 4$

receptively. Table (2.5) - (2.7) shows the result of $k = 6$ for different σ. It is seen that for $n = 25, 50, 75, 100$, The PCD of configuration (5,1) is the smallest.

Table 2.1: Monte Carlo Simulation Results; Based on 100,000 iterations, $k = 3$

parametric confi.	a	σ	λ	n^*	$P(\mu_0 = 0)$	Std(P)	$P(\mu_0 = 1)$	Std(P)
(0,3)	0.1	1	10	25	0.9388	0.0008	0.9393	0.00076
				50	0.9947	0.00023	0.9949	0.00022
				75	0.9996	0.00006	0.9996	0.00006
				100	0.9999	0.00002	0.9999	0.00002
(1,2)	0.1	1	10	25	0.9048	0.0009	0.9045	0.00093
				50	0.9928	0.00027	0.9925	0.00027
				75	0.9993	0.00008	0.9994	0.00008
				100	0.9999	0.00002	0.9999	0.00003
(2,1)	0.1	1	10	25	0.8778	0.0010	0.8781	0.00103
				50	0.9896	0.00032	0.9899	0.00032
				75	0.9990	0.00010	0.9993	0.00008
				100	0.9999	0.00002	0.9999	0.00003
(3,0)	0.1	1	10	25	0.8835	0.0010	0.8830	0.00102
				50	0.9897	0.00032	0.9900	0.00032
				75	0.9992	0.00009	0.9992	0.00009
				100	0.9999	0.00002	0.9999	0.00002
(0,3)	0.1	2	20	25	0.7843	0.0013	0.7862	0.00130
				50	0.9388	0.00076	0.9373	0.00077
				75	0.9824	0.00042	0.9824	0.00042
				100	0.9948	0.00023	0.9948	0.00023
(1,2)	0.1	2	20	25	0.6710	0.0015	0.6713	0.00149
				50	0.9046	0.00093	0.9048	0.00093
				75	0.9732	0.00051	0.9725	0.00052
				100	0.9926	0.00027	0.9917	0.00029
(2,1)	0.1	2	20	25	0.6037	0.0015	0.6009	0.00155
				50	0.8780	0.00103	0.8785	0.00103
				75	0.9647	0.00058	0.9648	0.00058
				100	0.9898	0.00032	0.9898	0.00032
(3,0)	0.1	2	20	25	0.6451	0.0015	0.6483	0.00151
				50	0.8843	0.00100	0.8851	0.00101
				75	0.9651	0.00058	0.9648	0.00058
				100	0.9900	0.00031	0.9897	0.00032

Table 2.2: Monte Carlo Simulation Results; Based on 100,000 iterations, $k = 4$

configuration	a	σ	λ	n^*	$P(\mu_0 = 0)$	Std(P)	$P(\mu_0 = 5)$	Std(P)
(0,4)	0.5	5	10	25	0.93322	0.00079	0.93569	0.00077
				50	0.99450	0.00023	0.99458	0.00023
				75	0.99959	0.00006	0.99952	0.00007
				100	0.99998	0.00001	0.99997	0.00002
(1,3)	0.5	5	10	25	0.89844	0.00096	0.89840	0.00096
				50	0.99104	0.00029	0.99178	0.00029
				75	0.99926	0.00009	0.99929	0.00008
				100	0.99994	0.00002	0.99991	0.00003
(2,2)	0.5	5	10	25	0.86480	0.00108	0.86651	0.00079
				50	0.98846	0.00033	0.98895	0.00033
				75	0.99883	0.00011	0.99912	0.00009
				100	0.99994	0.00002	0.99989	0.00003
(3,1)	0.5	5	10	25	0.84397	0.00115	0.84102	0.00116
				50	0.98664	0.00036	0.98658	0.00036
				75	0.99880	0.00011	0.99883	0.00011
				100	0.99983	0.00004	0.99986	0.00004
(4,0)	0.5	5	10	25	0.84947	0.00113	0.84893	0.00113
				50	0.98706	0.00036	0.98702	0.00036
				75	0.99901	0.00010	0.99895	0.00010
				100	0.99993	0.00003	0.99996	0.00002
(0,4)	0.5	10	20	25	0.77136	0.00133	0.76967	0.00133
				50	0.93460	0.00078	0.93545	0.00078
				75	0.98103	0.00043	0.98120	0.00043
				100	0.99439	0.00024	0.99459	0.00023
(1,3)	0.5	10	20	25	0.64677	0.00151	0.65191	0.00151
				50	0.89620	0.00097	0.89843	0.00096
				75	0.97078	0.00053	0.97062	0.00053
				100	0.99189	0.00028	0.99115	0.00030
(2,2)	0.5	10	20	25	0.56026	0.00157	0.55991	0.00157
				50	0.86628	0.00108	0.86633	0.00096
				75	0.96094	0.00062	0.96043	0.00062
				100	0.98864	0.00034	0.98889	0.00033
(3,1)	0.5	10	20	25	0.51220	0.00158	0.51370	0.00158
				50	0.84201	0.00115	0.84387	0.00115
				75	0.95215	0.00065	0.95384	0.00066
				100	0.98625	0.00037	0.98666	0.00036
(4,0)	0.5	10	20	25	0.56892	0.00157	0.57356	0.00156
				50	0.84617	0.00114	0.84920	0.00113
				75	0.95308	0.00066	0.95484	0.00066
				100	0.98691	0.00036	0.98672	0.00036

Table 2.3: Monte Carlo Simulation Results of (μ_1, μ_2, μ_3); Based on 100,000 iterations, indifference zone=(10,11), $\sigma = 20$, $k = 3$

n	$\bar{P}\left(Std(\bar{P})\right)$					
	(10,11,11)	(10,11,12)	(10,11,13)	(11,11,11)	(11,11,12)	(11,11,13)
25	0.41335	0.48531	0.48598	0.59729	0.64196	0.64395
	0.00155	0.00158	0.00158	0.00155	0.00151	0.00151
50	0.67283	0.71570	0.72001	0.78557	0.80801	0.80851
	0.00148	0.00142	0.00142	0.00129	0.00124	0.00124
75	0.82113	0.84562	0.84718	0.88515	0.89694	0.89916
	0.00121	0.00114	0.00113	0.00100	0.00096	0.00095
100	0.90534	0.91611	0.91819	0.93853	0.94395	0.94568
	0.00092	0.00087	0.00086	0.00076	0.00072	0.00071
125	0.94865	0.95600	0.95651	0.96666	0.97040	0.97116
	0.00069	0.00064	0.00064	0.00056	0.00053	0.00052

Table 2.4: Monte Carlo Simulation Results of $(\mu_1, \mu_2, \mu_3, \mu_4)$; Based on 100,000 iterations, indifference zone=(10,11), $\sigma = 20$, $k = 4$

n	$\bar{P}\left(Std(\bar{P})\right)$					
	(8,10,10,10)	(9,10,10,10)	(10,10,10,10)	(11,11,11,11)	(11,11,11,12)	(11,11,11,13)
25	0.41906	0.42450	0.36387	0.56953	0.59850	0.60071
	0.00156	0.00156	0.00152	0.00156	0.00155	0.00154
50	0.64654	0.64219	0.56897	0.77152	0.78349	0.78661
	0.00151	0.00151	0.00156	0.00132	0.00130	0.00129
75	0.79178	0.78955	0.73688	0.87818	0.88499	0.88575
	0.00128	0.00128	0.00139	0.00103	0.00101	0.00100
100	0.88452	0.88180	0.84897	0.93566	0.93831	0.93861
	0.00101	0.00102	0.00113	0.00077	0.00076	0.00075
125	0.93632	0.93619	0.91496	0.96433	0.96709	0.96754
	0.00077	0.00077	0.00088	0.00058	0.00056	0.00056
150	0.96531	0.96488	0.95484	0.98114	0.98154	0.98206
	0.00057	0.00058	0.00065	0.00043	0.00042	0.00042

Table 2.5: Monte Carlo Simulation Results; Based on 100,000 iterations, $\sigma = 5$, $k = 6$

configuration	a	σ	λ	$P(Std(P))$			
				$n = 25$	$n = 50$	$n = 75$	$n = 100$
(0,6)	0.5	5	20	0.9299	0.9944	0.9995	0.9999
				0.00081	0.00024	0.00007	0.00002
(1,5)	0.5	5	20	0.8899	0.9910	0.9993	0.9999
				0.00099	0.00030	0.00008	0.00003
(2,4)	0.5	5	20	0.8546	0.9876	0.9991	0.9999
				0.00112	0.00035	0.00010	0.00003
(3,3)	0.5	5	20	0.8261	0.9849	0.9989	0.9999
				0.00120	0.00039	0.00010	0.00004
(4,2)	0.5	5	20	0.7919	0.9828	0.9987	0.9999
				0.00128	0.00041	0.00011	0.00003
(5,1)	0.5	5	20	0.7765	0.9801	0.9983	0.9999
				0.00132	0.00044	0.00013	0.00004
(6,0)	0.5	5	20	0.7859	0.9799	0.9983	0.9999
				0.00130	0.00044	0.00013	0.00003

Table 2.6: Monte Carlo Simulation Results; Based on 100,000 iterations, $\sigma = 10$, $k = 6$

configuration	a	σ	λ	$P(Std(P))$			
				$n = 25$	$n = 50$	$n = 75$	$n = 100$
(0,6)	0.5	10	10	0.7570	0.9314	0.9796	0.9943
				0.00136	0.00080	0.00045	0.00024
(1,5)	0.5	10	10	0.6255	0.8918	0.9685	0.9906
				0.00153	0.00098	0.00055	0.00031
(2,4)	0.5	10	10	0.5259	0.8564	0.9579	0.9875
				0.00158	0.00111	0.00064	0.00035
(3,3)	0.5	10	10	0.4529	0.8216	0.9485	0.9852
				0.00157	0.00121	0.00070	0.00038
(4,2)	0.5	10	10	0.3988	0.7924	0.9383	0.9824
				0.00155	0.00128	0.00076	0.00042
(5,1)	0.5	10	10	0.3780	0.7755	0.9306	0.9799
				0.00153	0.00132	0.00080	0.00044
(6,0)	0.5	10	10	0.4524	0.7841	0.9329	0.9801
				0.00157	0.00130	0.00079	0.00044

Table 2.7: Monte Carlo Simulation Results; Based on 100,000 iterations, $\sigma = 15$, $k = 6$

configuration	a	σ	λ	$\bar{P}(Std(\bar{P}))$			
				$n = 25$	$n = 50$	$n = 75$	$n = 100$
(0,6)	0.5	15	20	0.6316	0.8341	0.9303	0.9698
				0.00153	0.00118	0.00081	0.00054
(1,5)	0.5	15	20	0.4517	0.7492	0.8912	0.9538
				0.00157	0.00137	0.00099	0.00066
(2,4)	0.5	15	20	0.3297	0.6780	0.8563	0.9363
				0.00149	0.00148	0.00111	0.00077
(3,3)	0.5	15	20	0.2565	0.6130	0.8212	0.9202
				0.00138	0.00154	0.00121	0.00086
(4,2)	0.5	15	20	0.2142	0.5703	0.7936	0.9083
				0.00130	0.00157	0.00128	0.00091
(5,1)	0.5	15	20	0.2106	0.5416	0.7762	0.8967
				0.00129	0.00158	0.00132	0.00096
(6,0)	0.5	15	20	0.3226	0.5812	0.7838	0.8975
				0.00148	0.00156	0.00130	0.00096

2.4 Single-stage Procedure

As seen in the previous section, based on Monte Carlo Simulations, we have concluded that *least favorable configuration* is not unique for different k and different α. However, it is shown that LFC is associated with a parametric configuration when each location parameter of the k populations is either $\theta + \delta_1$ or $\theta + \delta_2$. Next, all different configurations under this situation need to be considered and compared for a fixed α to determine the LFC. In this section, a single-stage procedure is introduced to partition k two-parameter exponential populations with a required correction probability satisfied.

Theorem 2.4.1 *Assuming σ is known, the partition problem* (2.2) *has*

$$P\Big[CD|\boldsymbol{\theta}^0(r), \sigma\Big] \geq P^* \tag{2.11}$$

using the partition rule (2.4)*, provided the sample size is at least $n^* = \frac{b\sigma}{a}$. The constant $b = b(k, P^*)$ is a solution of an integral equation* (2.14) *and determined by* (2.15).

Proof: Recall the formulation 2.7, we have

$$P\Big[CD|\boldsymbol{\theta}^0(r),\ \sigma\Big] = P\Big[Z_i - Z_0 < \frac{an}{\sigma},\ Z_0 - Z_j < \frac{an}{\sigma},\ \ 1 \le i \le r,\ \ r+1 \le j \le k\Big].$$

Let us define

$$b = \frac{an}{\sigma}, \tag{2.12}$$

which can be considered as a fixed positive constant if a, n, σ are already determined. If only a and σ are given, in order to reach a required power the value of b can be considered as a critical constant to solve, which can be used to determine the sample size needed to fulfill the partition process.

With (2.12), the probability of making correct decision can be expressed as

$$\begin{aligned}
&P\Big[CD|\boldsymbol{\theta}^0(r),\ \ \sigma\Big] \\
&= P\Big[Z_i - Z_0 < b,\ Z_0 - Z_j < b,\ \ 1 \le i \le r,\ \ r+1 \le j \le k\Big] \\
&= P\Big[Z_i < b + Z_0,\ Z_j > -b + Z_0,\ \ 1 \le i \le r,\ \ r+1 \le j \le k\Big] \\
&= \int_0^\infty P\Big[Z_i < b+z,\ Z_j > -b+z,\ \ 1 \le i \le r,\ \ r+1 \le j \le k\Big] f(z)\ \mathrm{d}z \\
&= \int_0^\infty \Big\{\prod_{i=1}^r P(Z_i < b+z)\Big\}\Big\{\prod_{j=1}^{k-r} P(Z_j > -b+z)\Big\} f(z)\ \mathrm{d}z \\
&= \int_0^b F(b+z)^r f(z) dz + \int_b^\infty F(b+z)^r [1 - F(-b+z)]^{k-r} f(z)\ \mathrm{d}z,
\end{aligned} \tag{2.13}$$

where F(.) and f(.) are respectively the cumulative function(cdf) and probability density function(pdf) of the standard exponential distribution, which are expressed as

$$F(z) = 1 - e^{-z}, \quad f(z) = e^{-z}.$$

The least value of $P\Big[CD|\boldsymbol{\theta}^0(r),\ \ \sigma\Big]$ is attained for some unknown r in the range $r = 0, \ldots, k$. So in order to solve for b values, we perform a numerical computation to determine the value of r and then to solve for b. Equivalently, we have to compute

the critical constant $b^* = b_{k,\alpha}$ by solving $P\Big[CD|\boldsymbol{\theta}^0(r),\ \sigma\Big] = 1-\alpha$ for some value of $r \in (0, 1, \ldots, k)$ which minimize $P(CD)$ given some specified α. Following algorithm was implemented to compute the values of critical constants numerically.

Algorithm:

Step 1: For fixed values of k and α, compute numerically the values of critical constants $b(k, \alpha)$ that satisfy the probability requirement $P_r(CD|\boldsymbol{\theta}^0(r),\ \sigma) \geq 1-\alpha$ for $r = 0, 1, \ldots, k$. In other words, let $b_r(k, \alpha)$ be the solution of the integral equation:

$$\int_0^b F(b+z)^r f(z)\ \mathrm{d}z + \int_b^\infty F(b+z)^r [1-F(-b+z)]^{k-r} f(z)\ \mathrm{d}z = P^* \tag{2.14}$$

when $r = 0, 1, \ldots, k$.

Step 2: Let

$$b^* = \max\left(b_0, b_1, \ldots, b_k\right). \tag{2.15}$$

Then the optimum value of the critical constant satisfying the probability requirement $P\Big[CD|\boldsymbol{\theta}^0(r),\ \sigma\Big] = 1-\alpha$, irrespective of r, is b^*.

The values of critical constant obtained in step 2 above for different values of α are tabulated for different k in table 2.8. Next, it can be shown that probability requirement (2.9) is satisfied if n satisfies that

$$n \geq \frac{b\sigma}{a}. \tag{2.16}$$

The expression for the sample size $n^* = \frac{b\sigma}{a}$, will be defined as the optimal fixed sample size, had σ been known. This completes the proof of the Theorem 1.

■

Along with the introducing of process of single-stage selection and the simulations in section 2.2, here author gives some theoretical proof of LFC for the case when $k = 3$ and $k = 4$. Recall the result of (2.4.13), the probability of making correct decision can be

Table 2.8: Value of constant b satisfying (2.14)

k	P^*						
	0.50	0.75	0.80	0.90	0.95	0.975	0.99
1	0.0000	0.6931	0.9162	1.6094	2.3025	2.9957	3.9120
2	0.6463	1.3755	1.6026	2.3009	2.9953	3.6887	4.6051
3	0.9776	1.7451	1.9793	2.6918	3.3934	4.0905	5.0091
4	1.2135	2.0075	2.2468	2.9693	3.6760	4.3756	5.2958
5	1.3989	2.2126	2.4556	3.1854	3.8956	4.5970	5.5182
6	1.5524	2.3815	2.6273	3.3626	4.0754	4.7781	5.7000
7	1.6837	2.5254	2.7734	3.5128	4.2276	4.9313	5.8538
8	1.7988	2.6507	2.9005	3.6433	4.3597	5.0641	5.9870
9	1.9013	2.7619	3.0131	3.7586	4.4763	5.1813	6.1046
10	1.9937	2.8618	3.1142	3.8620	4.5806	5.2861	6.2097
11	2.0780	2.9525	3.2060	3.9556	4.6751	5.3811	6.3049
12	2.1554	3.0355	3.2899	4.0412	4.7614	5.4677	6.3918
13	2.2271	3.1122	3.3673	4.1200	4.8409	5.5475	6.4717
14	2.2938	3.1833	3.4392	4.1930	4.9145	5.6213	6.5457
15	2.3562	3.2497	3.5062	4.2611	4.9830	5.6901	6.6146
16	2.4149	3.3119	3.5689	4.3248	5.0472	5.7545	6.6791
17	2.4702	3.3705	3.6280	4.3847	5.1075	5.8149	6.7396
18	2.5226	3.4258	3.6837	4.4412	5.1643	5.8719	6.7967
19	2.5723	3.4782	3.7365	4.4947	5.2181	5.9258	6.8507
20	2.6396	3.5280	3.7867	4.5455	5.2691	5.9770	6.9020

expressed as

$$P\Big[CD|\boldsymbol{\theta}^0(r),\ \ \sigma\Big] = \int_0^b F(b+z)^r f(z)dz + \int_b^\infty F(b+z)^r[1-F(-b+z)]^{k-r} f(z)\ \mathrm{d}z.$$

Theorem 2.4.2 *Assuming σ is known, using the partition rule* (2.4) *to deal with the partition problem* (2.2)*, LFC is attached at $r = 2$ for $k = 3$ for any predetermined probability requirement P^*. Here r stands for the number of exponential populations to be partitioned which has a location parameter on the left boundary of the indifference zone.*

Proof: According to Theorem (2.1.1), LFC must be one of the parametric configurations that all the location parameters of the two-parameter exponential populations to be partitioned are on the boundary of the indifference zone. First, PCD for those parametric configurations can be obtained.

$$\begin{aligned} P(CD|r=0) &= \int_0^b (1-e^{-b-z})^0 e^{-z}dz + \int_b^\infty (1-e^{-b-z})^0 (e^{b-z})^3 e^{-z}dz \\ &= [-e^{-z}]_0^b + \int_b^\infty e^{3b}e^{-4z}dz \\ &= 1-e^{-b} + [-\frac{1}{4}e^{3b}e^{-4z}]_b^\infty \\ &= 1-\frac{3}{4}e^{-b}. \end{aligned}$$

$$\begin{aligned} P(CD|r=1) &= \int_0^b (1-e^{-b-z})^1 e^{-z}dz + \int_b^\infty (1-e^{-b-z})^1 (e^{b-z})^2 e^{-z}dz \\ &= 1-\frac{7}{6}e^{-b} + \frac{1}{4}e^{-3b}. \end{aligned}$$

$$\begin{aligned} P(CD|r=2) &= \int_0^b (1-e^{-b-z})^2 e^{-z}dz + \int_b^\infty (1-e^{-b-z})^2 (e^{b-z})^1 e^{-z}dz \\ &= 1-\frac{3}{2}e^{-b} + \frac{1}{3}e^{-2b} + \frac{1}{3}e^{-3b} - \frac{1}{12}e^{-5m}. \end{aligned}$$

$$P(CD|r=3) = \int_0^b (1-e^{-b-z})^3 e^{-z} dz + \int_b^\infty (1-e^{-b-z})^3 (e^{b-z})^0 e^{-z} dz$$
$$= 1 - \frac{3}{2}e^{-b} + e^{-2b} - \frac{1}{4}e^{-3b}.$$

Next we try to compare PCD for different configurations.

$$P(CD|r=0) - P(CD|r=3) = 1 - \frac{3}{4}e^{-b} - (1 - \frac{3}{2}e^{-b} + e^{-2b} - \frac{1}{4}e^{-3m})$$
$$= \frac{3}{4}e^{-b} - e^{-2b} + \frac{1}{4}e^{-3b}$$
$$= \frac{3}{4}e^{-2b}(e^b - 1) - \frac{1}{4}e^{-3b}(e^b - 1)$$
$$= (\frac{3}{4}e^{-2b} - \frac{1}{4}e^{-3b})(e^b - 1) > 0.$$

So $P(CD|r=0) > P(CD|r=3)$.

$$P(CD|r=1) - P(CD|r=2) = 1 - \frac{7}{6}e^{-b} + \frac{1}{4}e^{-3b} - (1 - \frac{3}{2}e^{-b} + \frac{1}{3}e^{-2b} + \frac{1}{3}e^{-3b} - \frac{1}{12}e^{-5b})$$
$$= \frac{1}{3}e^{-b} - \frac{1}{3}e^{-2b} - \frac{1}{12}e^{-3b} + \frac{1}{12}e^{-5b}$$
$$= (\frac{1}{3}e^{-b}(1-e^{-b}) - \frac{1}{12}e^{-3b}(1+e^{-b})(1-e^{-b})$$
$$= (\frac{1}{3}e^{-b} - \frac{1}{12}e^{-3b} - \frac{1}{12}e^{-4b})(1-e^{-b})$$
$$> (\frac{1}{3}e^{-b} - \frac{1}{6}e^{-3b})(1-e^{-b}) > 0.$$

So $P(CD|r=1) > P(CD|r=2)$.

$$P(CD|r=3) - P(CD|r=2) = 1 - \frac{3}{2}e^{-b} + e^{-2b} - \frac{1}{4}e^{-3b} - (1 - \frac{3}{2}e^{-b} + \frac{1}{3}e^{-2b} + \frac{1}{3}e^{-3b} - \frac{1}{12}e^{-5b})$$
$$= \frac{2}{3}e^{-2b} - \frac{7}{12}e^{-3b} + \frac{1}{12}e^{-5b}$$
$$> (\frac{8}{12}e^{-2b} - \frac{7}{12}(e^{-3b}) > 0.$$

So $P(CD|r=3) > P(CD|r=2)$. **Then we can conclude** that under $k=3$, for any P^*,

the LFC can be attained when $r = 2$ since $P(CD|r = 0) > P(CD|r = 1) > P(CD|r = 2) < P(CD|r = 3)$.

■

Theorem 2.4.3 *Assuming σ is known, using the partition rule (2.4) to deal with the partition problem (2.2) for $k = 4$, LFC can be attained at different value of r for different predetermined probability requirement P^*. Here r stands for the number of exponential populations to be partitioned which has a location parameter on the left boundary of the indifference zone.*

Proof: According to Theorem (2.1.1), LFC must be one of the parametric configurations that all the location parameters of the two-parameter exponential populations to be partitioned are on the boundary of the indifference zone. First, PCD for those parametric configurations can be obtained.

$$\begin{aligned} P(CD|r = 0) &= \int_0^b (1 - e^{-b-z})^0 e^{-z} dz + \int_b^\infty (1 - e^{-b-z})^0 (e^{b-z})^4 e^{-z} dz \\ &= 1 - \frac{4}{5} e^{-b}. \end{aligned}$$

$$\begin{aligned} P(CD|r = 1) &= \int_0^b (1 - e^{-b-z})^1 e^{-z} dz + \int_b^\infty (1 - e^{-b-z})^1 (e^{b-z})^3 e^{-z} dz \\ &= 1 - \frac{5}{4} e^{-b} + \frac{3}{10} e^{-3b}. \end{aligned}$$

$$\begin{aligned} P(CD|r = 2) &= \int_0^b (1 - e^{-b-z})^2 e^{-z} dz + \int_b^\infty (1 - e^{-b-z})^2 (e^{b-z})^2 e^{-z} dz \\ &= 1 - \frac{5}{3} e^{-b} + \frac{1}{3} e^{-2b} + \frac{1}{2} e^{-3b} - \frac{2}{15} e^{-5b}. \end{aligned}$$

$$P(CD|r=3) = \int_0^b (1-e^{-b-z})^3 e^{-z} dz + \int_b^\infty (1-e^{-b-z})^3 (e^{b-z})^1 e^{-z} dz$$
$$= 1 - 2e^{-b} + e^{-2b} + \frac{1}{4}e^{-3b} - \frac{1}{4}e^{-5b} + \frac{1}{20}e^{-7b}.$$

$$P(CD|r=4) = \int_0^b (1-e^{-b-z})^4 e^{-z} dz + \int_b^\infty (1-e^{-b-z})^4 (e^{b-z})^0 e^{-z} dz$$
$$= 1 - 2e^{-b} + 2e^{-2b} - e^{-3b} + \frac{1}{5}e^{-4b}.$$

Next we try to compare PCD for different configuration.

$$\begin{aligned} P(CD|r=0) - P(CD|r=1) &= 1 - \frac{4}{5}e^{-b} - (1 - \frac{5}{4}e^{-b} + \frac{3}{10}e^{-3b}) \\ &= \frac{9}{20}e^{-b} - \frac{3}{10}e^{-3b} \\ &> \frac{9}{20}e^{-b} - \frac{9}{20}e^{-3b} > 0 \\ &= (\frac{3}{4}e^{-2b} - \frac{1}{4}e^{-3b})(e^b - 1) > 0. \end{aligned}$$

So $P(CD|r=0) > P(CD|r=1)$.

$$\begin{aligned} P(CD|r=1) - P(CD|r=2) &= 1 - \frac{5}{4}e^{-b} + \frac{3}{10}e^{-3b} - (1 - \frac{5}{3}e^{-b} + \frac{1}{3}e^{-2b} + \frac{1}{2}e^{-3b} - \frac{2}{15}e^{-5b}) \\ &= \frac{5}{12}e^{-b} - \frac{1}{3}e^{-2b} - \frac{1}{5}e^{-3b} + \frac{2}{15}e^{-5b} \\ &= \frac{1}{3}e^{-b} + \frac{1}{12}e^{-b} - \frac{1}{3}e^{-2b} - \frac{1}{5}e^{-3b} + \frac{2}{15}e^{-5b} \\ &= \frac{1}{3}(e^{-2b} - e^{-2b}) + \frac{1}{12}e^{-5b}(e^4 b - \frac{12}{5}e^{2b} + \frac{8}{5}) \\ &= \frac{1}{3}(e^{-2b} - e^{-2b}) + \frac{1}{12}e^{-5b}[(e^{2b} - \frac{6}{5})^2 + \frac{4}{25}] > 0. \end{aligned}$$

So $P(CD|r=1) > P(CD|r=2)$.

$$P(CD|r=2) - P(CD|r=3) = 1 - \frac{5}{3}e^{-b} + \frac{1}{3}e^{-2b} + \frac{1}{2}e^{-3b} - \frac{2}{15}e^{-5b} - (1 - 2e^{-b} + e^{-2b}$$

$$
\begin{aligned}
&+ \frac{1}{4}e^{-3b} - \frac{1}{4}e^{-5b} + \frac{1}{20}e^{-7b}) \\
&= \frac{1}{3}e^{-b} - \frac{2}{3}e^{-2b} + \frac{1}{4}e^{-3b} + \frac{7}{60}e^{-5b} - \frac{1}{20}e^{-7b} \\
&= f(b).
\end{aligned}
$$

It is not easy to compare $P(CD|r=2)$ and $P(CD|r=3)$ by subtraction directly. However, it is easy to calculate $f(0.2) = -0.0062$ and $f(0.3) = 0.0026$. Since $f(m)$ is a continuous function, there must exist a number c between 0.2 and 0.3 such that $f(c) = 0$. Therefore, the sign of $f(b)$ are different for different b values. And we know for different configuration b values depends on the P^*. That implies, for different P^*, the sign of $f(m)$ can be positive or negative, resulting of which the relationship of $P(CD|r=2)$ and $P(CD|r=3)$ can be opposite. After performing amounts of simulations, it is obtained that there is a bound q(about 0.3) for P^* that $P(CD|r=0) > P(CD|r=1)$ when $P^* > q$ and $P(CD|r=0) < P(CD|r=1)$ when $P^* < q$. Continue to compare PCD,

$$
\begin{aligned}
P(CD|r=4) - P(CD|r=3) &= 1 - 2e^{-b} + 2e^{-2b} - e^{-3b} + \frac{1}{5}e^{-4b} - (1 - 2e^{-b} + e^{-2b} \\
&\quad + \frac{1}{4}e^{-3b} - \frac{1}{4}e^{-5b} + \frac{1}{20}e^{-7b}) \\
&= e^{-2b} - \frac{5}{4}e^{-3b} + \frac{1}{5}e^{-4b} + \frac{1}{4}e^{-5b} - \frac{1}{20}e^{-7b} \\
&= (1-e^{-b})^2(e^{-2b} + \frac{3}{4}e^{-3b}) + \frac{7}{10}e^{-4b} - \frac{1}{2}e^{-5b} - \frac{1}{20}e^{-7b} \\
&= (1-e^{-b})^2(e^{-2b} + \frac{3}{4}e^{-3b}) + (\frac{1}{2}e^{-4b} - \frac{1}{2}e^{-5b}) \\
&\quad + (\frac{1}{5}e^{-4b} - \frac{1}{20}e^{-7b}) > 0.
\end{aligned}
$$

So $P(CD|r=4) > P(CD|r=3)$. Combining the results above, we have $P(CD|r=0) > P(CD|r=1) > P(CD|r=2)?P(CD|r=3) < P(CD|r=4)$. **Then we can conclude** that under $k=4$, for different P^*, the LFC can be attained either at $r=2$ or $r=3$.

■

For the case $k > 4$, this kind of performance that LFC is not unique for different P^* also appear. This also shows the goodness of our Single-stage because it guarantees the critical constant b^* obtained comes from the LFC for any given P^* .

2.5 Monte Carlo Simulation Study of Single-stage Procedure

The following steps were taken to simulate the single-stage procedure (2.4.1).

Step 1 Generate random observations X_{il}, $i = 0, 1, \ldots, k,\ l = 1, \ldots, n$

Denote $A = ExpRandom(\sigma)$, which is random observation from exponential distribution with scale parameter σ. Then

$X_{il} = A_{il} + \mu_i$, $i = 0, 1, \ldots, k,\ l = 1, \ldots, n$;

Step 2 Compute T_i, $i = 0, 1, \ldots, k$

$T_0 = Min_{1 \le l \le n} X_{0l} = Min(A_0) + \mu_0$

$T_i = Min_{1 \le l \le r} X_{il} = Min(A_i) + \mu_i, \quad \text{where} \quad i = 0, 1, \ldots, r$

$T_j = Min_{1 \le l \le k-r} X_{jl} = Min(A_j) + \mu_j, \quad \text{where} \quad j = r+1, r+2, \ldots, k$;

Step 3 Calculate Flag

$\text{Flag}_i = 1 \quad \text{if} \quad T_i - T_0 < d,$

$\text{Flag}_i = 0 \quad \text{otherwise.}$

$\text{Flag}_j = 1 \quad \text{if} \quad T_j - T_0 > d,$

$\text{Flag}_j = 0 \quad \text{otherwise}$.

$\text{Flag} = \prod_{i=1}^{k} \text{Flag}_i$.

Flag = 1 indicates a correct partition, Flag = 0 indicates an incorrect partition

Step 4 Run step (1) − (3) N times and calculate $P(CD) = P(\text{Flag} = 1)$

It was shown in Theorem (2.1.1) that in order to attach the infimum of PCD all the location parameters of populations waiting to be partitioned should be located on either $\theta_0 + \delta_1$ or $\theta_0 + \delta_2$. Here the simulation process are also consistent with that conclusion. According to the step 2 and step 3 in the simulation process, we have

$$\begin{aligned} T_i - T_0 &= (Min(A_i) + \mu_i) - (Min(A_0) + \mu_0) \\ &= Min(A_i) - Min(A_0) + \mu_i - \mu_0, \\ T_j - T_0 &= (Min(A_j) + \mu_j) - (Min(A_0) + \mu_0) \\ &= Min(A_i) - Min(A_0) + \mu_j - \mu_0 \end{aligned}$$

In order to compare the different case that location components of the two-parameter exponential variables does and doesn't sit on the boundary of the indifference zone. First the prior case is analyzed. Let's suppose $\mu_i = \mu_0 + \delta_1,\ \mu_j = \mu_0 + \delta_2$, for $i = 1, \ldots, r$ and $j = r+1, \ldots, k$ then in step 3 to check whether $T_i - T_0 < d$ is true, it is equivalent as to check

$$\begin{aligned} Min(A_i) - Min(A_0) + \delta_1 &< \frac{\delta_1 + \delta_2}{2}, \\ Min(A_i) - Min(A_0) &< \frac{\delta_2 - \delta_1}{2} = a. \end{aligned}$$

Similarly to check whether $T_j - T_0 > d$ is true, it is equivalent as to check

$$\begin{aligned} Min(A_i) - Min(A_0) + \delta_2 &> \frac{\delta_1 + \delta_2}{2}, \\ Min(A_i) - Min(A_0) &> \frac{\delta_2 - \delta_1}{2} = -a. \end{aligned}$$

It is easy to infer that the probability of correctly partitioning only depends on σ, a and n, irrespective to the value of μ_0, δ_1 and δ_2. To detect the configuration component r, firstly

σ, a are fixed as some specified values, the PCD then is simulated out for all possible r under different n. By doing this manipulation, it is convenient to see the difference of PCD for different r when PCD is attaching different values due to different n.

Compared to the case above, next a case that not all populations' location parameter are on the boundary of the indifference set is discussed. Here $\mu_i = \mu_0 + \delta_1,\ \mu_j = \mu_0 + \delta_2 + \tau,\ \tau > 0$ for $i = 1, \ldots, r$ and $j = r+1, \ldots, k$ is assumed as the real setting of components. Following the step 3 of simulation process, the statements $T_i - T_0 < d$ and $T_j - T_0 > d$ are checked. As shown above, to check $T_i - T_0 < d$ is still equivalent to check $Min(A_i) - Min(A_0) < a$. However, to check whether $T_j - T_0 > d$ is achieved or not, it is equivalent as to check

$$Min(A_i) - Min(A_0) + \delta_2 + \tau > \frac{\delta_1 + \delta_2}{2},$$
$$Min(A_i) - Min(A_0) > \frac{\delta_2 - \delta_1}{2} - \tau = -a - \tau.$$

Because $-a > -a-\tau$, it is clear that while δ, a and n are all fixed, the PCD for all possible r of the case that not all population location parameter are on the boundary must be larger than PCD of the case that all location components are located on the boundary. In other words, the infimum must be attached under the case that all population needs partitioning should have a location parameter on the boundary of the indifference zone; under infimum of PCD, populations in "Bad Set" have a location component at $\mu_i = \mu_0 + \delta_1$ and populations in "Good Set" have a location component at $\mu_j = \mu_0 + \delta_2$.

In tables (2.9)-(2.12), the performance of single-stage selection procedure is summarized for different P^* and k under different parametric configurations. For preassigned P^*, after obtaining the value of constant b in table (2.8) with k, n^* can be calculated under the setting of σ and a. Next, Monte Carlo simulations are performed by assigning the selection size n^* for different parametric configurations. Respect to Theorem (2.1.1), LFC

is certainly included in those configurations even though it is not known before performing the simulation. It can be noticed that $\bar{P}$s are all above P^* and the corresponding standard deviations are very small, which indicates the probability requirement 2.9 is satisfied for single-stage procedure.

Table 2.9: Simulation of Single-stage Procedure (2.4.1); Based on 100,000 iterations, $P^* = 0.9$ and n^* =50, k=4

configuration	n^*	P	Std(P)	a	σ
(0,4)	50	0.95857	0.00063	0.5938	10
(1,3)	50	0.93611	0.00077	0.5938	10
(2,2)	50	0.91453	0.00088	0.5938	10
(3,1)	50	0.90056	0.00095	0.5938	10
(4,0)	50	0.90258	0.00093	0.5938	20
(0,4)	50	0.96017	0.00093	1.1877	20
(1,3)	50	0.93594	0.00103	1.1877	20
(2,2)	50	0.91582	0.00100	1.1877	20
(3,1)	50	0.90034	0.00076	1.1877	20
(4,0)	50	0.90392	0.00093	1.1877	20

Table 2.10: Simulation of Single-stage Procedure (2.4.1); Based on 100,000 iterations, $P^* = 0.95$ and n^* =50, k=4

configuration	n^*	P	Std(P)	a	σ
(0,4)	50	0.97963	0.00044	0.73521	10
(1,3)	50	0.96472	0.00056	0.73521	10
(2,2)	50	0.95789	0.00063	0.73521	10
(3,1)	50	0.94979	0.00069	0.73521	10
(4,0)	50	0.95069	0.00068	0.73521	20
(0,4)	50	0.97910	0.00045	1.47042	20
(1,3)	50	0.96951	0.00054	1.47042	20
(2,2)	50	0.95683	0.00064	1.47042	20
(3,1)	50	0.95022	0.00068	1.47042	20
(4,0)	50	0.95027	0.00068	1.47042	20

Table 2.11: Simulation of Single-stage Procedure (2.4.1); Based on 100,000 iterations, $P^* = 0.9$ and n^* =50, k=6

configuration	n^*	P	Std(P)	a	σ
(0,6)	50	0.96994	0.00054	0.67252	10
(1,5)	50	0.95430	0.00066	0.67252	10
(2,4)	50	0.93819	0.00076	0.67252	10
(3,3)	50	0.92275	0.00084	0.67252	10
(4,2)	50	0.91155	0.00089	0.67252	10
(5,1)	50	0.90043	0.00095	0.67252	10
(6,0)	50	0.90151	0.00094	0.67252	10
(0,6)	50	0.96924	0.00055	1.34504	10
(1,5)	50	0.95432	0.00066	1.34504	10
(2,4)	50	0.93879	0.00076	1.34504	10
(3,3)	50	0.92175	0.00085	1.34504	10
(4,2)	50	0.90991	0.00090	1.34504	20
(5,1)	50	0.89921	0.00095	1.34504	20
(6,0)	50	0.90216	0.00094	1.34504	20

Table 2.12: Simulation of Single-stage Procedure (2.4.1); Based on 100,000 iterations, $P^* = 0.95$ and n^* =50, k=6

configuration	n^*	P	Std(P)	a	σ
(0,6)	50	0.98549	0.00037	0.81509	10
(1,5)	50	0.97645	0.00048	0.81509	10
(2,4)	50	0.97044	0.00053	0.81509	10
(3,3)	50	0.96169	0.00060	0.81509	10
(4,2)	50	0.95441	0.00066	0.81509	20
(5,1)	50	0.94985	0.00069	0.81509	20
(6,0)	50	0.95049	0.00068	0.81509	20
(0,6)	50	0.98555	0.00037	1.63018	10
(1,5)	50	0.97713	0.00047	1.63018	10
(2,4)	50	0.96851	0.00055	1.63018	10
(3,3)	50	0.96233	0.00060	1.63018	10
(4,2)	50	0.95520	0.00065	1.63018	20
(5,1)	50	0.94972	0.00069	1.63018	20
(6,0)	50	0.95102	0.00068	1.63018	20

3

Multistage Methodology for the partition problem

In the next chapters we will consider two general types of sequential procedures, namely a purely sequential and a two-stage procedures to partition a set of two-parameter exponential populations based on comparisons with a control population. First the purely sequential procedure is derived and then its theoretical first-order and second-order asymptotic properties are derived. These properties are studied via Monte Carlo simulations and the efficiency of the proposed procedure is studied.

3.1 Purely Sequential Selection

In this section, we will construct a purely sequential procedure along the lines of Mukhopadhyay and Solanky (1994). One may also see Robbins et al. (1968), and Robbins (1959) to review a brief history of the purely sequential procedures. Recall that $n^* = \frac{b\sigma}{a}$, is the optimal fixed sample size required from each population, had σ been known. As before, based on the sample size n, Let X_{ij} denote the jth observation from the ith population, with the density function as below

$$f_X(x) = \sigma^{-1} exp\{-(x_{ij} - \theta_i)/\sigma\} I(x_{ij} > \theta_i). \tag{3.1}$$

Let us define the following

$$\begin{aligned} T_i &= Min_{1\le j\le n}\{X_{i1},...,X_{in}\}, \\ V_i &= (n-1)^{-1}\sum_{j=1}^{n}(X_{ij}-T_i), \\ \hat{\sigma} &= \sum_{i=0}^{k}\sum_{j=1}^{n}(X_{ij}-T_i)/((k+1)(n-1)). \end{aligned} \tag{3.2}$$

For the case σ is unknown, the basic idea is to keep updating estimates of σ until we arrive at some kind of "stability", that is the sample size n and the corresponding estimator $\hat{\sigma}$ of σ satisfies a certain relationship which is to be made precise very shortly.

We start with $m(\geq 2)$ observations from each of the k populations to be partitioned and the control population. Then keep taking one more observation at a time from each population according to the following stopping rule:

$$N = N(a) = \inf\left\{n \geq m : n \geq \frac{b\hat{\sigma}}{a}\right\}. \tag{3.3}$$

The constant b in above equation comes from (2.15). Next we will shown that the purely sequential procedure (3.3) is a well defined rule which will terminate with probability 1.

Theorem 3.1.1 *The purely sequential procedure* (3.3) *terminates with probability 1.*

Proof: Fox fixed $\underline{\theta}, \sigma, m$ and a, we have $P(N < \infty) = 1 - \lim_{n\to\infty} P(N > n) \geq 1 - \lim_{n\to\infty} P(n < \frac{b\hat{\sigma}}{a}) = 1$, since $\hat{\sigma} \to \sigma$ w.p. 1 as $n \to \infty$. That is, $P(N < \infty) = 1$, in other words the purely sequential procedure (3.3) terminates with probability one.

■

Based on the totality of all samples, that is having $X_{i1}, X_{i2}, \ldots, X_{ik}$ from $\pi_i, i = 1, \ldots, k$, we now implements the customary selection procedure P_N given by (2.4). Throughout, $\underline{\theta}$ and σ are kept fixed. In next, we derive some theoretical properties of purely sequential procedure.

3.2 Asymptotic Properties of Purely Sequential Procedure

In this section, we derive some essential properties for the purely sequential procedure as defined in (3.3). The derived properties show the efficiency of the proposed purely sequential procedure (3.3). We will also prove the probability of correct decision of the procedure is guaranteed.

Theorem 3.2.1 *For the purely sequential procedure* (3.3)*, we have as* $a \to 0$*:*

(i) $N/n^* \to 1$ w.p. *1;*

(ii) $E\left(N/n^*\right) \to 1$*;*

(iii) $n^{*\frac{1}{2}}\left(N - n^*\right) \to N\left(0, 1/\left(k+1\right)\right)$*;*

(iv) $\liminf P\left(CD\right) \geq P^*$ under the LFC*;*

where $n^* = b\sigma/a$ and the constant b comes from (2.14) and determined by (2.15).

Proof: Utilizing the Lemma 1 of Chow and Robbins (1965), it follows that as $a \to 0$, we have $N \to \infty$, w.p. 1, $\hat{\sigma}_N \to \sigma$ w.p. 1, and $\hat{\sigma}_{N-1} \to \sigma$ w.p. 1. Note that the purely sequential procedure (3.3) can be expressed as $N = N_\nu = \inf\left\{n \geq m : n > \Psi_\nu T_n\right\}$, with $\Psi_\nu = \frac{b}{a}$, $T_n = \hat{\sigma}$ and the basic inequality (equation 2.4.3 in Mukhopadhyay and Solanky (1994) simplifies to

$$\frac{b}{a}\hat{\sigma}_N \leq N \leq m + \frac{b}{a}\hat{\sigma}_{N-1}. \tag{3.4}$$

Dividing throughout (3.4) by n^* and taking limits as $a \to 0$ leads to the part (i). Next,

invoke the Helmert's orthogonal transformation, the $\hat{\sigma_N}$ can be expressed as

$$\begin{aligned}\hat{\sigma}_N &= \textstyle\sum_{i=0}^{k}\sum_{j=1}^{n}(X_{ij}-T_i)/\left((k+1)(n-1)\right)\\ &= (n-1)^{-1}\textstyle\sum_{i=1}^{n-1}W_i'\\ &= \tfrac{1}{2}\sigma(k+1)^{-1}(n-1)^{-1}\textstyle\sum_{i=1}^{n-1}\chi^2_{2(k+1)},\end{aligned} \tag{3.5}$$

where W_i's are i.i.d. $\frac{1}{2}\sigma(k+1)^{-1}(n-1)^{-1}\sum_{i=1}^{n-1}\chi^2_{2(k+1)}$ distributed random variables.

Let $W^* = \sup_{n\geq 2}\left\{(n-1)^{-1}\sum_{i=1}^{n-1}W'\right\}$. From the right hand side of the basic inequality (3.4), we can write $N \leq m+a^{-1}b\sigma_{N-1}$ as $N \leq m+a^{-1}bW^*$. That is, $Nn^{*-1} \leq m+\sigma^{-1}W^*$ for sufficiently small values of a such that n^{*-1} becomes smaller than unity. By Wiener's (1939) dominated ergodic theorem one concludes that $E(W^*) < \infty$. Now, the dominated convergence theorem and part (i) together imply part (ii) $E(N/n^*) \to 1$.

According to (3.5), it can be obtained the following

$$\begin{aligned}E(\hat{\sigma}_N) &= \sigma,\\ Var(\hat{\sigma}_N) &= \frac{1}{4}\sigma^2(k+1)^{-2}(N-1)^{-2}\cdot(N-1)\cdot 4(k+1)\\ &= \sigma^2(k+1)^{-1}(N-1)^{-1},\\ Std(\hat{\sigma}_N) &= \sigma(k+1)^{-\frac{1}{2}}(N-1)^{-\frac{1}{2}}.\end{aligned}$$

Then, one may easily obtain $\frac{N^{\frac{1}{2}}(\sigma_N-\sigma)}{\sigma(k+1)^{-\frac{1}{2}}}$ and $\frac{N^{\frac{1}{2}}(\sigma_{N-1}-\sigma)}{\sigma(k+1)^{-\frac{1}{2}}}$ both converge to $N(0,1)$ as $a \to 0$. Next, along the lines of Theorem 2.4.1 in Mukhopadhyay and Solanky (1994), we can let $\Psi_\nu = \frac{b}{a}$. Let $a' = \sigma$, $b' = (k+1)^{\frac{1}{2}}\sigma$ and then obtain $\frac{a'^{\frac{1}{2}}(N_\nu - a'\Psi_\nu)}{b'\Psi^{\frac{1}{2}}} \to N(0,1)$. The results can be expressed as:

$$\begin{aligned}&\tfrac{\sigma^{\frac{1}{2}}(N-n^*)}{(k+1)^{-\frac{1}{2}}\sigma(b/a)^{\frac{1}{2}}} \to N(0,1),\\ &\tfrac{N-n^*}{(k+1)^{-\frac{1}{2}}n^{*\frac{1}{2}}} \to N(0,1),\\ &n^{*-\frac{1}{2}}(N-n^*) \to N(0,\tfrac{1}{k+1}).\end{aligned}$$

Then part (iii) is proved. Next, to prove the part (iv), first note that the two properties of the LFC still hold. That is, for any N, the LFC is attached when all the populations partitioning have a location parameter either at $\mu_0 + \delta_1$ or $\mu_0 + \delta_2$. Then along with following the steps of Theorem 2.4.1, the $P(CD)$ based on the sample of size N can be written as

$$
\begin{aligned}
& P\Big[CD|\boldsymbol{\theta}^0(r),\ \ \sigma\Big] \\
& = P\Big[T_i - T_0 < d,\ T_j - T_0 > d,\ \ 0 < i \leq r,\ \ r < j \leq k|\boldsymbol{\theta}^0(r),\ \ \sigma\Big] \\
& = P\Big[\tfrac{T_i-\theta_i}{\sigma/N} - \tfrac{T_0-\theta_0}{\sigma/N} < \tfrac{d-\theta_i+\theta_0}{\sigma/N},\ \tfrac{T_j-\theta_j}{\sigma/N} - \tfrac{T_0-\theta_0}{\sigma/N} > \tfrac{d-\theta_j+\theta_0}{\sigma/N},\ \ 1 \leq i \leq r,\ \ r+1 \leq j \leq k\Big] \\
& = P\Big[Z_i - Z_0 < \tfrac{d-\delta_1}{\sigma/N},\ Z_j - Z_0 > \tfrac{d-\delta_2}{\sigma/N},\ \ 1 \leq i \leq r,\ \ r+1 \leq j \leq k\Big] \qquad (3.6) \\
& = P\Big[Z_i - Z_0 < \tfrac{a}{\sigma/N},\ Z_j - Z_0 > -\tfrac{a}{\sigma/N},\ \ 1 \leq i \leq r,\ \ r+1 \leq j \leq k\Big] \\
& = P\Big[Z_i - Z_0 < \tfrac{aN}{\sigma},\ Z_j - Z_0 > -\tfrac{aN}{\sigma},\ \ 1 \leq i \leq r,\ \ r+1 \leq j \leq k\Big] \\
& = E\Big[\textstyle\int_0^{b'} F(b'+z)^r f(z)\ \mathrm{d}z + \int_{b'}^{\infty} F(b'+z)^r [1 - F(-b'+z)]^{k-r} f(z)\ \mathrm{d}z\Big],
\end{aligned}
$$

where F(.) and f(.) are respectively the cumulative function(cdf) and probability density function(pdf) of the standard exponential distribution as $F(z) = 1 - e^{-z}$ and $f(z) = e^{-z}$ and $b' = \frac{aN}{\sigma}$.

Also, from part (i), one get $aN\sigma^{-1} \to b$ w.p.1 as $a \to 0$, and hence (3.6) together with the dominated convergence theorem will lead to part (iv). This complete the proof of Theorem 2. ■

Next, for the purely sequential procedure (3.3) we will derive a second-order expansion to determine the amount of over-sampling the procedure does asymptotically. The amount of oversampling β is defined below and also tabulated for the practitioners. We will also compare the validity of the asymptotic expression β for the small and moderate sample sizes.

Theorem 3.2.2 *For the purely sequential procedure* (3.3)*, we have as* $a \to 0$:

(i) $E(N) = n^* + \beta + o(1)$ for all $\theta \in \Omega(a)$ if $m \geq 2$ when $k \geq 2$;

(ii) $P(CD) = P^* + \frac{1}{n^*}\{b\beta g'(b) + \frac{b^2}{2(k+1)} g''(b)\} + \circ\left(\frac{1}{n^*}\right)$ under the LFC for $m \geq 2$ when $k \geq 2$.

where where $n^* = b\sigma/a$ and the constant b comes from (2.14) and tabulated in Table (2.8). $\beta = (k+1)^{-1}(\frac{1}{2}\tau - 1)$ with ν defined in (3.8). $g(x)$ is define in (3.9), $g'(x)$ and $g''(x)$ are defined in (3.11) and (3.12). The values of the constant $\beta = \beta(k)$ are provided in Table (3.1).

Proof: First note that using (3.5), we can rewrite $\hat{\sigma} = (n-1)^{-1} \sum_{i=1}^{n-1} W_i'$, where W_1', $W_2', \cdots$ are i.i.d. $\frac{1}{2}(k+1)^{-1} \sigma \chi^2_{2(k+1)}$ random variables. Let's write $W_i = 2(k+1)\sigma^{-1} W_i'$, with W_i being i.i.d. $\chi^2_{2(k+1)}$ random variables. Using this the purely sequential procedure could be rewritten as

$$
\begin{aligned}
N &= \inf\left\{n \geq m : n \geq \tfrac{b\hat{\sigma}}{a}\right\}, \\
N &= \inf\left\{n \geq m : n n^{*-1} \geq \sigma^{-1}(n-1)^{-1} \textstyle\sum_{i=1}^{n-1} W_i'\right\}, \\
N &= \inf\left\{n \geq m : 2(k+1) n^{*-1} n(n-1) \geq \textstyle\sum_{i=1}^{n-1} W_i\right\}.
\end{aligned}
$$

Note that $N = Q + 1$ where

$$
Q = \inf\left\{n \geq m - 1 : 2(k+1) n^{*-1} n^2 \left(n^{-1} + 1\right) \geq \sum\nolimits_{i=1}^{n} W_i\right\}. \tag{3.7}
$$

The stopping variable Q is of the form of Mukhopadhyay and Solanky (1994)s equation (2.4.7) with $\delta = 2$, $L_0 = 1$, $h^* = \frac{2(k+1)}{n^*}$, $\theta = E(W_i) = 2(k+1)$, $\tau^2 = E(W_i^2) - \theta^2 = 4(k+1)$, $\beta^* = (\delta - 1)^{-1} = 1$, $n_0^* = (\theta/h^*)^{\beta^*} = n^*$, $P = \beta^{*2}\tau^2\theta^{-2} = (k+1)^{-1}$, $b = k+1$, and ν is defined in the equation (2.4.9) in Mukhopadhyay and Solanky (1994) as

$$
\nu = \nu(k) = k + 2 - \sum\nolimits_{n=1}^{\infty} n^{-1} E\left[\max\left\{0, \chi^2_{2n(k+1)} - 4n(k+1)\right\}\right]. \tag{3.8}
$$

Next, note that the constant η as defined in the equation (2.4.10) in Mukhopadhyay and Solanky (1994) simplifies to

$$\begin{aligned}\eta &= \beta^*\theta^{-1}\nu - \beta^* L_0 - \frac{1}{2}\delta\beta^{*2}r^2\theta^{-2} \\ &= \frac{1}{2\,(k+1)}\nu - 1 - \frac{4\,(k+1)}{4\,(k+1)^2} \\ &= (k+1)^{-1}\,(\frac{1}{2}\nu - 1) - 1.\end{aligned}$$

Using the Theorem 2.4.8(v) of Mukhopadhyay and Solanky (1994) with $w = 1$ leads to

$$E\,(Q/n_0^*) = 1 + \eta n_0^{*-1} + o(n_0^{*-1})$$

Then it can be obtained that

$$\begin{aligned}E\,(N) &= E\,(Q) + 1 \\ &= 1 + n_0^* + \eta + \circ\,(1) \\ &= n^* + (k+1)^{-1}\,(\frac{1}{2}\nu - 1) + \circ\,(1)\,,\end{aligned}$$

if $m - 1 > (k+1)^{-1}$, that is, if $m > 1 + (k+1)^{-1}$. This is part (i). And $\beta = (k+1)^{-1}(\frac{1}{2}\tau - 1)$

In order to prove part (ii), let us denote a function g(x) as

$$g(x) = \int_0^x F(x+z)^r f(z)\,\mathrm{d}z + \int_x^\infty F(x+z)^r[1 - F(-x+z)]^{k-r} f(z)\,\mathrm{d}z. \qquad (3.9)$$

where F(.) and f(.) are respectively the cumulative function(cdf) and probability density function(pdf) of the standard exponential distribution; k is the number of populations to be partitioned and r is the value for LFC to be attained. Then we should have $E[g(\frac{aN}{\sigma})] = P$ and $g(b) = P^*$, where b comes from (2.14) and are tabulated in Table (2.8).

By expanding the series of F(.) and f(.), the expression function g(x) can be transformed to

$$
\begin{aligned}
&g(x) = \textstyle\int_0^x F(z+x)^r f(z)\,\mathrm{d}z + \int_x^\infty F(z+x)^r[1-F(z-x)]^{k-r} f(z)\,\mathrm{d}z \\
&= \textstyle\int_0^x (1-e^{-z-x})^r e^{-z}\,\mathrm{d}z + \int_x^\infty (1-e^{-z-x})^r (e^{-z+x})^{k-r} e^{-z}\,\mathrm{d}z \\
&= \textstyle\int_0^x \sum_{i=0}^r \binom{r}{i}(-1)^i e^{-iz-ix} e^{-z}\,\mathrm{d}z + \int_x^\infty \sum_{i=0}^r \binom{r}{i}(-1)^i e^{-iz-ix} e^{-(k-r)z+(k-r)x} e^{-z}\,\mathrm{d}z \\
&= \textstyle\sum_{i=0}^r \binom{r}{i}(-1)^i e^{-ix} \int_0^x e^{-(i+1)z}\,\mathrm{d}z + \sum_{i=0}^r \binom{r}{i}(-1)^i e^{(k-r-i)x} \int_x^\infty e^{-(k-r+1+i)z}\,\mathrm{d}z \qquad (3.10) \\
&= \textstyle\sum_{i=0}^r \binom{r}{i}(-1)^i (i+1)^{-1} e^{-ix}(1-e^{-(i+1)x}) + \sum_{i=0}^r \binom{r}{i}(-1)^i (k-r+1+i)^{-1} \\
&\quad e^{(k-r-i)x} e^{-(k-r+1+i)x} \\
&= \textstyle\sum_{i=0}^r \binom{r}{i}(-1)^i (i+1)^{-1}[e^{-ix} - e^{-(2i+1)x}] + \sum_{i=0}^r \binom{r}{i}(-1)^i (k-r+1+i)^{-1} e^{-(2i+1)x}.
\end{aligned}
$$

Then the first and second derivative of the function g(x) can be obtained as

$$
\begin{aligned}
g'(x) &= \textstyle\sum_{i=0}^r \binom{r}{i}(-1)^i (i+1)^{-1}[-ie^{-ix} + (2i+1)e^{-(2i+1)x}] \\
&\quad + \textstyle\sum_{i=0}^r \binom{r}{i}(-1)^i (k-r+1+i)^{-1}(-2i+1)e^{-(2i+1)x}.
\end{aligned} \qquad (3.11)
$$

$$
\begin{aligned}
g''(x) &= \textstyle\sum_{i=0}^r \binom{r}{i}(-1)^i (i+1)^{-1}[i^2 e^{-ix} - (2i+1)^2 e^{-(2i+1)x}] \\
&\quad + \textstyle\sum_{i=0}^r \binom{r}{i}(-1)^i (k-r+1+i)^{-1}(-2i+1)^2 e^{-(2i+1)x}.
\end{aligned} \qquad (3.12)
$$

Expanding g(x) at $x = b$, we can have

$$
g(x) = g(b) + g'(b)(x-b) + g''(Z(x))\frac{(x-b)^2}{2}, \qquad (3.13)
$$

where $Z(x)$ is a positive random variable such that $min(x,b) \le Z(x) \le max(x,b)$. Then we can have the following

$$
\begin{aligned}
E[g(\tfrac{aN}{\sigma})] &= E[g(b\tfrac{N}{n^*})] = g(b) + g'(b)E[(b\tfrac{N}{n^*} - b)] + \tfrac{1}{2}E[g''(Z(x))(b\tfrac{N}{n^*} - b)^2] \\
&= g(b) + \tfrac{bg'(b)}{n^*}E[(N-n^*)] + \tfrac{b^2}{2n^*}E[g''(Z(x))U^*].
\end{aligned} \qquad (3.14)
$$

Recall that $g(b) = P^*$ and $E(N) = n^* + (k+1)^{-1}\left(\frac{1}{2}\nu - 1\right) + \circ(1)$. And according to the Theorem 2.4.8(iv) of Mukhopadhyay and Solanky (1994) with $w = 1$ it can be obtained that $g''(Z(x))U^{*2}$ is uniformly integrable if $m - 1 > (k+1)^{-1}$. Also, $g''(Z(x))U^{*2} \to (k+1)^{-1}g''(b)\chi_1^2$ as $a \to 0$. Then, it can be obtained that for $m > 1 + (k+1)^{-1}$, that is, for $m \geq 2$,

$$\begin{aligned} P(CD) &= P^* + \tfrac{bg'(b)}{n^*}(\beta + \circ(1)) + \tfrac{b^2}{2n^*}((k+1)^{-1}g''(b) + \circ(1)) \\ &= P^* + \tfrac{1}{n^*}\{b\beta g'(b) + \tfrac{b^2}{2(k+1)}g''(b)\} + \circ\left(\tfrac{1}{n^*}\right). \end{aligned}$$

It should be noted in part(ii) of Theorem (3.2.2), the function $g(x)$ is defined under the least favorable configuration. The procedure (3.3) doesn't provide when LFC is attained for different value of k. The theoretical result is provided in case statisticians need to consider it for specified situations.

■

3.3 Monte Carlo Simulation Study of Purely Sequential Procedure

The purely sequential procedure (3.3), starts with $m \geq 2$ observations from each of the k populations and the control population. The procedure takes one additional sample at a time from each all the k populations and the control population until stopping rule (3.3) is satisfied. In this section, some simulation are performed under different setting of n^*, b, a and σ. In table (3.2-3.4), we took the value of $m = 5$ and $m = 10$, $k = 4$, and $P^* = 0.9$, $p^* = 0.95$. Respect to the value of k and P^*, the value of the design constant b was obtained from (2.8). Recall that the design parameters δ_1 and δ_2 are predetermined by the experimenter based on the definition of the Good and Bad populations. Note that

Table 3.1: The values of the constant β and ν as defined in Theorem 3.2.2

k	ν	β	k	ν	β
1	2.62634	0.17158	21	22.99967	0.47727
2	3.79489	0.29915	22	23.99976	0.47826
3	4.86155	0.35769	23	24.99983	0.47916
4	5.90474	0.39047	24	25.99987	0.48000
5	6.93362	0.41113	25	26.99991	0.48077
6	7.95334	0.42524	26	27.99993	0.48148
7	8.96699	0.43544	27	28.99995	0.48214
8	9.97653	0.44314	28	29.99997	0.48276
9	10.98326	0.44916	29	30.99997	0.48333
10	11.98803	0.45400	30	31.99998	0.48387
11	12.99141	0.45798	31	33.00000	0.48437
12	13.99383	0.46130	32	34.00000	0.48485
13	14.99556	0.46413	33	35.00000	0.48529
14	15.99680	0.46656	34	36.00000	0.48571
15	16.99770	0.46868	35	37.00000	0.48611
16	17.99834	0.47054	36	38.00000	0.48649
17	18.99880	0.47219	37	39.00000	0.48684
18	19.99913	0.47366	38	40.00000	0.48718
19	20.99937	0.47498	39	41.00000	0.48750
20	21.99955	0.47618	40	42.00000	0.48780

$a = \delta_2 - \delta_1$, which represents the length of indifference zone.

Table 3.2: Simulation of The purely Sequential Procedure (3.3); Based on 100,000 iterations, p^* =0.9, k=4, m =5

configuration	n^*	$\bar{n}$	Std(n)	P	Std(P)	b	a	σ	λ
(0,4)	29.6939	30.0798	0.0079	0.9588	0.0006	2.9693	0.5	5	10
(1,3)	29.6939	30.0661	0.0079	0.9359	0.0007	2.9693	0.5	5	10
(2,2)	29.6939	30.0582	0.0079	0.9166	0.0008	2.9693	0.5	5	10
(3,1)	29.6939	30.0764	0.0079	0.9020	0.0009	2.9693	0.5	5	10
(4,0)	29.6939	30.0783	0.0079	0.9035	0.0009	2.9693	0.5	5	10
(0,4)	59.3878	59.7425	0.0347	0.9549	0.0020	2.9693	0.5	10	20
(1,3)	59.3878	59.7787	0.0349	0.9356	0.0024	2.9693	0.5	10	20
(2,2)	59.3878	59.7613	0.0351	0.9152	0.0027	2.9693	0.5	10	20
(3,1)	59.3878	59.7628	0.0346	0.9012	0.0029	2.9693	0.5	10	20
(4,0)	59.3878	59.7767	0.0347	0.9019	0.0029	2.9693	0.5	10	20

Table 3.3: Simulation of The purely Sequential Procedure (3.3); Based on 100,000 iterations, p^* =0.95, k=4, m =5

configuration	n^*	$\bar{n}$	Std(n)	P	Std(P)	b	a	σ	λ
(0,4)	36.7605	37.1282	0.0087	0.9793	0.0004	3.6760	0.5	5	10
(1,3)	36.7605	37.1429	0.0088	0.9670	0.0005	3.6760	0.5	5	10
(2,2)	36.7605	37.1488	0.0087	0.9582	0.0006	3.6760	0.5	5	10
(3,1)	36.7605	37.1394	0.0087	0.9502	0.0006	3.6760	0.5	5	10
(4,0)	36.7605	37.1379	0.0088	0.9503	0.0006	3.6760	0.5	5	10
(0,4)	73.5210	73.8684	0.0392	0.9791	0.0014	3.6760	0.5	10	20
(1,3)	73.5210	73.9263	0.0393	0.9660	0.0018	3.6760	0.5	10	20
(2,2)	73.5210	73.8798	0.0389	0.9580	0.0020	3.6760	0.5	10	20
(3,1)	73.5210	73.9342	0.0388	0.9511	0.0021	3.6760	0.5	10	20
(4,0)	73.5210	73.9369	0.0390	0.9513	0.0021	3.6760	0.5	10	20

Table 3.4: Simulation of The purely Sequential Procedure (3.3); Based on 100,000 iterations, p^* =0.99, k=4, m =5

configuration	n^*	$\bar{n}$	Std(n)	$\bar{P}$	Std($\bar{P}$)	b	a	σ	λ
(0,4)	52.9580	53.3041	0.0333	0.9951	0.0007	5.2958	0.5	5	10
(1,3)	52.9580	53.3730	0.0329	0.9937	0.0008	5.2958	0.5	5	10
(2,2)	52.9580	53.2533	0.0335	0.9908	0.0009	5.2958	0.5	5	10
(3,1)	52.9580	53.3993	0.0332	0.9900	0.0010	5.2958	0.5	5	10
(4,0)	52.9580	53.3478	0.0330	0.9906	0.0010	5.2958	0.5	5	10
(0,4)	105.9160	106.3808	0.0463	0.9957	0.0006	5.2958	0.5	10	20
(1,3)	105.9160	106.2043	0.0458	0.9954	0.0007	5.2958	0.5	10	20
(2,2)	105.9160	106.3197	0.0461	0.9914	0.0009	5.2958	0.5	10	20
(3,1)	105.9160	106.2876	0.0465	0.9510	0.0009	5.2958	0.5	10	20
(4,0)	105.9160	106.2534	0.0465	0.9515	0.0009	5.2958	0.5	10	20

4

Two-stage Selection Procedure

In this chapter a two-stage procedures to partition a set of two-parameter exponential populations based on comparisons with a control population will be considered and compared with the purely sequential procedure derived in last chapter. First the two-stage procedure is derived and then its theoretical first-order and second-order asymptotic properties are derived. These properties are studied via Monte Carlo simulations and the efficiency of the proposed procedure is studied.

4.1 Two-stage Selection

In this section, we will propose a two-stage procedure for the partition problem introduced in the chapter 2 for the unknown σ case. The two-stage procedures are operationally more convenient to implement than the purely sequential procedures. Unlike for the purely sequential procedure, in which the experimenter has to decide whether or not to continue sampling after adding each observation, in the two-stage procedure the sample size is determined only once. Meaning, the experimenter would select a small pilot sample and then based on that pilot sample it is determined how many additional samples need to be collected. This feature of the two-stage procedure makes it more user friendly and operationally convenient. For more literature on the two-stage procedures, the reader is recommended to Solanky (2006) and Mukhopadhyay and Solanky (1994).

Next, we describe the a two-stage procedure for the partition problem introduced in the chapter 2.

Stage I. Let $m(\geq 2)$ denote the common starting sample size from k treatments and the control group. The procedure begins by taking a sample X_{ij} from $\pi_i; i = 0, \ldots, k; j = 1, \ldots, m$. Let

$$\begin{aligned} T_i &= Min_{1\leq j\leq m}\{X_{i1}, ..., X_{im}\}, \\ V_{im} &= (m-1)^{-1}\sum_{j=1}^{m}(X_{ij} - T_i), \\ \hat{\sigma}_m &= \sum_{i=0}^{k}\sum_{j=1}^{m}(X_{ij} - T_i)/((k+1)(m-1)). \end{aligned}$$

Note that $\hat{\sigma}_m$ is the pooled estimator of σ.

Stage II. Let us define

$$N = N(a) = max\left\{m, \left[\frac{h\hat{\sigma}_m}{a}\right]\right\}. \tag{4.1}$$

The constant $h = h(k; m; P^*)$ is a positive constant defined in (4.4) and $[x]$ denotes the smallest integer greater than or equal to x. If $N = m$, we do not take any more samples from each population. However, if $N > m$, then we sample the difference in the second stage, that is, we take $(N - m)$ more samples from each π_i, $i = 0, \cdots, k$ in the second stage. Based on the totality of all samples, that is having $X_{i1}, \cdots, X_{iN}$ from π_i, we now implements the selection procedure $\wp$ given by (2.4). Throughout, $\underline{\theta}$ and σ are kept fixed.

Fox fixed $\underline{\theta}, \sigma, m$ and a, we have $P(N < \infty) = 1 - \lim_{n\to\infty} P(N > n) \geq 1 - \lim_{n\to\infty} P(n < \frac{h\hat{\sigma}_m}{a}) = 1$. That is, $P(N < \infty) = 1$, in other words the two-stage procedure (4.1) terminates with probability one.In next, we derive some theoretical properties of the two-stage procedure.

4.2 Asymptotic Properties of Two-stage Procedure

Theorem 4.2.1 *For the two-stage procedure* (4.1)*, with h defined in* (4.4)*, we have:*

(i) $P\,(CD) \geq P^*$ *for all* $\underline{\theta} \in \Omega(a)$*;*

(ii) $h\sigma a^{-1} \leq E(N) \leq m + h\sigma a^{-1}$*;*

(iii) $E\,(N/n^*) \to hb^{-1}$ as $a \to 0$*;*

(iv) $\liminf P\,(CD) \geq P^*$ *for all* $\underline{\theta} \in \Omega(a)$ as $a \to 0$*;*

where $n^* = b\sigma/a$ and the constant b comes from (2.15).

Proof: First we notice the basic inequality respect to the definition of N,

$$h\sigma_m a^{-1} \leq N \leq m + h\sigma_m a^{-1}.$$

Take expectation for each part of the inequality turns out to be part(ii) since $E(\hat{\sigma}_m) = \sigma$. Next, dividing throughout the expression by n^* and taking limits as $a \to 0$ leads to the part (iii).

Next, invoke the Helmert's orthogonal transformation, the $\hat{\sigma_m}$ can be expressed as

$$\begin{aligned}
\hat{\sigma}_m &= \textstyle\sum_{i=0}^{k} \sum_{j=1}^{m} (X_{ij} - T_i)/\left((k+1)(m-1)\right) \\
&= \tfrac{1}{2}\sigma(k+1)^{-1}(m-1)^{-1} \textstyle\sum_{i=1}^{m-1} \chi^2_{2(k+1)} \\
&= \tfrac{1}{2}\sigma(k+1)^{-1}(m-1)^{-1}\chi^2_{2(k+1)(m-1)},
\end{aligned}$$

According to the expression above, it is easy to have $f\hat{\sigma}_m/\sigma$ is χ^2_f distributed, where $f = 2(k+1)(m-1)$ is the degree of freedom. To prove the part (i), first note that theorem (2.1) holds for the two-stage selection procedure (4.1). That is, for any N determined by (4.1), the LFC is attached when all the populations partitioning have a location parameter either at $\mu_0 + \delta_1$ or $\mu_0 + \delta_2$. In addition, the distribution of N doesn't depend on $\boldsymbol{\theta}$; $I(N = n)$ is independent of $(T_{1n}, \cdots, T_{kn})$ for every fixed $n \geq 2$. Then for two-stage procedure, the $P(CD)$ based on the selection rule (2.4) with the initial sample size m and

stopping sample size N can be written as

$$
\begin{aligned}
&P\Big[CD|\boldsymbol{\theta}^0(r),\ \sigma\Big] \\
&= P\Big[T_i - T_0 < d,\ T_j - T_0 > d,\ \ 0 < i \le r,\ \ r < j \le k|\boldsymbol{\theta}^0(r),\ \ \sigma\Big] \\
&= P\Big[\tfrac{T_i-\theta_i}{\sigma/N} - \tfrac{T_0-\theta_0}{\sigma/N} < \tfrac{d-\theta_i+\theta_0}{\sigma/N},\ \tfrac{T_j-\theta_j}{\sigma/N} - \tfrac{T_0-\theta_0}{\sigma/N} > \tfrac{d-\theta_j+\theta_0}{\sigma/N},\ \ 1 \le i \le r,\ \ r+1 \le j \le k\Big] \\
&= P\Big[Z_i - Z_0 < \tfrac{d-\delta_1}{\sigma/N},\ Z_j - Z_0 > \tfrac{d-\delta_2}{\sigma/N},\ \ 1 \le i \le r,\ \ r+1 \le j \le k\Big] \\
&= P\Big[Z_i - Z_0 < \tfrac{aN}{\sigma},\ Z_j - Z_0 > -\tfrac{aN}{\sigma},\ \ 1 \le i \le r,\ \ r+1 \le j \le k\Big] \\
&\ge P\Big[Z_i - Z_0 < \tfrac{a}{\sigma}\tfrac{h\hat{\sigma}_m}{a},\ Z_j - Z_0 > \tfrac{-a}{\sigma}\tfrac{h\hat{\sigma}_m}{a},\ \ 1 \le i \le r,\ \ r+1 \le j \le k\Big] \\
&= P\Big[Z_i - Z_0 < hf^{-1}\chi^2_f,\ Z_j - Z_0 > -hf^{-1}\chi^2_f,\ \ 1 \le i \le r,\ \ r+1 \le j \le k\Big] \\
&= \int_0^\infty P\Big[Z_i - Z_0 < hf^{-1}t,\ Z_j - Z_0 > -hf^{-1}t,\ \ 1 \le i \le r,\ \ r+1 \le j \le k\Big] g(t)\ \mathrm{d}t \\
&= \int_0^\infty \int_0^\infty P\Big[Z_i < z + hf^{-1}t,\ Z_j > z - hf^{-1}t,\ \ 1 \le i \le r,\ \ r+1 \le j \le k\Big] g(t)f(z)\ \mathrm{d}t\mathrm{d}z \\
&= \int_0^\infty \int_0^\infty \Big\{\textstyle\prod_{i=1}^r P(Z_i < z + hf^{-1}t)\Big\}\Big\{\textstyle\prod_{i=1}^{k-r} P(Z_j > z - hf^{-1}t)\Big\} f(z)g(t)\ \mathrm{d}z\mathrm{d}t \\
&= \int_0^\infty \int_0^{hf^{-1}t} \{F(z + hf^{-1}t)\}^r f(z)g(t)\ \mathrm{d}z\mathrm{d}t \\
&+ \int_0^\infty \int_{hf^{-1}t}^\infty \{F(z + hf^{-1}t)\}^r \{1 - F(z - hf^{-1}t)\}^{k-r} f(z)g(t)\ \mathrm{d}z\mathrm{d}t,
\end{aligned}
$$

where F(.) and f(.) are respectively the cumulative function(cdf) and probability density function(pdf) of the standard exponential distribution as $F(z) = 1 - e^{-z}$ and $f(z) = e^{-z}$; and g(t) is the probability density function of χ^2_f. Here we define h as below

$$h = \frac{aN}{\hat{\sigma}_m} \tag{4.2}$$

In the above expression, the critical constant h depends upon the value of an integer $r(0 \le r \le k)$ and we have to compute the critical constant h that minimizes $P[CD]$ for $r = 0, 1, \cdots, k$ and simultaneously satisfies the probability requirement in order to maintain significance lever at α.

The least value of $P[CD]$ is attained for some unknown r in the range $r = 0, \ldots, k$. So in order to solve for h values, we perform a numerical way to detect r and deter-

mine h. Equivalently, we have to compute the critical constant $h^* = h_{k,m,\alpha}$ by solving $P\Big[CD|\boldsymbol{\theta}^0(r),\ \sigma\Big] = 1-\alpha$ for some value of $r \in (0, 1, \ldots, k)$ which minimize $P(CD)$ given some specified α. Following algorithm is implemented to compute the values of critical constants numerically.

Algorithm:

Step 1 For fixed values of k, m and α, compute numerically the values of critical constants $h(k, m, \alpha)$ that satisfy the probability requirement $P_r(CD|\boldsymbol{\theta}^0(r),\ \sigma) = 1-\alpha$ for $r = 0, 1, \ldots, k$. In other words, let $h_r(k, m, \alpha)$ be the solution of the integral equation:

$$\begin{aligned} &\int_0^\infty \int_0^{hf^{-1}t} \{F(z+hf^{-1}t)\}^r f(z)g(t)\ \mathrm{d}z\mathrm{d}t \\ &+ \int_0^\infty \int_{hf^{-1}t}^\infty \{F(z+hf^{-1}t)\}^r \{1-F(z-hf^{-1}t)\}^{k-r} f(z)g(t)\ \mathrm{d}z\mathrm{d}t = P^*, \end{aligned} \tag{4.3}$$

where $r = 0, 1, \ldots, k$.

Step 2 Let

$$h^* = \max\left(h_0, h_1, \ldots, h_k\right). \tag{4.4}$$

The optimum value of the critical constant satisfying the probability requirement $P\Big[CD|\boldsymbol{\theta}^0(r),\ \sigma\Big] = 1-\alpha$, irrespective of r, is h^*.

Next, to reduce the complexity of solving the integral equation (4.3) in step 1, the following steps are performed. Let I_1 and I_2 be the first and second integral of of the left hand side

of the equation. Then I_1 and I_2 can be simplified as

$$\begin{aligned}
I_1 &= \int_0^\infty \int_0^{hf^{-1}t} \{F(z+hf^{-1}t)\}^r f(z)g(t)\,\mathrm{d}z\mathrm{d}t \\
&= \int_0^\infty \int_0^{hf^{-1}t} (1-e^{-z-hf^{-1}t})^r e^{-z} g(t)\,\mathrm{d}z\mathrm{d}t \\
&= \int_0^\infty \int_0^{hf^{-1}t} \textstyle\sum_{i=0}^r \binom{r}{i}(-1)^i e^{-iz-hf^{-1}it} e^{-z} g(t)\,\mathrm{d}z\mathrm{d}t \\
&= \textstyle\sum_{i=0}^r \binom{r}{i}(-1)^i \int_0^\infty e^{-hf^{-1}it} g(t) \int_0^{hf^{-1}t} e^{-(i+1)z}\,\mathrm{d}z\mathrm{d}t \\
&= \textstyle\sum_{i=0}^r \binom{r}{i}(-1)^i (i+1)^{-1} \int_0^\infty e^{-hf^{-1}it}(1-e^{-(i+1)hf^{-1}t} g(t)\,\mathrm{d}t \\
&= \textstyle\sum_{i=0}^r \binom{r}{i}(-1)^i (i+1)^{-1} \int_0^\infty (e^{-hf^{-1}it} - e^{-(2i+1)hf^{-1}t}) g(t)\,\mathrm{d}t \\
&= \textstyle\sum_{i=0}^r \binom{r}{i}(-1)^i (i+1)^{-1} \Big[(2hf^{-1}i+1)^{f/2} - (2(2i+1)hf^{-1}+1)^{f/2}\Big].
\end{aligned} \tag{4.5}$$

Similarly, I_2 can be simplified as

$$\begin{aligned}
I_2 &= \int_0^\infty \int_{hf^{-1}t}^\infty \{F(z+hf^{-1}t)\}^r \{1-F(z-hf^{-1}t)\}^{k-r} f(z)g(t)\,\mathrm{d}z\mathrm{d}t \\
&= \int_0^\infty \int_{hf^{-1}t}^\infty (1-e^{-z-hf^{-1}t})^r (e^{-z+hf^{-1}t})^{k-r} e^{-z} g(t)\,\mathrm{d}z\mathrm{d}t \\
&= \int_0^\infty \int_{hf^{-1}t}^\infty \textstyle\sum_{i=0}^r \binom{r}{i}(-1)^i e^{-iz-hf^{-1}it} e^{-(k-r)z+(k-r)hf^{-1}t} e^{-z} g(t)\,\mathrm{d}z\mathrm{d}t \\
&= \textstyle\sum_{i=0}^r \binom{r}{i}(-1)^i \int_0^\infty e^{-hf^{-1}it} e^{(k-r)hf^{-1}t} g(t) \int_{hf^{-1}t}^\infty e^{-(k-r+1+i)z}\,\mathrm{d}z\mathrm{d}t \\
&= \textstyle\sum_{i=0}^r \binom{r}{i}(-1)^i (k-r+1+i)^{-1} \int_0^\infty e^{-hf^{-1}it} e^{(k-r)hf^{-1}t} g(t) e^{-(k-r+1+i)hf^{-1}t}\,\mathrm{d}t \\
&= \textstyle\sum_{i=0}^r \binom{r}{i}(-1)^i (k-r+1+i)^{-1} \int_0^\infty e^{-(2i+1)hf^{-1}t} g(t)\,\mathrm{d}t \\
&= \textstyle\sum_{i=0}^r \binom{r}{i}(-1)^i (k-r+1+i)^{-1} \Big[2(2i+1)hf^{-1}+1\Big]^{f/2}.
\end{aligned} \tag{4.6}$$

Combining (4.5) and (4.6), we have the following

$$\begin{aligned}
&\sum_{i=0}^r \binom{r}{i}(-1)^i \Big[(i+1)^{-1}(2hf^{-1}i+1)^{f/2} - (i+1)^{-1}[2(2i+1)hf^{-1}+1]^{f/2} \\
&+ (k-r+1+i)^{-1}[2(2i+1)hf^{-1}+1]^{f/2}\Big] = P^*.
\end{aligned} \tag{4.7}$$

That is, alternatively one can determine $h_r = h(k, m, \alpha)$ such that is satisfy (4.7). After comparing the h values for different $r(0 \le r \le k)$ under fixed different k m and α, the values of critical constant are obtained. Table (4.1)-(4.3) presents the values of critical

constant for $m = 2, \ldots, 20$, $k = 2, \ldots, 10$ and $P^* = 0.9,\ 0.95,\ 0.99$. Then, it is straight forward known that probability requirement (2.9) is satisfied if n satisfies that

$$n \geq \frac{h\hat{\sigma}_m}{a} \tag{4.8}$$

Along with the expression of (4.1), this procedure certainly guaranteeing the requirement of correct partition rate. This completes the proof of part(i).

■

Table 4.1: Critical Constant value of h, as defined in (4.4) for $P^* = 0.9$

m	k								
	2	3	4	5	6	7	8	9	10
2	3.4223	3.7513	3.9423	4.0805	4.1914	4.2857	4.3685	4.4428	4.5104
3	2.7939	3.1714	3.4186	3.6046	3.7548	3.8815	3.9914	4.0887	4.1761
4	2.6163	3.0014	3.2612	3.4590	3.6194	3.7549	3.8723	3.9761	4.0693
5	2.5327	2.9203	3.1855	3.3884	3.5535	3.6930	3.8139	3.9208	4.0167
6	2.4841	2.8729	3.1409	3.3468	3.5145	3.6563	3.7793	3.8879	3.9854
7	2.4523	2.8417	3.1116	3.3193	3.4887	3.6320	3.7563	3.8661	3.9646
8	2.4300	2.8197	3.0908	3.2999	3.4704	3.6148	3.7400	3.8506	3.9498
9	2.4133	2.8034	3.0754	3.2853	3.4568	3.6019	3.7278	3.8390	3.9387
10	2.4005	2.7907	3.0634	3.2741	3.4462	3.5919	3.7183	3.8300	3.9301
11	2.3903	2.7806	3.0538	3.2651	3.4377	3.5839	3.7107	3.8228	3.9233
12	2.3820	2.7724	3.0460	3.2578	3.4308	3.5774	3.7046	3.8170	3.9177
13	2.3751	2.7656	3.0396	3.2517	3.4251	3.5720	3.6994	3.8121	3.9130
14	2.3693	2.7598	3.0341	3.2465	3.4203	3.5674	3.6951	3.8079	3.9091
15	2.3643	2.7549	3.0294	3.2421	3.4161	3.5635	3.6914	3.8044	3.9057
16	2.3600	2.7506	3.0254	3.2383	3.4125	3.5601	3.6881	3.8013	3.9028
17	2.3563	2.7469	3.0218	3.2350	3.4094	3.5571	3.6853	3.7987	3.9002
18	2.3530	2.7436	3.0187	3.2320	3.4066	3.5545	3.6828	3.7963	3.8979
19	2.3500	2.7407	3.0160	3.2294	3.4041	3.5522	3.6806	3.7942	3.8959
20	2.3474	2.7381	3.0135	3.2271	3.4019	3.5501	3.6787	3.7923	3.8941

Table 4.2: Critical Constant value of h, as defined in (4.4) for $P^* = 0.95$

m	k								
	2	3	4	5	6	7	8	9	10
2	5.1038	5.2493	5.3150	5.3651	5.4110	5.4553	5.4984	5.5403	5.5810
3	3.8767	4.2012	4.4088	4.5650	4.6921	4.8004	4.8953	4.9800	5.0568
4	3.5504	3.9084	4.1471	4.3285	4.4760	4.6009	4.7096	4.8061	4.8930
5	3.4001	3.7712	4.0230	4.2154	4.3719	4.5044	4.6195	4.7214	4.8130
6	3.3138	3.6917	3.9506	4.1491	4.3108	4.4476	4.5663	4.6713	4.7656
7	3.2578	3.6398	3.9032	4.1056	4.2706	4.4101	4.5312	4.6382	4.7343
8	3.2185	3.6033	3.8698	4.0748	4.2421	4.3835	4.5062	4.6147	4.7120
9	3.1895	3.5762	3.8449	4.0519	4.2208	4.3637	4.4876	4.5972	4.6954
10	3.1671	3.5553	3.8257	4.0342	4.2044	4.3484	4.4732	4.5836	4.6825
11	3.1494	3.5387	3.8104	4.0201	4.1933	4.3361	4.4617	4.5727	4.6722
12	3.1398	3.5252	3.7979	4.0086	4.1806	4.3261	4.4524	4.5639	4.6638
13	3.1230	3.5139	3.7876	3.9991	4.1717	4.3178	4.4446	4.5565	4.6568
14	3.1129	3.5045	3.7789	3.9910	4.1642	4.3108	4.4380	4.5503	4.6509
15	3.1043	3.4964	3.7714	3.9841	4.1578	4.3048	4.4323	4.5450	4.6459
16	3.0969	3.4894	3.7650	3.9781	4.1523	4.2996	4.4274	4.5403	4.6415
17	3.0904	3.4833	3.7593	3.9729	4.1474	4.2951	4.4232	4.5363	4.6377
18	3.0847	3.4779	3.7544	3.9683	4.1431	4.2911	4.4194	4.5328	4.6343
19	3.0796	3.4732	3.7500	3.9642	4.1393	4.2875	4.4161	4.5296	4.6313
20	3.0751	3.4689	3.7460	3.9606	4.1359	4.2844	4.4131	4.5268	4.6286

Table 4.3: Critical Constant value of h, as defined in (4.4) for $P^* = 0.99$

m	k								
	2	3	4	5	6	7	8	9	10
2	9.1644	9.8849	9.3020	8.9395	8.7007	8.5367	8.4205	8.3366	8.2753
3	6.9220	6.9439	6.9584	6.9816	7.0116	7.0457	7.0819	7.1190	7.1562
4	6.0115	6.2082	6.3400	6.4458	6.5370	6.6182	6.6920	6.7598	6.8228
5	5.6129	5.8768	6.0563	6.1968	6.3142	6.4162	6.5067	6.5883	6.6629
6	5.3899	5.6887	5.8936	6.0530	6.1851	6.2985	6.3985	6.4880	6.5691
7	5.2476	5.5675	5.7882	5.9595	6.1007	6.2216	6.3275	6.4221	6.5075
8	5.1489	5.4830	5.7144	5.8938	6.0414	6.1673	6.2774	6.3755	6.4639
9	5.0765	5.4207	5.6598	5.8451	5.9973	6.1270	6.2402	6.3408	6.4315
10	5.0210	5.3728	5.6178	5.8076	5.9634	6.0959	6.2114	6.3140	6.4064
11	4.9773	5.3349	5.5845	5.7778	5.9364	6.0711	6.1885	6.2927	6.3864
12	4.9418	5.3042	5.5574	5.7536	5.9144	6.0509	6.1699	6.2753	6.3700
13	4.9126	5.2788	5.5350	5.7335	5.8961	6.0342	6.1544	6.2608	6.3565
14	4.8880	5.2574	5.5161	5.7166	5.8807	6.0201	6.1413	6.2486	6.3451
15	4.8670	5.2391	5.5000	5.7021	5.8676	6.0080	6.1301	6.2382	6.3353
16	4.8489	5.2233	5.4861	5.6896	5.8562	5.9976	6.1204	6.2292	6.3268
17	4.8332	5.2096	5.4739	5.6787	5.8463	5.9885	6.1120	6.2213	6.3194
18	4.8194	5.1975	5.4632	5.6691	5.8376	5.9804	6.1046	6.2144	6.3129
19	4.8071	5.1868	5.4537	5.6606	5.8298	5.9733	6.0979	6.2082	6.3071
20	4.7962	5.1773	5.4453	5.6530	5.8229	5.9669	6.0920	6.2027	6.3019

4.3 Monte Carlo Simulation Study of Two-stage Procedure

The two-stage procedure (4.1), starts with $m \geq 2$ at the first stage observations from each of the k populations and the control population. The procedure takes $N - m$ additional samples at the second stage from each all the k populations and the control population. In this section, some simulation are performed under different setting of n^*, h, a and σ. In table (4.4-4.8), we took the value of $m = 5$ and $m = 10$, $k = 4$, and $P^* = 0.9$, $p^* = 0.95$ and $p^* = 0.99$. Respect to the value of k and P^*, the value of the design constant h was obtained from (3.8). Recall that the design parameters δ_1 and δ_2 are predetermined by the experimenter based on the definition of the Good and Bad populations. Note that $a = \delta_2 - \delta_1$, which represents the length of indifference zone.

Table 4.4: Simulation of Two-stage Procedure (4.1); Based on 100,000 iterations, p^* =0.95, k=4, m =5

configuration	n^*	$\bar{n}$	Std(n)	$\bar{P}$	Std($\bar{P}$)	b	h	a	σ	λ
(0,4)	36.7605	40.7309	0.0284	0.9801	0.0004	3.6760	4.0230	0.5	5	10
(1,3)	36.7605	40.7130	0.0285	0.9690	0.0005	3.6760	4.0230	0.5	5	10
(2,2)	36.7605	40.7430	0.0285	0.9587	0.0006	3.6760	4.0230	0.5	5	10
(3,1)	36.7605	40.7202	0.0284	0.9518	0.0006	3.6760	4.0230	0.5	5	10
(4,0)	36.7605	40.7110	0.0283	0.9536	0.0006	3.6760	4.0230	0.5	5	10
(0,4)	73.521	80.9339	0.0179	0.9802	0.0001	3.6760	4.0230	0.5	10	20
(1,3)	73.521	80.9671	0.0179	0.9689	0.0001	3.6760	4.0230	0.5	10	20
(2,2)	73.521	80.9757	0.0180	0.9587	0.0001	3.6760	4.0230	0.5	10	20
(3,1)	73.521	80.9655	0.0179	0.9515	0.0002	3.6760	4.0230	0.5	10	20
(4,0)	73.521	80.9202	0.0179	0.9525	0.0002	3.6760	4.0230	0.5	10	20

4.4 Concluding Remarks and Future Work

As well as proposing a convenient partition rule this dissertation provides the sample size needed for practitioners in order to reach a prerequired probability of correct decision,

Table 4.5: Simulation of Two-stage Procedure (4.1); Based on 100,000 iterations, p^* =0.9, k=4, m =5

configuration	n^*	$\bar{n}$	Std(n)	$\bar{P}$	Std($\bar{P}$)	b	h	a	σ	λ
(0,4)	29.6939	32.3596	0.0225	0.9603	0.0006	2.9639	3.1855	0.5	5	10
(1,3)	29.6939	32.3532	0.0225	0.9381	0.0007	2.9639	3.1855	0.5	5	10
(2,2)	29.6939	32.3567	0.0225	0.9187	0.0008	2.9639	3.1855	0.5	5	10
(3,1)	29.6939	32.3665	0.0225	0.9049	0.0009	2.9639	3.1855	0.5	5	10
(4,0)	29.6939	32.3583	0.0225	0.9075	0.0009	2.9639	3.1855	0.5	5	10
(0,4)	59.3878	64.2152	0.0451	0.9594	0.0002	2.9639	3.1855	0.5	10	20
(1,3)	59.3878	64.2055	0.0449	0.9359	0.0002	2.9639	3.1855	0.5	10	20
(2,2)	59.3878	64.2158	0.0449	0.9167	0.0003	2.9639	3.1855	0.5	10	20
(3,1)	59.3878	64.2119	0.0451	0.9020	0.0003	2.9639	3.1855	0.5	10	20
(4,0)	59.3878	64.1813	0.0450	0.9051	0.0003	2.9639	3.1855	0.5	10	20

Table 4.6: Simulation of Two-stage Procedure (4.1); Based on 100,000 iterations, p^* =0.99, k=4, m =5

configuration	n^*	$\bar{n}$	Std(n)	$\bar{P}$	Std($\bar{P}$)	b	h	a	σ	λ
(0,4)	52.958	61.0838	0.0135	0.9962	0.0004	5.2958	6.0563	0.5	5	10
(1,3)	52.958	61.0423	0.0135	0.9939	0.0005	5.2958	6.0563	0.5	5	10
(2,2)	52.958	61.0672	0.0135	0.9921	0.0006	5.2958	6.0563	0.5	5	10
(3,1)	52.958	61.0735	0.0135	0.9904	0.0006	5.2958	6.0563	0.5	5	10
(4,0)	52.958	61.0761	0.0135	0.9905	0.0006	5.2958	6.0563	0.5	5	10
(0,4)	105.916	121.6386	0.0271	0.9961	0.0001	5.2958	6.0563	0.5	10	20
(1,3)	105.916	121.6365	0.0271	0.9938	0.0001	5.2958	6.0563	0.5	10	20
(2,2)	105.916	121.6433	0.0270	0.9919	0.0001	5.2958	6.0563	0.5	10	20
(3,1)	105.916	121.6309	0.0271	0.9901	0.0001	5.2958	6.0563	0.5	10	20
(4,0)	105.916	121.6166	0.0270	0.9902	0.0001	5.2958	6.0563	0.5	10	20

Table 4.7: Simulation of Two-stage Procedure (4.1); Based on 100,000 iterations, p^* =0.9, k=4, m =10

configuration	n^*	$\bar{n}$	Std(n)	$\bar{P}$	Std($\bar{P}$)	b	h	a	σ	λ
(0,4)	29.6939	31.1381	0.0045	0.9611	0.0001	2.9639	3.0634	0.5	5	10
(1,3)	29.6939	31.1534	0.0045	0.9384	0.0002	2.9639	3.0634	0.5	5	10
(2,2)	29.6939	31.1329	0.0045	0.9194	0.0002	2.9639	3.0634	0.5	5	10
(3,1)	29.6939	31.1397	0.0045	0.9050	0.0002	2.9639	3.0634	0.5	5	10
(4,0)	29.6939	31.1279	0.0045	0.9072	0.0002	2.9639	3.0634	0.5	5	10
(0,4)	59.3878	61.7847	0.0091	0.9599	0.0001	2.9639	3.0634	0.5	10	20
(1,3)	59.3878	61.7586	0.0091	0.9370	0.0002	2.9639	3.0634	0.5	10	20
(2,2)	59.3878	61.7827	0.0091	0.9173	0.0002	2.9639	3.0634	0.5	10	20
(3,1)	59.3878	61.7739	0.0091	0.9030	0.0002	2.9639	3.0634	0.5	10	20
(4,0)	59.3878	61.7687	0.0091	0.9053	0.0002	2.9639	3.0634	0.5	10	20

Table 4.8: Simulation of Two-stage Procedure (4.1); Based on 100,000 iterations, p^* =0.95, k=4, m =10

configuration	n^*	$\bar{n}$	Std(n)	P	Std(P)	b	h	a	σ	λ
(0,4)	36.7605	38.7659	0.0057	0.9806	0.0001	3.6760	3.8257	0.5	5	10
(1,3)	36.7605	38.7657	0.0057	0.9697	0.0001	3.6760	3.8257	0.5	5	10
(2,2)	36.7605	38.7491	0.0057	0.9597	0.0001	3.6760	3.8257	0.5	5	10
(3,1)	36.7605	38.7552	0.0057	0.9524	0.0002	3.6760	3.8257	0.5	5	10
(4,0)	36.7605	38.7549	0.0057	0.9533	0.0002	3.6760	3.8257	0.5	5	10
(0,4)	73.5210	77.0131	0.0114	0.9800	0.0001	3.6760	3.8257	0.5	10	20
(1,3)	73.5210	77.0022	0.0114	0.9690	0.0001	3.6760	3.8257	0.5	10	20
(2,2)	73.5210	76.9935	0.0114	0.9590	0.0001	3.6760	3.8257	0.5	10	20
(3,1)	73.5210	76.9941	0.0114	0.9509	0.0002	3.6760	3.8257	0.5	10	20
(4,0)	73.5210	77.0104	0.0113	0.9519	0.0002	3.6760	3.8257	0.5	10	20

which is selecting "Good" and "Bad" two-parameter exponential populations with a control respect to the location parameter when scaling parameter is unknown. When scaling parameter is unknown, this dissertation proposes two sequential procedures to implement which also guarantees the PCD under the same partition rule. Of course, more studies on the proposed model remain to be done. They include the following several aspects.

(i) This dissertation assumes the scale parameter stays the same among all exponential populations as well as the control population. But heteroscedasiticity may exist for the two-parameter exponential populations partitioning.

(ii) This dissertation assumes that statisticians are selecting the sample size a vector at a time. This consideration is made since no prior knowledge is given on the cost of selecting. For example, when comparing several drugs is needed, the cost of collecting a patient taking one drug may be more expensive then collecting a patient taking a different drug. Under such circumstances, the procedure proposed in this dissertation can be improved.

(iii) In this dissertation, a single population is considered in the control set. But it is possible to have more than one control in some real cases.

Bibliography

[1] Bain, L.J., Engelhardt, M., (1991). *Statistical Analysis of Reliability and Life Testing Models*. Marcel Dekker, New York.

[2] Bechhofer, R. E. (1954). A Single-sample multiple decision procedure for ranking means of normal populations with known variances, *Annals of Mathematical Statistics* 25, 16-39.

[3] Chow, Y.S. and Robbins, H. (1965). On the asymptotic theory of fixed width sequential confidence intervals for the means, *Annals of Mathematical Statistics* 36: 457-462.

[4] Cohen, A. C., and Helm, F. R. (1973). Estimators in the exponential distribution, *Technometrics* 15, 415-418.

[5] Datta, S. and Mukhopadhyay, N. (1998). Second-order asymptotics for multistage methodologies in partitioning a set of normal populations having a common unknown variance, *Statistics and Decisions* 16, 191-205.

[6] Desu, M. M., Narula, S. C. and Villarreal, B. (1977). A two-stage procedure for selecting the best of k exponential distributions, *Communication Statistics* 6, 1231-1243.

[7] Gupta, S.S. (1956). On a decision rule for a problem in ranking means. Ph.D. diss., University of North Carolina-Chapel Hill.

[8] Johnson, N.L., Kotz, S., Balakrishnan, N., (1994). *Continuous Univariate Distributions*. Wiley, New York.

[9] Lam, K., (1987). Subset selection of normal populations under heteroscedasticity. *Proceedings of the Second International Advanced Seminar/Workshop on Inference Procedures Associated with Statistical Ranking and Selection*, Sydney, Australia, August 1987.

[10] Lam, K., (1988). An improved two-stage selection procedure. *Communications in Statistics - Simulation and Computation* 17 (3), 995-1006.

[11] Lawless, J.F. (2003). *Statistical Models and Methods for Lifetime Data*. Wiley, New York.

[12] Lawless, J.F., Singhal, K., (1980). Analysis of data from life test experiments under an exponential model. *Naval Research Logistics Quarterly 27*, 323-334.

[13] Mukhopadhyay, N. (1986). On selecting the best exponential population, *Journal of Indian Statistical Association* 24, 31-41.

[14] Mukhopadhyay, N. and Solanky, T.K.S. (1994). *Multistage Selection and Ranking Procedures*. Marcel Dekker: New York.

[15] Ng, C.K., Lam, K., Chen, H.J., (1993). Multiple comparison of exponential location parameters with the best under type II censoring, *American Journal of Mathematical and Management Sciences* 12, 383-402.

[16] Raghavachari, M. and Starr, N. (1970). Selection problems for some terminal distributions, *Metron* 28, 185-197.

[17] Ray, W. D. (1957). Sequential confidence intervals for the mean of a normal population with unknown variance, *Journal of the Royal Statistical Society, Series B* 19, 133-143.

[18] Robbins, H., Sobel, M. and Starr, N. (1968). A sequential procedure for selecting the best of k populations, *Annals of Mathematical Statistics* 39, 88-92.

[19] Solanky, T.K.S. (2006). A two-stage procedure with elimination for partitioning a set of normal populations with respect to a control, *Sequential Analysis* 25, 297-310.

[20] Tong, Y.L. (1969). On Partitioning a Set of Normal Populations by Their Locations with Respect to a Control, *Annals of Mathematical Statistics* 40, 1300-1324.

[21] Wiener, N. (1939). The egodic theorem, *Duke Mathematical Journal* 5, 1-18.

[22] Woodroofe, M.. (1977) Second order approximations for sequential point and interval estimation. *Annals of Statistics* 10, 984-995.

[23] Wu, S.F., Lin, Y.P., Yu, Y.R., (2010). One-stage multiple comparisons with the control for exponential location parameters under heteroscedasticity, *Computational Statistics and Data Analysis* 54 1372-1380.

[24] Zelen, M. (1966). Application of exponential models to problems in cancer research, *Journal of the Royal Statistical Society* 129, 368-398.

[25] Zelen, M.. and Dannemiller, M. C. (1961). The robustness of life testing procedures derived from the exponential distribution, *Technometrics* 3, 29-49.

Vita

The author, *Rui Wang*, was born in Chibi, Hubei province, China. He obtained his Bachelor's degree in Electronic and Information Engineering from Huazhong University of Science and Technology in 2010. He registered as a graduate student in Department of Mathematics, University of New Orleans in 2012. After finished the master program, he continued to pursue the doctoral degree of Engineering and Applied Science. During Ph.D program, he worked on generalization of partition problem in statistics. One of the applications of his research is to help doctors and medical researchers in identifying good drugs and compare efficacy of the new developed drugs with other existing drugs. His main research interests include

- Survival analysis,
- Clinical trials,
- Experimental design,
- Statistical partition,