STATISTICAL INFERENCE FOR TREATMENTS VERSUS A CONTROL

JIANAN PENG

Statistical Inference for Treatments versus a Control

by

©Jianan Peng

A thesis submitted to the
School of Graduate Studies
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Department of Mathematics and Statistics
Memorial University of Newfoundland

January 2002 Submitted

St. John's Newfoundland Canada

Abstract

The treatments versus a control problem occurs in many scientific fields, with a major portion in medical research. Its primary goal is to determine if the response to one or more treatments differ from the response to a control or existing standard and if so, further to identify which treatments are better than the control. In many experiments, one often has a prior knowledge that the treatments are at least as effective as the control. It is well known that utilization of ordering information increases the efficiency of statistical inference procedures. The aim of this thesis is to develop some new statistical inference procedures for the problem by utilizing the prior information.

In particular, simultaneous confidence lower bounds for the differences between treatment means and the control mean are considered. Efficient computation algorithms are proposed to obtain the optimal lower bounds between the best treatment mean and the control mean. Multiple contrast tests which take account of the prior knowledge play an important role in this thesis.

Power studies via simulation compare the new proposed procedures with Dunnett's procedure and the likelihood ratio test. The new proposed procedures are also illustrated by some real data sets.

Acknowledgments

I owe my profound gratitude to my supervisor, Professor Chu-In Charles Lee. He has been very generous with his ideas and time. His insight, dedication, and continuing support have made it possible for me to complete this endeavor. I also thank him and his wife Mrs. Lisa Lee for their kind concerns for my child Chantal Xiangting Peng.

Many thanks go to Professors Edgar G. Goodaire and Paul Peng as members of my supervisor Commitee. I feel indebted to Professor Hong Wang for his financial support and help during my graduate study. I am also grateful to Professor Veeresh Gadag for his financial support in 1996. My thanks also go to Dr. Lin Liu for helpful discussions with her.

I would like to thank the Department of Mathematics and Statistics, and Professors Bruce Watson and Herb Gaskill, the past and present Department Heads, for providing me with a friendly atmosphere and necessary facilities to finish my program.

I am grateful to the School of Graduate Studies, Memorial University of Newfoundland for supporting me with a Graduate Fellowship. I also thank the Department of Computing and Communication for making available to me the high performance computing facilities.

I would like to thank Dr. Hari Mukerjee, Dr. Hong Wang and Dr. Alwell Oyet who have spent their valuable time in reading my thesis. Their suggestions for the better presentation of this thesis are greatly appreciated.

I express my thanks to Drs Roy Bartlett, Louise Dionne, Hong Wang and Steve

Xiaoshen Yan for their impressive lectures.

I extend my appreciation to Mrs. Cathy Gaskill for her time and patience to help improve my English. My appreciation also goes to Dr. Nabil Shalaby for his encouragement and help.

I would also like to thank Professors Ming Liao and Rong Wu for their encouragement.

I'm especially grateful to my family: to my wife, Guiwen, whose patience, undedrstanding and love have sustained me through the rough times. And to my child, Xiangting, who brings me joy. Finally, I would like to express my sincere appreciation to my parents for their care.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The problem of comparing several treatment populations with a control population occurs frequently in medical trials and other experiments. For example, in comparative clinical trials, different therapies are often compared with a standard therapy or placebo to determine which therapy increases the survival time of patients with a certain disease. One of the main pharmaceutical activities is the search for new drugs that are more effective ("better") than a standard drug or placebo: in this case treatments can be different dose levels of a new drug and the control can be a standard drug or placebo.

Traditionally, a common tool in analyzing data from these studies is a test of homogeneity of the treatment means and the control mean as in the Analysis of Variance. However, such homogeneity tests (whether or not they yield statistically significant results) usually do not supply the type of conclusion that the experimenter truly desires. Furthermore, should a significant result be obtained , the experimenter's problems have only just begun since the experimenter is seldom satisfied with terminating the analysis at this point; in particular, he or she may want to determine which treatment is better than the control or to see which treatment can

be considered best in some well-defined sense of the term best. Moreover, there may be a question as to whether testing a null hypothesis of homogeneity and estimating a parameter are appropriate formulations of the problem. In many medical trials, the most important problem is to correctly identify the best treatment or treatments. Hence, a formulation such as a simultaneous statistical inference problem or ranking and selection problem ought to be more realistic in these cases. Formulating the statistical problem in terms of selection or multiple comparison would seem to be particularly pertinent if the choice of a therapy for a group of patients is to be made on the basis of the trial results or if the treatments to be studied in a later trial depend upon which treatment turns out to be superior to the others in the initial trial. In this thesis, we will explore some new methods in simultaneous statistical inference based on univariate response under the assumptions of normality and homogeneous variances regarding "treatments versus a control" problems. If the homogeneous variance assumption is questionable, one may use nonparametric methods as in the paper of Chakraborti and Desu (1991) which includes methods to handle censored data.

The field of multiple comparisons which forms a part of a broader subject of simultaneous statistical inference has progressed tremendously in the last 40 years. A large number of theoretical developments have taken place including different approaches to error rate control, step-wise testing procedures, modified Bonferroni procedures, resampling methods and so on. The research is motivated by problems in many fields such as medicine, education, quality control, biology, genetics, and the physical sciences. A number of statistical procedures have been proposed to test whether any of k treatments are different from a control and most of them

are multiple comparisons procedures. Dunnett's procedure (1955) is surely the best known and most widely used. Many generalizations of Dunnett's procedure have been made. For instance, in various medical and industrial fields, it is of interest to compare K_t test treament groups and K_c control groups to determine which treament is better than which control. Shaffer (1977) generalized Dunnett's method to allow for more than one control and presented two-sided $100(1-\alpha)\%$ simultaneous confidnece intervals to compare K_t to K_c controls. Hoover (1991) computed Shaffer's (1977) two-sided critical values for $K_c = 2$ with equal sample size. He also developed a one-sided procedure for comparing the means of K_t treatments to K_c controls and computed the corresponding critical values for $K_c = 2$ with equal sample size. Solorzano and Spurrier (1999) developed simultaneous one-sided confidnece intervals comparing K_t treatments and K_c controls for unequal sample sizes. Cheung and Holland (1991, 1992) extended Dunnett's procedure (1955) to the case of more than one group of treatments, each group containing several treatments compared with a specified treatment, with the error rate covering all groups and treatment comparisons simultaneously. Using the closure method of Marcus, Peritz and Gabriel (1976), Dunnett's method (1955) has been developed into a step-down and a step-up procedure in Dunnett and Tamhane (1992) and Liu (1997a). Giani and Straßburger (2000) proposed multiple comparison procedures to discriminate good, equivalent and bad treatments with respect to the control. Procedures of testing simultaneously for superiority and equivalence of a new treatment compared with $k \geq 1$ standard treatments in a clinical trial have been recently developed by Dunnett and Tamhane (1997) and Kwong (2001).

A multiple comparison procedure makes one or more assertions. Any incorrect

assertion given by the multiple comparison procedure may lead to an incorrect decision. Various criteria exist by which one can assess the performance of a multiple comparisons procedure. One can consult Hochberg and Tamhane (1987). In this thesis, *error rate* of a multiple comparison procedure is defined to be the supremum of the probability of making at least one incorrect assertion, i.e.,

$$\text{error rate} = \sup_{\mu} P_{\mu}\{\text{at least one incorrect assertion}\}.$$

Hochberg and Tamhane (1987) calls it strong familywise error rate (FWE).

There are several books on statistical simultaneous inference available. The book by Miller (1981) is highly accessible but does not cover recent developments. The book by Hochberg and Tamhane (1987) is an excellent resource for research statisticians but seems formidable for typical practitioners. The book by Westfall and Young (1993) capably shows how modern computers enable one to adjust the p-values of tests of hypotheses for multiplicity; however, it is often desirable to go beyond stating p-values and infer the direction and the magnitude of the differences among the treatments being compared. The latest book by Hsu (1996) emphasizes proper application of the latest methods for confident directions inference and confidence interval inference empowered by modern computers.

Multiple comparisons procedure has its roots in ranking and selection, but the aim of ranking and selection is different from that of simultaneous confidence intervals. It is well-known that there are two principal formulations developed in the statistical literature regarding ranking and selection. One approach, the indifference zone selection, was suggested by Bechhofer (1954). The second approach, subset selection, has been presented by Gupta and Sobel (1958) and Gupta (1965). The

indifference zone approach has as its goal to select or indicate the best treatment without a control. The probability of selecting the best treatment is at least P^* whenever the best treatment is at least δ^* (> 0) away from the second best treatment. This minimal probability P^* can only be guaranteed if the common sample size n is large enough. So the indifference zone approach is useful at the experimental design stage in order to determine this common sample size n. When there is a control, Dunnett (1984) proposed the corresponding indifference zone approach to select the best treatment which is at the same time better than the control. The subset selection has as its goal to select a non-empty subset, as small as possible, from the k treatments without a control in order to include the best treatment or the treatments which are bettter than the control, with a certain confidence. The size of the subset is random. The confidence requirement has to be met for all parameter configurations. For book-length discussions on ranking and selection, see Gibbons, Olkin and Sobel (1977) and Bechhofer, Santner and Goldsman (1995).

Frequently there is a prior knowledge that the treatments are at least as effective as the control. This type of prior knowledge may come from past experiences or it may arise in medical experiments where, for example, a higher dose level of a drug has a larger effect on patients; some treatments are known to have survival times at least as long as that of the control. The prior knowledge can be expressed as $\mu_0 \leq \mu_i, i = 1, \cdots, k$, where μ_0 denotes the mean value from the control and μ_i denotes the mean value from treatment i $(i = 1, \cdots, k)$. This type of inequality constraints is typically called simple tree order in order restriction inference.

Statistical inference under order restrictions is an important field in statistics. Many types of problems are concerned with identifying meaningful structure in real

world situations. Structure characterized by order restrictions arises in numerous settings and has many useful applications. For example, the failure rate of a component may increase as it ages; survival times for treatments are longer than that of a control (see Singh and Wright (1998)); treatment responses may be stochastically dominated by a control; or treatments present simple tree uniform stochastic ordering (see Park (1992)). The books of Barlow, Bartholomew, Brenner and Brunk (1972), and Robertson, Wright and Dykstra (1988) are two classical monographs on this field and contain many important problems.

Utilizing ordering information increases the efficiency of statistical inference procedures. The one-tailed, two-sample t-test provides a familiar example in which the procedure which utilizes the prior knowledge dominates procedures which ignore this knowledge. It is well known that the isotonic regression reduces total square error and maximum absolute error. Lee (1981) poineered the problem of pointwise mean square error for the normal means with a simple order. He showed that in this case mean square error is reduced for every individual mean by using order restricted MLE's. Lee (1988) also observed that these pointwise properties do not hold, in general, for partial order restrictions such as simple tree ordering. Why should mean square error be reduced by the isotonic regression at each point for a simple order and not for the simple tree partial order? The possible explanation is that the number of lower sets is linear in the number of populations for the simple order, while for the simple tree partial order it is exponential in the number of populations.

The test of homogeneity against the simple tree order has been well developed. The classical likelihood ratio tests $\bar{\chi}^2$ or $\bar{E}^2$ for testing the homogeneity of partially

ordered means from several normal populations, first proposed by Bartholomew (1959, 1961), is known to possess generally superior operating characteristics to those of its competitors, see Robertson, Wright and Dykstra (1988). However, difficulties in computing the restricted maximum likelihood estimates and determining the null distributions of the test statistics make LRTs difficult to implement in many instances, particularly when the sample sizes are unequal. Recently, Miwa, Hayter and Liu (2000) developed some programs to numerically compute the level probabilities for the simple order. Therefore, it is of considerable interest to approximate the null distributions of $\bar{\chi}^2$ or $\bar{E}^2$. For example, Wright and Tran (1985) discussed that the equal-weights distributions of these test statistics provide reasonable approximations for the case of unequal sample sizes if the sample sizes are not too different for the simple tree ordering. The amount of variation in the null $\bar{\chi}^2$ and $\bar{E}^2$ distributions, as functions of the weight vector, can be determined by obtaining the sharp upper and lower bounds for the appropriate tail probabilities in terms of the partial order under consideration such as considered in Lee, Robertson and Wright (1993). The complexity of the null $\bar{\chi}^2$ and $\bar{E}^2$ distributions also motivates one to seek other test procedures. A variety of other procedures have been proposed, most of which are based on one or more contrasts among the sample means. These include the multiple contrast tests of Mukerjee, Robertson, and Wright (1987) and ad hoc tests proposed by Williams (1971, 1972). Conaway, Pillers, Robertson, and Sconing (1991) used a circular cone to approximate the LRT and they developed a test which has the advantages of being easier to compute and can be used with unequal sample sizes. Tang and Lin (1997) used an orthant to approximate S_{01}. McDermott (1999) proposed a class of tests based on an improved orthant approximation which can be

viewed as generalizations of the multiple contrast tests of Mukerjee, Robertson, and Wright (1987).

As far as making statistical inference is concerned, a confidence interval provides a visual perspective unmatched by a point estimate or a test statistic. Dunnett's procedure (1955) probably is the first one for simultaneous confidence intervals for simple tree restriction, but he did not realize this. Bohrer (1967) showed how the usual simultaneous two-sided Scheffé bounds on all linear functions of certain parameters can be sharpened if attention is restricted to only linear combinations of normal means whose coefficients are known to be nonnegative. Bohrer and Francis (1972) further developed simultaneous one-sided confidence bounds in this restricted setting. Marcus and Peritz (1976) also developed a methodology for obtaining simultaneous confidence intervals for linear combinations of normal means with certain restrictions on the coefficients. They obtained the Bohrer-Francis confidence intervals as a special case of this procedure. Marcus (1978) developed a set of simultaneous confidence bounds (SCB) for simple order and simple tree order in the case of known variance. Korn (1982) studied confidence bands for monotone dose-response curves without a control. With the assumption that the response means are monotone nondecreasing $\mu_1 \leq \mu_2 \leq \ldots \leq \mu_k$, the $100(1-\alpha)\%$ simultaneous confidence intervals for $\mu_i (i = 1, 2, \ldots, k)$ were given as follows

$$\max_{l \leq i} \{\bar{Y}_l - m_{k,\nu} S/\sqrt{n_l}\} \leq \mu_i \leq \min_{j \geq i} \{\bar{Y}_j + m_{k,\nu} S/\sqrt{n_j}\},$$

where $m_{k,\nu}$ is the upper α point of the studentized maximum modulus distribution with parameters k and ν (see Miller (1981)). Under the same assumption as in Korn (1982), Schoenfeld (1986) sought confidence intervals for each individual μ_i

and also simultaneous confidence intervals for all the means. For a given mean μ_i, his upper and lower bounds are the maximum and minimum values of x such that the hypotheses $x < \mu_i$ and $\mu_i < x$ were accepted by their respective likelihood ratio tests. Schoenfeld's method for finding simultaneous confidence intervals was based on an idea by Lee (1984). Lee (1996) proposed the generalized studentized maximum modulus procedure and used it to seek a confidence band for each individual μ_i by incorporating the monotonicity of the response means. Lee's method gains much over the methods by Schoenfeld's (1986) and Korn's (1982). Marcus and Talpetz (1992) further proposed an alternative test statistic and used it to construct a set of SCB, but their procedure is inferior to the Dunnett's procedure in comparing μ_i to the control mean μ_0. Marcus and Genizi (1994) derived the simultaneous confidence intervals of normal means of the form $\sum_{i=1}^{k} c_i\mu_i$ for umbrella contrasts $c_1 \leq c_2 \leq \ldots \leq c_h \geq \ldots \geq c_k$ and $\sum_{i=1}^{k} n_i c_i = 0$, where h is called the peak of the umbrella order. Berk and Marcus (1996) summarized the results of SCB for simple order, simple tree order and umbrella order. Hayter and Liu (1999) also proposed a test statistic to develop simultaneous confidence intervals for all the ordered pairwise differences $\mu_j - \mu_i$ for $1 \leq i < j \leq h$ and $h \leq j < i \leq k$, where the peak h is known. Hayter and Liu's (1999) method is a generalization of Hayter's (1990) procedure.

When the homogeneity hypothesis is rejected in favour of the simple tree alternative, it implies that there exists at least one treatment better than the control. Let $\mu_{best} = \max_{1\leq i\leq k} \mu_i$ be the mean of the best treatment. Since $\mu_{best} - \mu_0$ is the largest difference between any treatment mean and the control mean, the confidence lower bound for $\mu_{best} - \mu_0$ is bounded below by that for any $\mu_i - \mu_0$, $(i = 1, \ldots, k)$ or their non-negative linear combinations. If this maximized confidence lower bound

for $\mu_{best} - \mu_0$ is positive, then μ_{best} is significantly larger than μ_0. The sharpest simultaneous confidence lower bound for $\mu_{best} - \mu_0$ can provide useful information regarding treatments and a control. The construction of the simultaneous confidence lower bound of $\mu_{best} - \mu_0$ is a particularly useful inference method that has not been considered before and is the main topic of this thesis.

In Chapter 2, we will introduce some basic results on least squares regression and particularly, the isotonic regression. We will present two algorithms that have been used extensively in studying and computing the isotonic regression, namely, the pool-adjacent-violators algorithm and the minimum-lower-sets algorithm. The likelihood ratio tests for testing the simple tree restriction will be given. Some advances in simultaneous confidence intervals under the simple tree order in the literature will also be included. The Kuhn-Tucker Equivalence Theorem, which will be used in Chapter 4 and Chapter 7, will also be presented.

In Chapter 3, we will propose a new test statistic to compare factor means and the control mean in two-factor experiments. The new test statistic can be inverted to yield sharp simultaneous confidence lower bounds for the differences of such means. The new test is unbiased, monotone and consistent. Its power compares favorably with Dunnett's test and the LRT.

In Chapter 4, we will study simultaneous confidence lower bounds for two cases. One, the difference between the best treatment and the control for treatments which are at least as good as the control and two, where no restriction is placed on the treatment means and the control mean. The evaluation of the simultaneous confidence lower bound for the difference between the best treatment mean and the control mean is a concave programming problem subject to homogeneous linear

inequality constraints. Two efficient computation algorithms will be derived.

In Chapter 5, we will address the cone order monotonicity problem. The Likelihood ratio test S_{01} for testing homogeneity of treatment means and the control mean $H_0 : \mu_0 = \mu_1 = \ldots = \mu_k$ with the alternative restricted by the simple tree ordering $\Omega = \{\mu \in R^{k+1} : \mu_0 \leq \mu_i \ (i = 1, \ldots, k)\}$ is not cone order monotone (COM[Ω]). In this chapter, we will propose a likelihood ratio test statistic S_{01}^o for testing homogeneity with the alternative restricted by $\Omega^o = \{\mu : \mu_0 \leq \bar{\mu} \leq \mu_i, i = 1, \ldots, k\}$, where $\bar{\mu} = \sum_{i=0}^{k} n_i\mu_i / \sum_{i=0}^{k} n_i$. LRT S_{01}^o and its modification are COM[Ω^o]. We will also offer an alternative test procedure which is COM[Ω] and COM[Ω^o] and has competitive power performance.

In Chapter 6, a test statistic based on Dunnett's procedure will be proposed for testing an interval hypothesis $H_0^I : d(\mu) \leq \delta$ vs $H_1^I : d(\mu) > \delta$ under $\mu_i \geq \mu_0$ $(i = 1, \ldots, k)$, where the standardized difference of the means between the best treatment and the control is measured by $d(\mu) = (\mu_{best} - \mu_0)/\sigma$, The quantity $d(\mu)$ can be used to measure the dispersion among treatment means and the control mean and is useful for assessing the equivalence of the treatment means and the control mean. Numerical quadrature will be employed to obtain the tabulated percentage points of the test statistic for testing H_0^I versus H_1^I. We will also show how to construct a confidence lower bound for $d(\mu)$.

In Chapter 7, we will consider the problem of identifying the minimum effective dose in dose-response studies. Assessing monotone dose-response relationship is frequently encountered in practice in the context of actively proving a significant monotonous dependence of the reponse on increasing doses or treatments. But the monotone dose-response assumption is not always satisfied. In some situations

there may be a negative response at low doses, then it is appropriate to make a partial monotone dose-response assumption such as $\mu_q \leq \mu_{q+1} \leq \ldots \leq \mu_k$, here q is prespecified through an experimenter's prior knowledge. The popular Williams' trend test is not applicable in this case. We will propose a multiple contrast test and use it to develop a stepwise method which is more powerful than Hsu and Berger's DR method under the partial monotone dose-response assumption.

Finally, Chapter 8 will provide a discussion of the results of this thesis and further areas of research relating to the problem.

Chapter 2

Statistical Inference Under Order Restrictions

2.1 Introduction

Isotonic regression problem arises from the maximum likelihood estimation of normal means under an order restriction and it plays a very important role in the order restricted inference. Its usefulness is greatly enhanced by the fact that it solves a wide variety of restricted estimation problems in which the objective function may take many different forms other than the sum of squares. Its application includes maximum likelihood estimation of ordered normal variances, ordered binomial parameters (bioassay), ordered Poisson means, ordered multinomial parameters as well as a variety of problems from other areas, such as reliability theory and density estimation. It is also well known that test procedures such as the likelihood ratio tests which utilize the prior ordering information dominate procedures which ignore this information. Moreover, simultaneous confidence intervals for restricted setting can be much shorter than that of Scheffé's procedure. In addition, solutions to many other optimization problems can be expressed in terms of isotonic regression, see

Dykstra and Lee (1991).

The problem of developing algorithms for isotonic regression has received a great deal of attention. In fact, isotonic regression is a quadratic programming problem and there is an extensive literature on methods for computing solutions. The problem of computing the isotonic regression is a special case and a number of efficient algorithms have been proposed.

The most widely used algorithm for simple order is the pool-adjacent-violators algorithm (PAVA) first published by Ayer, Brunk, Ewing, Reid and Silverman (1955). PAVA can be used to develop algorithms for some other least squares problem as we shall see in Chapter 7. PAVA is a very efficient algorithm but it does not apply in general to partially ordered isotonic regression. For general partially ordered isotonic regression the most well known algorithm is the minimum-lower-sets algorithm of Brunk (1955). Several other algorithms have been developed for partial orders to increase the efficiency of the computation, such as the minimum violator algorithm due to Thompson (1962), the min-max algorithm due to Lee (1983), the minimum lower sets algorithm of Qian (1992), among others.

Many of the methods of statistical inference are derived from the experimental situation in which one wishes to compare several normal treatments with a normal control or standard. In many cases it is believed that the treatments are at least as effective as the control. Procedures have been developed to make use of this prior information in statistical inference. For instance, if the control and treatment populations are normal with a common unknown variance, then Dunnett's (1955) one-sided procedures can be used to test homogeneity of the means versus a simple tree alternative. Bartholomew's (1959, 1961) $\bar{\chi}^2$ or $\bar{E}^2$ test is the likelihood ratio

test of homogeneity with a simple tree alternative. Raubertas, Lee, and Nordheim (1986) generalized Bartholomew's tests to allow hypotheses involving homogeneous linear inequality restrictions. If the $\bar{\chi}^2$ and $\bar{E}^2$ test rejects the homogeneity null hypothesis then one may wish to carry out a multiple comparison procedure to determine which means are different or which simultaneous confidence bounds may be of interest.

In Section 2.2 we first review some concepts and preliminary results of projections on closed convex cones. Concepts of partial order such as simple tree order and isotonic regressions are given in Section 2.3. We also present the above mentioned two algorithms in Section 2.4. In Section 2.5, we focus on the test of hypotheses involving simple tree order with likelihood ratio tests or their modifications. In Section 2.6, we give a brief review regarding simultaneous confidence intervals under simple tree order. Since the evaluation of the simultaneous confidence lower bounds under simple tree order is a maximization problem subject to some constraints, in Section 2.7 we present the Kuhn-Tucker Equivalence Theorem.

2.2 Projections on Closed Convex Cones

2.2.1 Convex Sets, Cones and Dual Cones

Let R^k be a k-dimensional Euclidean space with the inner product defined by

$$< f, g >_{\mathbf{w}} = \sum_{i=1}^{k} f_i g_i w_i, \quad \forall f, g \in R^k. \tag{2.1}$$

where $\mathbf{w} = (w_1, \ldots, w_k)$ is a vector of weights such that $w_i > 0$, $i = 1, 2, \ldots, k$ and $\sum_{i=1}^{k} w_i = 1$. If $w_1 = \ldots = w_k = 1/k$, we omit the subscription and use $< f, g >= \sum_{i=1}^{k} f_i g_i$. A subset C of R^k is said to be *convex* if $(1 - \lambda) f + \lambda g \in C$

whenever $f, g \in C$ and $0 \leq \lambda \leq 1$. It is well known that the intersection of an arbitrary number of convex sets is still convex. A subset $\mathcal{C}$ of R^k is called a *cone* if it is closed under nonnegative scalar multiplication, i.e., $\lambda f \in \mathcal{C}$ when $f \in \mathcal{C}$ and $\lambda \geq 0$. Note that a cone is not necessarily "pointed." For example, subspaces of R^k are special cones. So are the open and closed half-spaces corresponding to a hyperplane containing the origin.

For a convex cone $\mathcal{C}$, the subset $\mathcal{C}^*$ of R^k defined by

$$\mathcal{C}^* = \{g \in R^k : < g, f >_w = \sum_{i=1}^{k} g_i f_i w_i \leq 0, \quad \forall f \in \mathcal{C}\}. \tag{2.2}$$

is called the *Fenchel dual* or *polar* of $\mathcal{C}$. In particular, if $\mathcal{C} = S$ is a subspace of R^k, then

$$S^* = S^{\perp} = \{g \in R^k : < g, f >_w = 0, \forall f \in S \subset R^k\}. \tag{2.3}$$

It can be shown that $\mathcal{C}^*$ is also a convex cone and furthermore, it is closed. For any two subsets A, B of R^k, denote $A + B$ the direct sum of sets A, B, i.e., $A + B = \{f + g | f \in A, g \in B\}$. Let $\mathcal{C}$, $\mathcal{C}_1$ and $\mathcal{C}_2$ be convex cones. We have the following results,

(a) $\mathcal{C} \subset (\mathcal{C}^*)^*$, and $\mathcal{C} = (\mathcal{C}^*)^*$ if $\mathcal{C}$ is closed: (2.4)

(b) $(-\mathcal{C})^* = -\mathcal{C}^*$; (2.5)

(c) $\mathcal{C}_1^* \subset \mathcal{C}_2^*$ if $\mathcal{C}_1 \supset \mathcal{C}_2$ (2.6)

(d) $(\mathcal{C}_1 + \mathcal{C}_2)^* = \mathcal{C}_1^* \cap \mathcal{C}_2^*$; (2.7)

(e) $(\mathcal{C}_1 \cap \mathcal{C}_2)^* = \mathcal{C}_1^* + \mathcal{C}_2^*$ if the latter is closed, (2.8)

see Rockafellar (1970, page 146).

2.2.2 Properties of Projections on Closed Convex Cones

For any closed convex cone $\mathcal{C} \subset R^k$ and a given vector $g \in R^k$, the solution to the problem

$$\text{Minimize } \{\sum_{i=1}^{k}(g_i - f_i)^2 w_i\} \tag{2.9}$$

subject to $f \in \mathcal{C}$ uniquely exists. This unique solution, denoted by $P_{\mathbf{w}}(g|\mathcal{C})$, is called the least squares projection of $\mathbf{g}$ onto $\mathcal{C}$ with weight $\mathbf{w}$.

Utilizing Theorem 8.2.7 of Robertson, Wright and Dykstra (1988), we now state as the following theorem.

Theorem 2.2.1 *Let $\mathcal{C}$ be a closed convex cone in R^k, and let $g, g^* \in R^k$. Then $g^* = P_{\mathbf{w}}(g|\mathcal{C})$ if and only if $g^* \in \mathcal{C}$ and*

$$< g - g^*, g^* >_{\mathbf{w}} = 0, \tag{2.10}$$

and

$$< g - g^*, f >_{\mathbf{w}} \leq 0. \quad \forall f \in \mathcal{C}. \tag{2.11}$$

2.2.3 Polyhedral Cones

Let $\mathcal{K}$ be a cone in R^k. $\mathcal{K}$ is a *polyhedral cone* if $\mathcal{K} = \{\mathbf{x} \in R^k : \ < \mathbf{b_i}, \mathbf{x} > \geq 0, i = 1, 2, \ldots, m\} = \{\mathbf{x} \in R^k : \mathbf{Bx} \geq \mathbf{0}\}$. Here B is the $m \times k$ matrix whose ith row is $\mathbf{b_i}$, while $-\mathbf{b_i}$ is called the generator of its *Fenchel dual* $\mathcal{K}^*$. We will tacitly assume that the set of generators $\{\mathbf{b_i}, i = 1, 2, \ldots, m\}$ are non-redundant, i.e., no proper subset of the set of $\mathbf{b_i}$ determines $\mathcal{K}^*$. For the simple tree cone $\mathcal{K} = \{\mathbf{x} : x_0 \leq x_i, i = 1, 2, \ldots, k\}$, $\mathbf{b_i} = (b_{i0}, b_{i1}, \ldots, b_{ik}) = (-1, 0, \ldots, 0, 1, 0, \ldots, 0), i = 1, 2, \ldots, k$, where $b_{i,i} = 1$.

2.3 Partial Order of Finite Sets and Isotonic Regressions

2.3.1 Partial Order and Simple Tree Order

Let $\mathcal{X}$ be a finite set $\{x_1, x_2, \ldots, x_k\}$. A binary relation $\preceq$ on $\mathcal{X}$ is a *partial order* on $\mathcal{X}$ if

1. it is *reflexive*: $x \preceq x$ for $x \in \mathcal{X}$;

2. it is *transitive*: $x, y, z \in \mathcal{X}$, $x \preceq y$ and $y \preceq z$ imply $x \preceq z$;

3. it is *antisymmetric*: $x, y \in \mathcal{X}$, $x \preceq y$ and $y \preceq x$ imply $x = y$.

A binary relation $\preceq$ on $\mathcal{X}$ is called a *simple order* if it is reflexive, transitive, antisymmetric and

4. it is *comparable*: $x, y \in \mathcal{X}$ implies that either $x \preceq y$ or $y \preceq x$.

Let $\mathcal{X} = \{x_0, x_1, \ldots, x_k\}$ and define the partial order $\preceq$ on $\mathcal{X}$ by $x_0 \preceq x_i$ $(i = 1, 2, \ldots, k)$ with no relationship between x_i and x_j for $i, j \geq 1$. This partial order restriction is called the *simple tree order*.

Note that there may be noncomparable elements for a partial order. A partial order usually arises when vector comparisons are involved. Simple order and simple tree order $x_0 \preceq x_i$ $(i = 1, 2, \ldots, k)$ are partial orders that are frequently encountered in applications. As an important partial order, the simple tree order arises in sampling situations where one wishes to compare several treatments with a control or a standard making use of prior information that all of the treatment means are at least as large as the control mean. (The case in which all of the treatment means

are no larger than the control mean is included by changing the signs of all the means.) For example, survival times of different therapies are at least as large as that of a standard therapy or placebo; increasing dose levels are frequently expected to produce stronger than or at least equal effects as a zero-dose control.

2.3.2 Isotonic Regression

A real-valued function f on a finite set $\mathcal{X}$ is said to be *isotonic* with respect to the partial ordering $\preceq$ on $\mathcal{X}$ if $x, y \in \mathcal{X}$ and $x \preceq y$ imply $f(x) \leq f(y)$.

Let g be a given function on $\mathcal{X}$ and w a given positive weight function on $\mathcal{X}$. An isotonic function g^* on $\mathcal{X}$ is called an *isotonic regression* of g with weight w if it minimizes

$$\sum_{x \in \mathcal{X}} [g(x) - f(x)]^2 w(x)$$

for all functions f on $\mathcal{X}$ which are isotonic.

A real-valued function on a finite set $\mathcal{X}$ can be considered as a point of a Euclidean space which has as its dimension the number of points in $\mathcal{X}$. In this setting, the collection, $\mathcal{I}$, of all isotonic functions on $\mathcal{X}$ with respect to a given partial order is a closed convex cone and the isotonic regression g^* is the closest point of $\mathcal{I}$ to g with distance induced by the inner product

$$< f, g >_{\mathbf{w}} = \sum_{i=1}^{k} f_i g_i w_i,$$

in other words, $g^* = P_{\mathbf{w}}(g|\mathcal{I})$. The existence and uniqueness then follow from the general theory of projection on closed convex cones described earlier in this chapter.

2.3.3 Properties of Isotonic Regression

The isotonic regression has a number of important properties. Some of them are given below.

Theorem 2.3.1 *Suppose g_1 and g_2 are isotonic functions on X such that $g_1(x) \leq g(x) \leq g_2(x)$ for all $x \in X$, and if g^* is an isotonic regression of g, then also $g_1(x) \leq g^*(x) \leq g_2(x)$ for all $x \in X$. In particular, if a and b are constants such that $a \leq g(x) \leq b$ for all $x \in X$, then also $a \leq g^*(x) \leq b$ for $x \in X$. (Th. 1.3.4. Robertson, Wright, and Dykstra (1988))*

Suppose g and w are functions on X, the weighted average of g over the nonempty subset A of X is defined as follows

$$\mathrm{Av}(A) = \frac{\sum_{x \in A} w(x) g(x)}{\sum_{x \in A} w(x)}.$$

While $\mathrm{Av}(A)$ depends on g, this is not explicit in the notation. Let $[g^* = c]$ denote $\{x \in X : g^*(x) = c\}$.

Theorem 2.3.2 *If c is any real number and if the set $[g^* = c]$ is nonempty, then $c = Av([g^* = c])$. (Th 1.3.5, Robertson, Wright, and Dykstra, R. L. (1988))*

Theorem 2.3.3 *For an arbitrary real-valued function, Ψ, defined on the reals,*

$$< g - g^*, \Psi(g^*) >= 0.$$

(Th 1.3.6, Robertson, Wright, and Dykstra (1988))

Theorem 2.3.2 reduces the problem of computing g^* to finding the sets on which g^* is constant (i.e. its *level sets*). There are a number of algorithms in computing

isotonic regressions and we will introduce two of them in the next section that have been extensively used, namely the *pool-adjacent-violators algorithm* (PAVA) for simple order and the *minimum-lower-sets algorithm* for any partial order such as the simple tree order.

2.4 Algorithms for Isotonic Regression

2.4.1 Pool-Adjacent-Violators Algorithm

Let X be a finite set $\{x_1, x_2, \ldots, x_k\}$ with a simple order $x_1 \preceq x_2 \preceq \ldots \preceq x_k$. Then a real valued function f on X is isotonic if and only if $f(x_1) \leq f(x_2) \leq \ldots \leq f(x_k)$. Let g be a given function on X and w a given positive weight function on X. The PAVA starts with g. If g is isotonic, then $g^* = g$. Otherwise, there must exist an index i such that $g(x_{i-1}) > g(x_i)$. These two values are then replaced by their weighted average, namely $\mathrm{Av}(\{i-1, i\})$ and the two weights $w(x_{i-1})$ and $w(x_i)$ are replaced by $w(x_{i-1}) + w(x_i)$. If this new set of $k-1$ values is isotonic, then $g^*(x_{i-1}) = g^*(x_i) = \mathrm{Av}(\{i-1, i\})$ and $g^*(x_j) = g(x_j)$ otherwise. If this new set is not isotonic then this process is repeated using the new values and weights until an isotonic set of values is obtained.

2.4.2 Minimum Lower Sets Algorithm

A subset L of X is called a lower set with respect to the partial order $\preceq$ if $y \in L$ and $x \preceq y$ imply $x \in L$. A subset U of X is called a upper set with respect to the partial order $\preceq$ if $x \in U$ and $x \preceq y$ imply $y \in U$. We denote the class of all lower sets by $\mathcal{L}$ and the class of all upper sets by $\mathcal{U}$. A subset B of X is a level set if and only if there exists a lower set L and an upper set U such that $B = L \cap U$. The

minimum lower sets algorithm for isotonic regression is described next.

Select a lower set L_1 such that $Av(L_1) \leq Av(L)$ for all lower sets L. Suppose L_1' is another lower set having this property. Using the property of Av which is a strict Cauchy mean value function, $L_1 \cup L_1'$ is another lower set of minimum average. Therefore, the union of all lower sets of minimum average is the largest lower set of minimum average. Let L_1, and also B_1, denote this lower set. This level set is the set on which g^* assumes its smallest value:

$$g^*(x) = Av(B_1) = \min\{Av(L) : L \in \mathcal{L}\}$$

for $x \in B_1$. Now consider the averages of level sets of the form $L \cap L_1^c$, level sets consisting of lower sets with L_1 subtracted. Select again the largest of these level sets of minimum average, say $B_2 = L_2 \cap L_1^c$. The level set B_2 is the set on which g^* assumes its next smallest value:

$$g^*(x) = Av(B_2); \ x \in B_2.$$

This process is continued until X is exhausted.

In order to illustrate the above algorithm, let us consider the simple tree defined on $X = \{x_0, x_1, \ldots, x_k\}$ with $\preceq$ by $x_0 \preceq x_i, i = 1, 2, \ldots, k$. The nonempty lower sets consist of $\{x_0\}$ and $\{x_0\} \cup A$ with A any nonempty subset of $\{x_1, x_2, \ldots, x_k\}$. Thus there are 2^k nonempty lower sets. On the other hand, if an upper set contains x_0, it must contain $\{x_1, x_2, \ldots, x_k\}$ as a subset. Thus, there are 2^k nonempty upper sets consisting of X and all of the nonempty subsets of $\{x_1, x_2, \ldots, x_k\}$. The minimum lower sets algorithm applied to the simple tree order yields the following algorithm for computing the isotonic regresion g^* of a given function with weights

$w(w > 0)$. If $g(x_0) \leq g(x_i), i = 1, 2, \ldots, k$, then $g^* = g$. Otherwise, arrange the values $g(x_1), g(x_2), \ldots, g(x_k)$ in ascending order ($g(x_0)$ is not included). Denote these values by $g(x_{(1)}) \leq g(x_{(2)}) \leq \ldots g(x_{(k)})$ and let $w(x_{(i)})$ denote the weight corresponding to $g(x_{(i)})$ for $i = 1, 2, \ldots, k$. Next, find the smallest positive integer j for which

$$A_j = \frac{w(x_0)g(x_0) + \sum_{i=1}^{j} w(x_{(i)})g(x_{(i)})}{w(x_0) + \sum_{i=1}^{j} w(x_{(i)})} < g(x_{(j+1)})$$

Such an integer will exist unless $A_{k-1} \geq g(x_{(k)})$ and in this case set $j = k$. Now $g^*(x_0) = A_j$ and the value of $g^*(x_i)$ is either A_j or $g(x_i)$ depending upon which one is larger.

2.5 Testing the Simple Tree Order

In this section, the likelihood ratio tests (LRTs) for homogeneity of normal means with the simple tree alternative are introduced. If the simple tree order imposed on the alternative is in question, one may wish to test this order restriction as the null hypothesis with an unrestricted alternative.

Let $X = \{0, 1, 2, \ldots, k\}$ and assume that the simple tree order $\preceq$ is defined on X. Let μ_i be the mean of the ith normal population with variance σ_i^2 for $i = 0, 1, 2, \ldots, k$. We are interested in the following hypotheses

$$H_0 : \quad \mu_0 = \mu_1 = \ldots = \mu_k,$$

$$H_1 : \quad \mu_0 \leq \mu_i \text{ for all } i = 1, 2, \ldots, k.$$

and

$$H_2 : \text{ no restrictions on the means.}$$

Suppose that Y_{ij} is a normally distributed random variable with unknown mean μ_i and variance of the form $\sigma_i^2 = a_i\sigma^2$ with $a_0, a_1, \ldots, a_k$ known and σ^2 unkown for

$i = 0, 1, 2, \ldots, k$ and $j = 1, 2, \ldots, n_i$, also assume that the Y_{ij} are independent. For the case of equal but unknown variances, one would set $a_0 = a_1 = \ldots = a_k = 1$. Here index 0 refers to a control and index $i (i = 1, \ldots, k)$ refers to treatment i. Suppose also that S^2 is an estimator for σ^2 which is independent of $\bar{\mathbf{Y}} = (\bar{Y}_0, \bar{Y}_1, \bar{Y}_2, \ldots, \bar{Y}_k)$ with $\frac{\nu S^2}{\sigma^2} \sim \chi^2_\nu$ and $\nu = \sum_{i=0}^{k} n_i - (k+1) > 0$ (χ^2_ν denotes a chi-squared variable with ν degrees of freedom). With $w_i = n_i/a_i$ for $i \in N$, the maximum likelihood estimate $\hat{\mu}$ of μ_i under H_0 is $\sum_{i=0}^{k} w_i \bar{Y}_i / \sum_{i=0}^{k} w_i$. The restricted maximum likelihood estimate of $\boldsymbol{\mu}$ subject to H_1 is denoted by $\boldsymbol{\mu}^* = (\mu_0^*, \mu_1^*, \ldots, \mu_k^*)$. It is the isotonic regression of $\bar{\mathbf{Y}} = (\bar{Y}_0, \ldots, \bar{Y}_k)$ under H_1 with weight $\mathbf{w} = (w_0, w_1, \ldots, w_k)$. The unrestricted maximum likelihood estimate of $\boldsymbol{\mu}$ is $\bar{\mathbf{Y}}$.

In many experiments, one often has a priori knowledge that the treatments are at least as effective as the control, that is, one can assume that the treatments may be beneficial, but they are certainly not detrimental. Then the objective of the experiment is to determine if at least one of the treatments is more effective than the control, and the one-sided test of H_0 against H_1, but not H_0, is more appropriate. The likelihood ratio test (LRT) rejects H_0 in favour of $H_1 - H_0$ for large values of

$$\bar{E}^2_{01} = \frac{\sum_{i=o}^{k} w_i(\mu_i^* - \hat{\mu})^2}{\sum_{i=0}^{k} w_i(\bar{Y}_i - \hat{\mu})^2 + \nu S^2}, \tag{2.12}$$

and if σ^2 is known, the LRT of H_0 versus $H_1 - H_0$ rejects H_0 in favour of $H_1 - H_0$ for large values of

$$\bar{\chi}^2_{01} = \sum_{i=0}^{k} w_i(\mu_i^* - \hat{\mu})^2/\sigma^2. \tag{2.13}$$

For testing H_1 versus H_2, the LRT rejects H_1 for large values of

$$\bar{E}^2_{12} = \frac{\sum_{i=o}^{k} w_i(\bar{Y}_i - \mu_i^*)^2}{\sum_{i=0}^{k} w_i(\bar{Y}_i - \mu_i^*)^2 + \nu S^2}, \tag{2.14}$$

and if σ^2 is known, the LRT of H_1 versus H_2 rejects H_1 for large values of

$$\bar{\chi}^2_{12} = \sum_{i=0}^{k} w_i(\bar{Y}_i - \mu_i^*)^2/\sigma^2. \tag{2.15}$$

Using the properties of projection, it can be shown that

$$\sum_{i=0}^{k} w_i(\bar{Y}_i - \hat{\mu})^2/\sigma^2 = \bar{\chi}^2_{01} + \bar{\chi}^2_{12}.$$

Accordingly, with $Q(\nu) = \nu S^2/\sigma^2 \sim \chi^2_\nu$, $\bar{E}^2_{01} = \bar{\chi}^2_{01}/\left(\bar{\chi}^2_{01} + \bar{\chi}^2_{12} + Q(\nu)\right)$. Therefore, rejecting for large values of $\bar{E}^2_{01}$ is equivalent to rejecting for large values of

$$S_{01} = \frac{\nu \bar{E}^2_{01}}{1 - \bar{E}^2_{01}} = \frac{\nu \bar{\chi}^2_{01}}{\bar{\chi}^2_{12} + Q(\nu)}. \tag{2.16}$$

A more straightforward approach to testing H_0 versus $H_1 - H_0$ is to replace σ^2 by S^2 in (2.16), which is proposed by Wright (1988). Wright called it the modified likelihood ratio test (MLRT). The MLRT rejects H_0 for large values of

$$T_{01} = \sum_{i=0}^{k} \frac{w_i(\mu_i^* - \hat{\mu})^2}{S^2} \tag{2.17}$$

Dividing the numerator and the denominator of (2.16) by ν, noting that for each ω in the underlying probability space $\bar{\chi}^2_{12}/\nu \to 0$ as $\nu \to \infty$, and using $Q(\nu)/\nu \to 1$ in probability as $\nu \to \infty$, we see that $S_{01} \to \bar{\chi}^2_{01}$ in probability as $\nu \to \infty$. Again applying the consistency of S^2 for σ^2, it follows that $T_{01} \to \bar{\chi}^2_{01}$ in probability as $\nu \to \infty$. Hence, the LRT and MLRT are asymptotically ($\nu \to \infty$) equivalent. The LRT of H_1 versus $H_2 - H_1$ rejects H_1 for large values of

$$\bar{E}^2_{12} = \frac{\bar{\chi}^2_{12}}{\bar{\chi}^2_{12} + Q(\nu)}. \tag{2.18}$$

or equivalently for large values of

$$S^2_{12} = \frac{\bar{\chi}^2_{12}}{Q(\nu)/\nu}. \tag{2.19}$$

Note that S_{12}^2 converges in distribution to $\bar{\chi}_{12}^2$ as $\nu \to \infty$. It is more convenient to table the critical values for S_{01} and S_{12}. The null distributions of $\bar{\chi}_{01}^2$, $\bar{\chi}_{12}^2$, S_{01}, S_{12} and T_{01} are given by the following theorem.

Theorem 2.5.1 *For $\mu \in H_0$ and ν a positive integer.*

$$P[\bar{\chi}_{01}^2 \geq c] = \sum_{l=2}^{k-1} P(l, k+1; \mathbf{w}) P\left[\chi_{l-1}^2 \geq c\right],$$

$$P[\bar{\chi}_{12}^2 \geq c] = \sum_{l=1}^{k} P(l, k+1; \mathbf{w}) P\left[\chi_{k-1-l}^2 \geq c\right].$$

$$P[S_{01} \geq c] = \sum_{l=2}^{k-1} P(l, k+1; \mathbf{w}) P\left[F_{l-1,N-l} \geq \frac{c(N-l)}{\nu(l-1)}\right],$$

$$P[S_{12} \geq c] = \sum_{l=1}^{k} P(l, k+1; \mathbf{w}) P\left[F_{k+1-l,\nu} \geq \frac{c}{k+1-l}\right],$$

$$P[T_{01} \geq c] = \sum_{l=2}^{k+1} P(l, k+1; \mathbf{w}) P\left[F_{l-1,\nu} \geq \frac{c}{l-1}\right]$$

for any $c > 0$, where $\nu = \sum_{i=0}^{k} n_i - (k+1)$.

The mixture coefficient $P(l, k+1; \mathbf{w})$ in the above theorem is called the level probability. It is defined as the probability that there are exactly l distinct values for the MLE $\boldsymbol{\mu}^*$ satisfying the simple tree order when H_0 is true. The values of $P(l, k+1; \mathbf{w})$ depend on the sample sizes n_i and the population variances through the weight $\mathbf{w}$. There is a recursive formula for computing $P(l, k+1; \mathbf{w})$ in Theorem 2.4.1 of Robertson, Wright and Dykstra (1988). If the weights are equal, the level probabilities for the simple tree ordering are less complex. For the equal-weight case,

suppress the $\mathbf{w}$, denote the level probabilties by $P(l, k+1)$. Using that recursive formula, one can show that

$$P(l, k+1) = \binom{k}{l-1} P(l, k-l+2) P\big(l, l; (k-l+2), 1, \ldots, 1\big). \tag{2.20}$$

The last factor in the above expression has been tabulated by Ruben (1954) for l and k with $0 \le k \le 49$ and $k-9 \le l \le k+1$. Numerical values of $P(l, k+1)$ with $k \le 19$ are given in Table A.11 of Robertson, Wright and Dykstra (1988). Wright and Tran (1985) discussed approximate procedures for $P(l, k+1; \mathbf{w})$ with unequal weights.

2.6 Simultaneous Interval Estimations for Treatments Versus a Control

Simultaneous interval estimations for treatments versus a control, particularly simultaneous one-sided confidence bounds (SCB) for various classes of contrast between normal means $\sum_{i=0}^{k} n_i c_i \mu_i$ with $\sum_{i=0}^{k} n_i c_i = 0$, have been inverted from test procedures used for testing H_0 versus $H_1 - H_0$. Assume that $\bar{Y}_i, i = 0, 1, \ldots, k$, are normal variates with mean μ_i and variance σ^2/n_i. Dunnett (1955) obtained SCB for all many-one contrasts with $c_i \ge 0$ $(i = 1, 2, \ldots, k)$ and $\sum_{i=1}^{k} n_i c_i = 1$, which is

$$\sum_{i=0}^{k} n_i c_i \mu_i \ge \sum_{i=0}^{k} n_i c_i \bar{Y}_i - d^{\alpha}_{k,\nu} S \sqrt{\frac{1}{n_0} + \frac{1}{n_i}}, \tag{2.21}$$

where $d^{\alpha}_{k,\nu}$ is the critical value of Dunnett's test statistic. Marcus (1978) employed $\bar{\chi}^2_{01}$ to provide SCB for all simple tree contrasts with coefficients satisfying $c_0 \le c_i$ $(i = 1, 2, \ldots, k)$ which is

$$\sum_{i=0}^{k} n_i c_i \mu_i \ge \sum_{i=0}^{k} n_i c_i \bar{Y}_i - \sigma t_{\alpha} \Big(\sum_{i=0}^{k} n_i c_i^2 \Big)^{1/2} \tag{2.22}$$

where t_α is the positive square root of the critical value for $\bar{\chi}^2_{01}$ under the simple tree order. Marcus (1978) also used $\bar{\chi}^2_{12}$ to build SCB for all contrasts with coefficients satisfying $c_i \geq 0$ $(i = 1, 2, \ldots, k)$ which is

$$\sum_{i=0}^{k} n_i c_i \mu_i \geq \sum_{i=0}^{k} n_i c_i \bar{Y}_i - \sigma \tilde{t}_\alpha \Big(\sum_{i=0}^{k} n_i c_i^2 \Big)^{1/2} \tag{2.23}$$

where $\tilde{t}_\alpha$ is the positive square root of the critical value for $\bar{\chi}^2_{12}$ under the simple tree order. Marcus and Talpaz (1992) proposed the test statistic

$$M_k = \frac{\sqrt{n}(\max_{1 \leq i \leq k} \mu_i^* - \mu_0^*)}{S}$$

to test H_0 versus $H_1 - H_0$ and used it to form the following simultaneous one-sided confidence bounds

$$\sum_{i=0}^{k} n_i c_i \mu_i \geq \sum_{i=0}^{k} n_i c_i \bar{Y}_i - g_{k,\nu,\alpha} S n^{-1/2} \Big(\sum_{i=0}^{k} |c_i| \Big) \tag{2.24}$$

where $c_0 \leq c_i$ $(i = 1, \ldots, k)$, and $g_{k,\nu,\alpha}$ is the upper αth percentile of the test statistic M_k.

With the prior knowledge of the simple tree order, the lower bound of $\sum_{i=0}^{k} n_i c_i^* \mu_i$ can be improved to

$$\max \Big\{ \sum_{i=0}^{k} n_i c_i \bar{Y}_i - \sigma t_\alpha \Big(\sum_{i=0}^{k} n_i c_i^2 \Big)^{1/2} \Big\} \tag{2.25}$$

subject to $\sum_{i=0}^{k} n_i c_i \mu_i \leq \sum_{i=0}^{k} n_i c_i^* \mu_i$, $\sum_{i=0}^{k} n_i c_i = 0$, $\sum_{i=0}^{k} n_i c_i^* = 0$, and $\mu_0 \leq \mu_i$ $(i = 1, \ldots, k)$, here $\mathbf{c}^*$ is a given vector. One can also similarly get the improved lower bound from (2.23). When $\sum_{i=0}^{k} n_i c_i^* \mu_i = \mu_{best} - \mu_0$, here $\mu_{best} = \max_{1 \leq i \leq k} \mu_i$, the improved lower bound is of particular interest and has not been studied in the literature. This thesis will explore it in Chapter 4.

2.7 Kuhn-Tucker Equivalence Theorem

The evaluation of the simultaneous confidence lower bounds such as (2.25) is a maximization problem subject to a mixture of equality and inequality constraints. To be specific, let $\mathbf{x}$ be an $n \times 1$ vector and $H(\mathbf{x})$ be an $m \times 1$ vector whose components $h_1(\mathbf{x}), \ldots, h_m(\mathbf{x})$ are differentiable concave functions for $\mathbf{x} \geq \mathbf{0}$. Let $g(\mathbf{x})$ be another differentiable concave function. The Kuhn-Tucker Equivalence Theorem enables us to find an $\mathbf{x}^o$ that maximizes $g(\mathbf{x})$ constrained by $H(\mathbf{x}) \geq \mathbf{0}$ and $\mathbf{x} \geq \mathbf{0}$. A vector $\mathbf{x}$ is said to be feasible if $\mathbf{x}$ satisfies all the contraints. The optimal value of the problem is the maximum of $g(\mathbf{x})$ over the sets of feasible points. Those feasible points which attain the optimal value are called optimal solutions. Let $\frac{\partial \phi}{\partial \mathbf{x}^o}$ and $\frac{\partial \phi}{\partial \mathbf{u}^o}$ denote the partial derivative evaluated at a specific point $\mathbf{x}^o$ and $\mathbf{u}^o$, respectively.

Theorem 2.7.1 *(Theorem 3 in Kuhn and Tucker (1951)) Let $h_1(\mathbf{x}), \ldots, h_m(\mathbf{x}), g(\mathbf{x})$ be concave and differentiable for $\mathbf{x} \geq \mathbf{0}$. Let $\phi(\mathbf{x}, \mathbf{u}) = g(\mathbf{x}) + \mathbf{u}'H(\mathbf{x})$. Then $\mathbf{x}^o$ maximizes $g(\mathbf{x})$ constrained by $H(\mathbf{x}) \geq \mathbf{0}$ and $\mathbf{x} \geq \mathbf{0}$ if and only if $\mathbf{x}^o$ and $\mathbf{u}^o$ satisfy the following conditions:*

(1) $\frac{\partial \phi}{\partial \mathbf{x}^o} \leq \mathbf{0}$, $[\frac{\partial \phi}{\partial \mathbf{x}^o}]'\mathbf{x}^o = 0, \mathbf{x}^o \geq \mathbf{0}$;

(2) $\frac{\partial \phi}{\partial \mathbf{u}^o} \geq \mathbf{0}$, $[\frac{\partial \phi}{\partial \mathbf{u}^o}]'\mathbf{u}^o = 0, \mathbf{u}^o \geq \mathbf{0}$.

When the constraints $H(\mathbf{x}) \geq \mathbf{0}, \mathbf{x} \geq \mathbf{0}$ are changed to the following three cases, some modifications are needed:

Case 1 : $H(\mathbf{x}) \geq \mathbf{0}$.

In this case , letting $\phi(\mathbf{x}, \mathbf{u}) = g(\mathbf{x}) + \mathbf{u}'H(\mathbf{x})$ defined for all $\mathbf{x}$ and constrained only by $\mathbf{u} \geq \mathbf{0}$, condition (1) should be replaced by (1*) $\frac{\partial \phi}{\partial \mathbf{x}^o} = \mathbf{0}$.

Case 2 : $H(\mathbf{x}) = \mathbf{0}, \mathbf{x} \geq \mathbf{0}$.

In this case , letting $\phi(\mathbf{x}, \mathbf{u}) = g(\mathbf{u}'H(\mathbf{x})$ defined for all $\mathbf{u}$ and constrained only by $\mathbf{x} \geq \mathbf{0}$, condition (2) should be replaced by (2^*) $\frac{\partial \phi}{\partial \mathbf{u}'} = \mathbf{0}$.

Case 3 : $H(\mathbf{x}) = \mathbf{0}$.

In this case , letting $\phi(\mathbf{x}, \mathbf{u}) = g(\mathbf{x}) + \mathbf{u}'H(\mathbf{x})$ defined for all $\mathbf{u}$ and $\mathbf{x}$ without any constraints, conditions (1) and (2) should be replaced by (1^*) and (2^*). This case corresponds to the usual method of Lagrange multipliers.

Chapter 3

A Multiple Comparisons Procedure for Detecting Differences Between Treatments and a Control in Two-factor Experiments

3.1 Introduction

Many situations in pharmaceutical research and other fields require comparing several treatment means with a control mean, or a standard. A number of statistical procedures have been proposed for applications involving treatments versus a control, of which the best known is Dunnett's (1955) multiple comparisons procedure. Many generalizations of Dunnett's procedure have been made. Shaffer (1977) extended Dunnett's procedure to yield simultaneous confidence intervals for all linear contrasts among the k treatment means and the control mean which are shorter than the intervals obtained by using the Tukey and Scheffé methods. Dunnett and Tamhane (1991) generalized it to unbalanced one-way layout in step-down fashion.

Liu (1996) developed a group sequential procedure for comparing several treatments with a control. The aforementioned results only involve one-way layouts. Cheung and Holland (1991 and 1992) generalized Dunnett's procedure to make comparisons simultaneously in each of r independent groups with each group consisting of some treatments and a control. Usually in multi-factor experiments it is of interest to make inference for factor means such as Bechhofer and Dunnett (1987). Hence, comparison of combined treatment means with the control mean is of interest. Miller (1981) suggested using weighted average of pairwise comparisons. Cochran and Cox (1957) compared the average effect of sulphur with the control in a scab index data. In that experiment, researchers would be interested in the effectiveness of the spring application or the fall application. Notice that Cheung and Holland's procedures (1991 and 1992) do not apply to this case.

We consider two-way no presence of interation models with I levels of Factor A and J levels of Factor B, and a control or a standard, labeled 0. The scab index data in Cochran and Cox (1957) and the experiment of antidotes effectiveness for the sown species in Bofinger and Mengersen (1988) are typical applications. We make the usual assumptions for the analysis of variance, that independent observations $X_{01}, \ldots, X_{0n_0}$ from the control group and observations $X_{ij1}, \ldots, X_{ijn}, i = 1, \ldots, I, j = 1, \ldots, J$ from the treatment groups are normally distributed with means μ_0 and μ_{ij} respectively, and a common unknown variance σ^2. The treatment means can be expressed as $\mu_{ij} = \mu + \alpha_i + \beta_j$ with $\sum_{i=1}^{I} \alpha_i = 0$ and $\sum_{j=1}^{J} \beta_j = 0$, here μ is the grand mean of the treatment groups, α_i (β_j) is the effect of the ith (jth) level of Factor A (B). Let $\bar{X}_{0\cdot} = \frac{\sum_{l=1}^{n_0} X_{0l}}{n_0}$ be the sample mean of the control group with sample size n_0, and let $\bar{X}_{ij\cdot}$ be the sample mean of the treatment group at

the ith level of Factor A and the jth level of Factor B with sample size n, that is, $\bar{X}_{ij\cdot} = \frac{\sum_{r=1}^{n} X_{ijr}}{n}$. In these experiments we are interested in comparing μ_{ij} with μ_0 and in comparing factor means $\mu_{i\cdot}$, $\mu_{\cdot j}$ and the grand mean μ with μ_0, where $\mu_{i\cdot} = \sum_{j=1}^{J} \mu_{ij}/J$, $\mu_{\cdot j} = \sum_{i=1}^{I} \mu_{ij}/I$, $\mu = \sum_{i=1}^{I}\sum_{j=1}^{J} \mu_{ij}/IJ$.

Let $N = n_0 + nIJ$ be the total sample size, and let $\nu = N - I - J$ be the degrees of freedom for the usual pooled estimator S^2 of the common variance σ^2,

$$S^2 = \frac{1}{\nu}\Big[\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{r=1}^{n}(X_{ijr} - \bar{X}_{ij\cdot})^2 + \sum_{i=1}^{I}\sum_{j=1}^{J} n(\bar{X}_{ij\cdot} - \bar{X}_{i\cdot\cdot} - \bar{X}_{\cdot j\cdot} + \bar{X}_{\cdots})^2 + \sum_{l=1}^{n_0}(X_{0l} - \bar{X}_{0\cdot})^2\Big].$$

where $\bar{X}_{i\cdot\cdot} = \frac{\sum_{j=1}^{J}\sum_{r=1}^{n} X_{ijr}}{nJ}$, $\bar{X}_{\cdot j\cdot} = \frac{\sum_{i=1}^{I}\sum_{r=1}^{n} X_{ijr}}{nI}$, $\bar{X}_{\cdots} = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{r=1}^{n} X_{ijr}}{nIJ}$. The statistic S^2 is independent of the mean vector $\bar{X} = (\bar{X}_{0\cdot}, \bar{X}_{11\cdot}, \cdots, \bar{X}_{IJ\cdot})$, and $\nu S^2/\sigma^2$ has a chi-squared distribution with ν degrees of freedom.

Let the null hypothesis be H_0: $\mu_0 = \mu_{11} = \cdots = \mu_{IJ}$ and let the alternative hypothesis be H_1: $\mu_0 \le \mu_{ij}$ $(i = 1, \ldots, I, j = 1, \ldots, J)$ with at least one strict inequality. The hypothesis H_1 is known in the literature as a simple tree ordering. Dunnett (1955) proposed the test that rejects H_0 for a large value of

$$D_k = \max_{1\le i\le I, 1\le j\le J} \frac{\bar{X}_{ij\cdot} - \bar{X}_{0\cdot}}{S\sqrt{\frac{1}{n_0} + \frac{1}{n}}} \tag{3.1}$$

and gave its critical values for equal sample size, where $k = IJ$. The likelihood ratio test of H_0 under the assumed simple tree order has been considered by Robertson and Wright (1985) and Conaway, Pillars, Robertson and Sconing (1991). Tang and Lin (1997) proposed an approximate likelihood ratio test for the problem. The power of LRT is generally good, but its computation is a challenging task. Partly because of the difficulties involved in applying LRTs for ordered hypotheses, some researchers such as Abelson and Tukey (1963) considered testing homogeneity versus

the simple tree alternative using tests based on single contrasts. One advantage of contrast test is that contrast test statistic is normally distributed with easily computed mean and variance under both the null and alternative hypotheses. While the contrast tests are very simple to use, their power characteristics are such that they cannot be recommended in general as competitors to the LRT. Certain multiple contrast tests have excellent power properties. Mukerjee, Robertson and Wright (1987) studied a class of multiple contrast tests and compared their powers to those of Dunnett's test and Abelson and Tukey's single-contrast test. Marcus and Talpetz (1992) developed a set of simultaneous one-sided confidence bounds (SCB) for all simple tree contrasts. However, the above multiple contrast tests except Dunnett's procedure are not suitable for comparing individual treatment mean μ_{ij} with μ_0 and in comparing factor means $\mu_{i.}$, $\mu_{.j}$ with μ_0 simultaneously.

Dunnett's procedure can be used to construct confidence lower bounds for the differences of treatment means and a control mean. The aim of this chapter is to propose a new test which is powerful and it can be used to construct simultaneous confidence lower bounds for the differences of factor means and a control mean in two-factor experiments. The new multiple contrast test statistic is given in Section 3.2, the power comparison is conducted in Section 3.3, simultaneous confidence lower bounds inverted from the proposed multiple contrast test statistic are presented in Section 3.4, in Section 3.5 a numerical example is employed to illustrate the gains of the new procedure, the extension of the procedure applied to many other designs is given in Section 3.6, and finally a conclusion is presented in Section 3.7.

3.2 A Multiple Contrast Test

For an experiment with two factors and a control, we are interested in comparing μ_{ij}, $\mu_{i.}$, $\mu_{.j}$ and μ with μ_0. The proposed test statistic rejected H_0 in favor of $H_1 - H_0$ if the test statistic

$$G_{k,\mathcal{C}} = \max_{C \in \mathcal{C}} \frac{\bar{X}_C - \bar{X}_0}{S\sqrt{\frac{1}{n_0} + \frac{1}{n_C}}} \tag{3.2}$$

is large, where $\bar{X}_C$ denotes the combined treatment mean with cells in C, n_C is the combined sample size in C and $k = IJ$. When $I \geq 2$ and $J \geq 2$, the collection $\mathcal{C}$ consists of $(I+1)(J+1)$ subsets which include all singletons $\{(i,j)\}$, all ith row $\{(i,1),(i,2),\ldots,(i,J)\}$, all jth column $\{(1,j),(2,j),\ldots,(I,j)\}$ and the set of all treatments $\{(i,j) : i = 1,\ldots,I, j = 1,\ldots,J\}$. When $I = 1$ or $J = 1$, the collection $\mathcal{C}$ consists of $I \times J + 1$ subsets.

Let the $1-\alpha$ percentile of the distribution of $G_{k,\mathcal{C}}$ be denoted by $c^{\alpha}_{k,\mathcal{C},\nu}$ when H_0 is true. One may evaluate these percentiles by a numerical integration of $k+1$ dimensions such as Genz (1992). Due to the complexity of our acceptance region $\{G_{k,\mathcal{C}} \leq c^{\alpha}_{k,\mathcal{C},\nu}\}$, these percentiles were calculated by simulation. The simulated percentiles are provided in Table 3.1 for $\alpha = .10$, .05 and .01, $k = 2, 4, 6, 8, 9, 10$, and the degrees of freedom ν=5, 7, 10, 15, 20, 25, 30, 40, 60, 100, and ∞ for the equal sample size case. While one run of 1,000,000 iterations is sufficient for the levels $\alpha = .10$, and $\alpha = .05$, and also for the level $\alpha = .01$ with $df = \infty$, up to seven runs were used for the level $\alpha = .01$ with $df = 5$. The accuracy employed here is that the simulated tail probabilities at $c^{\alpha}_{k,\mathcal{C},\nu} + 0.01$ and $c^{\alpha}_{k,\mathcal{C},\nu} - 0.01$ lay below and above α respectively by more than three standard deviations.

It is frequently encountered that the sample size for the control is larger than

the sample sizes for the treatments, particularly in medical and biological research. It is not uncommon to have twice the sample size for the control group and Table 3.2 is for the case $n_1 = \cdots = n_k = n$ and $n_0 = 2n$.

It is straight forward from (3.1) and (3.2) that $G_{k,\mathcal{C}} \geq D_k$. The $1-\alpha$ percentile $c^{\alpha}_{k,\mathcal{C},\nu}$ is larger than its counterpart $d^{\alpha}_{k,\nu}$ of the Dunnett's procedure. But the differences are relatively small. For example, for $\alpha = .05$, the differences lie between .02 and .06 for equal sample size case.

Let $\pi(\mu)$ be the power of the new test $G_{k,\mathcal{C}}$ at $\mu = (\mu_0, \mu_{11}, \ldots, \mu_{IJ})$. The new test possesses the following characteristics for equal sample size cases as well as unequal sample size cases.

Theorem 3.2.1 *If $\mu_{1ij} - \mu_{10} \leq \mu_{2ij} - \mu_{20}$, $i = 1, \ldots, I, j = 1, \ldots, J$, then $\pi(\mu_1) \leq \pi(\mu_2)$. Furthermore, $G_{k,\mathcal{C}}$ is unbiased and consistent.*

Proof. Let Y_0 and Y_{ij}, $i = 1, \ldots, I, j = 1, \ldots, J$ be independent normal variates with mean 0 and variance σ^2/n_0 and σ^2/n, respectively. Let $\bar{X}^{\mu_1}_{0\cdot} = Y_0 + \mu_{10}$, $\bar{X}^{\mu_1}_{ij\cdot} = Y_{ij} + \mu_{1ij}$, $\bar{X}^{\mu_2}_{0\cdot} = Y_0 + \mu_{20}$, and $\bar{X}^{\mu_2}_{ij\cdot} = Y_{ij} + \mu_{2ij}$, where $\mu_{1ij} - \mu_{10} \leq \mu_{2ij} - \mu_{20}$, $i = 1, \ldots, I, j = 1, \ldots, J$. We may rewrite the statistic $G_{k,\mathcal{C}}$ as

$$G_{k,\mathcal{C}} = \max\left\{ \max_{i,j} \frac{\bar{X}_{ij\cdot} - \bar{X}_{0\cdot}}{S\sqrt{\frac{1}{n_o} + \frac{1}{n}}}, \max_{i} \frac{\bar{X}_{i\cdot\cdot} - \bar{X}_{0\cdot}}{S\sqrt{\frac{1}{n_o} + \frac{1}{nJ}}}, \max_{j} \frac{\bar{X}_{\cdot j\cdot} - \bar{X}_{0\cdot}}{S\sqrt{\frac{1}{n_o} + \frac{1}{nI}}}, \frac{\bar{X}_{\cdots} - \bar{X}_{0\cdot}}{S\sqrt{\frac{1}{n_o} + \frac{1}{nIJ}}} \right\}.$$

Consider the first case, we have that

$$\max_{i,j} \frac{\bar{X}^{\mu_1}_{ij\cdot} - \bar{X}^{\mu_1}_{0\cdot}}{S\sqrt{\frac{1}{n_o} + \frac{1}{n}}} = \max_{i,j} \frac{(Y_{ij} - Y_0) + (\mu_{1ij} - \mu_{10})}{S\sqrt{\frac{1}{n_o} + \frac{1}{n}}}.$$

Since $\mu_{1ij} - \mu_{10} \leq \mu_{2ij} - \mu_{20}$, $i = 1, \ldots, I, j = 1, \ldots, J$,

$$\max_{i,j} \frac{\bar{X}^{\mu_1}_{ij\cdot} - \bar{X}^{\mu_1}_{0\cdot}}{S\sqrt{\frac{1}{n_o} + \frac{1}{n}}} \leq \max_{i,j} \frac{\bar{X}^{\mu_2}_{ij\cdot} - \bar{X}^{\mu_2}_{0\cdot}}{S\sqrt{\frac{1}{n_o} + \frac{1}{n}}}.$$

Similarly, we may obtain

$$\max_i \frac{\bar{X}^{\mu_1}_{i\cdot\cdot} - \bar{X}^{\mu_1}_{0\cdot}}{S\sqrt{\frac{1}{n_o}+\frac{1}{nJ}}} \le \max_i \frac{\bar{X}^{\mu_2}_{i\cdot\cdot} - \bar{X}^{\mu_2}_{0\cdot}}{S\sqrt{\frac{1}{n_o}+\frac{1}{nJ}}},$$

$$\max_j \frac{\bar{X}^{\mu_1}_{\cdot j\cdot} - \bar{X}^{\mu_1}_{0\cdot}}{S\sqrt{\frac{1}{n_o}+\frac{1}{nI}}} \le \max_j \frac{\bar{X}^{\mu_2}_{\cdot j\cdot} - \bar{X}^{\mu_2}_{0\cdot}}{S\sqrt{\frac{1}{n_o}+\frac{1}{nI}}},$$

and

$$\frac{\bar{X}^{\mu_1}_{\cdots} - \bar{X}^{\mu_1}_{0\cdot}}{S\sqrt{\frac{1}{n_o}+\frac{1}{nIJ}}} \le \frac{\bar{X}^{\mu_2}_{\cdots} - \bar{X}^{\mu_2}_{0\cdot}}{S\sqrt{\frac{1}{n_o}+\frac{1}{nIJ}}}.$$

Hence

$$G^{\mu_1}_{k,\mathcal{C}} \le G^{\mu_2}_{k,\mathcal{C}}.$$

Therefore, $\pi(\mu_1) = P(G^{\mu_1}_{k,\mathcal{C}} \ge c^{\alpha}_{k,\mathcal{C},\nu}) \le P(G^{\mu_2}_{k,\mathcal{C}} \ge c^{\alpha}_{k,\mathcal{C},\nu}) = \pi(\mu_2)$. The result holds for any collection $\mathcal{C}$.

If $\mu = (\mu_0, \mu_{11}, \ldots, \mu_{IJ})$ is such that $\mu_{ij} \ge \mu_0$, $i = 1, \ldots, I, j = 1, \ldots, J$, then so is $\mu - \mathbf{a} = (\mu_0 - a, \mu_{11} - a, \ldots, \mu_{IJ} - a)$ for any constant vector $\mathbf{a} = (a, \ldots, a)$. By the monotonicity of the new test, $\pi(\mu) = P(G^{\mu}_{k,\mathcal{C}} \ge c^{\alpha}_{k,\mathcal{C},\nu}) \ge \pi(a) = \alpha$. Hence, $G_{k,\mathcal{C}}$ is unbiased.

Since

$$\begin{aligned}
\max_{i,j} \frac{\mu_{ij} - \mu_0}{S\sqrt{\frac{1}{n_o}+\frac{1}{n}}} &\le \max_C \frac{\mu_C - \mu_0}{S\sqrt{\frac{1}{n_o}+\frac{1}{n_C}}} \\
&= \max_C \frac{(\bar{X}^{\mu}_C - \bar{X}^{\mu}_{0\cdot}) - (Y_C - Y_0)}{S\sqrt{\frac{1}{n_o}+\frac{1}{n_C}}} \\
&\le \max_C \frac{\bar{X}^{\mu}_C - \bar{X}^{\mu}_{0\cdot}}{S\sqrt{\frac{1}{n_o}+\frac{1}{n_C}}} + \max_C \frac{-(Y_C - Y_0)}{S\sqrt{\frac{1}{n_o}+\frac{1}{n_C}}}.
\end{aligned}$$

for any given ϵ we can find a positive a such that

$$P\Big\{ \max_C \frac{-(Y_C - Y_0)}{S\sqrt{\frac{1}{n_o}+\frac{1}{n_C}}} \le a \Big\} \ge 1 - \epsilon.$$

For a positive constant b, if

$$\max_{i,j} \frac{\mu_{ij} - \mu_0}{S\sqrt{\frac{1}{n_0} + \frac{1}{n}}} \geq (a + b),$$

then

$$P\Big\{ \max_{C} \frac{\bar{X}_C^{\mu} - \bar{X}_0^{\mu}}{S\sqrt{\frac{1}{n_0} + \frac{1}{n_C}}} \geq b \Big\} \geq 1 - \epsilon.$$

Therefore, the power $\pi(\mu)$ of $G_{k,\mathcal{C}}$ converges uniformly to one as $\max_{i,j} \frac{\mu_{ij} - \mu_0}{\sqrt{\frac{1}{n_0} - \frac{1}{n}}} \to \infty$ if $\mu \in H_1 - H_0$. □

3.3 Power Comparisons

The powers of the new test, Dunnett's test and the LRT are investigated in this section. The Monte Carlo method is used with 10,000 iterations. The standard errors are at most 0.005. For simplicity, we consider the equal sample size case with $\sigma^2/n = 1$, n is the common sample size. Three configurations are considered: the center direction $(-k, 1, \ldots, 1)$, the edge direction $(-1, k, -1, \ldots, -1)$, and the direction of pairwise comparison $(-1, 1, 0, \ldots, 0)$ which lies in the middle of the center direction and the edge direction. The center direction and the edge direction tend to attain the maximum and the minimum power respectively for all the tests. The simulated powers of the new test, Dunnett's test and LRT are provided in Table 3.3, where $\Delta^2 = n[(\mu_0 - \bar{\mu})^2 + \sum_{i=1}^{I} \sum_{j=1}^{J} (\mu_{ij} - \bar{\mu})^2]/\sigma^2$ is the noncentrality parameter and $\bar{\mu} = (\mu_0 + \sum_{i=1}^{I} \sum_{j=1}^{J} \mu_{ij})/(IJ+1)$. The new test is shown to be the most powerful one-sided test along the center direction when treatment means are approximately equal and are larger than the control mean. At $k = 6$, the percentages of the power

of the new test are 7.71, 17.84, and 15.09 higher than those of the LRT, for $\Delta = 1, 2,$ and 3, respectively. Here experimenters are often interested in comparing μ with μ_0. The scab index data in Section 3.5 indicates such a tendency. It is also the most powerful one along the pairwise directions $(-1, 1, 0, \ldots, 0), \ldots$ or $(-1, 0, \ldots, 0, 1)$ when treatment means are larger than the control mean and one treatment mean is larger than the remaining treatment means. At $k = 6$, the percentages of the power of the new test are 2.93, 8.23, and 10.26 higher than those of the LRT, for $\Delta = 1, 2,$ and 3, respectively. However, LRT is the clear choice along the edge direction when one treatment is effective while the remainings are not.

Consider the configuration when treatment means are larger than the control mean and all treatment means of the ith level of Factor A are larger than the remaining treatment means, the direction is an average of J pairwise directions. For example, the average of (-1; 1,0,0:0,0,0), (-1;0,1,0;0,0,0), and (-1:0,0,1:0,0,0) is (-1: 1/3,1/3,1/3;0,0,0). Table 3.3 indicates that the new test is more powerful than the LRT and Dunnett's procedure along the center direction and pairwise directions. Therefore, the new test is the most powerful one in comparing $\mu_{1.}$ with μ_0. The scab index data in Table 3.6 exihibits that the mean of the three fall applications, $\mu_{1.}$, is significantly below the control mean μ_0. Consequently, the new test will be the most powerful one to detect such a difference, as shown in the following section.

3.4 Simultaneous Confidence Lower Bounds

It is usual to say that a treatment is better than the control when that treatment mean is larger than the control mean. Dunnett's one-sided simultaneous confidence lower bounds for the difference between each treatment mean μ_{ij} and the control

mean μ_0 is

$$\mu_{ij} - \mu_0 \geq \bar{X}_{ij\cdot} - \bar{X}_{0\cdot} - d^{\alpha}_{k,\nu} S \sqrt{\frac{1}{n_0} + \frac{1}{n}}. \tag{3.3}$$

The weighted averages of (3.3) may be used as in Miller (1981) and the corresponding simultaneous confidence lower bounds are

$$\mu_{i\cdot} - \mu_0 \geq \bar{X}_{i\cdot\cdot} - \bar{X}_{0\cdot} - d^{\alpha}_{k,\nu} S \sqrt{\frac{1}{n_0} + \frac{1}{n}}. \tag{3.4}$$

Similarly for the cases of $\mu_{\cdot j} - \mu_0$ and $\mu - \mu_0$. Our new test statistic $G_{k,\mathcal{C}}$ in (3.2) has the following $100(1-\alpha)\%$ simultaneous confidence lower bounds

$$\mu_{ij} - \mu_0 \geq \bar{X}_{ij\cdot} - \bar{X}_{0\cdot} - c^{\alpha}_{k,\mathcal{C},\nu} S \sqrt{\frac{1}{n_0} + \frac{1}{n}}. \tag{3.5}$$

$$\mu_{i\cdot} - \mu_0 \geq \bar{X}_{i\cdot\cdot} - \bar{X}_{0\cdot} - c^{\alpha}_{k,\mathcal{C},\nu} S \sqrt{\frac{1}{n_0} + \frac{1}{nJ}}, \tag{3.6}$$

and similarly for the cases of $\mu_{\cdot j} - \mu_0$ and $\mu - \mu_0$. The one-sided simultaneous lower bounds (3.3), (3.4), (3.5), and (3.6) are used without assuming that $\mu_{ij} \geq \mu_0$. In the numerical example in Section 3.5, combinations of treatments include a fall (or spring) application of 300, 600, or 1200lb of sulphur per acre $\mu_{1\cdot}$ ($\mu_{2\cdot}$): a fall or a spring application of 300lb (600lb or 1200lb) per acre of sulphur $\mu_{\cdot 1}$ ($\mu_{\cdot 2}$ or $\mu_{\cdot 3}$); among others. Comparing (3.6) with (3.4), we observe that $c^{\alpha}_{k,\mathcal{C},\nu}\sqrt{\frac{1}{n_0} + \frac{1}{Jn}}$ is generally smaller than $d^{\alpha}_{k,\nu}\sqrt{\frac{1}{n_0} + \frac{1}{n}}$ when the number of combined treatments J is at least 2. Our new test statistic can provide sharper simultaneous confidence lower bounds of combinations of certain treatments means and the control mean than those of Dunnett's method.

The scab index experiment in Cochran and Cox (1957) (see Section 3.5) is such an example where a treatment is said to be better than the control when that treatment

mean is smaller than the control mean. The corresponding one-sided simultaneous confidence lower bounds to (3.3), (3.4), (3.5), and (3.6) are

$$\mu_0 - \mu_{ij} \geq \bar{X}_{0\cdot} - \bar{X}_{ij\cdot} - d^{\alpha}_{k,\nu} S \sqrt{\frac{1}{n_0} + \frac{1}{n}}, \tag{3.7}$$

$$\mu_0 - \mu_{i\cdot} \geq \bar{X}_{0\cdot} - \bar{X}_{i\cdot\cdot} - d^{\alpha}_{k,\nu} S \sqrt{\frac{1}{n_0} + \frac{1}{n}}, \tag{3.8}$$

$$\mu_0 - \mu_{ij} \geq \bar{X}_{0\cdot} - \bar{X}_{ij\cdot} - c^{\alpha}_{k,\mathcal{C},\nu} S \sqrt{\frac{1}{n_0} + \frac{1}{n}}, \tag{3.9}$$

and

$$\mu_0 - \mu_{i\cdot} \geq \bar{X}_{0\cdot} - \bar{X}_{i\cdot\cdot} - c^{\alpha}_{k,\mathcal{C},\nu} S \sqrt{\frac{1}{n_0} + \frac{1}{nJ}}, \tag{3.10}$$

respectively.

3.4.1 Efficiency of Confidence Lower Bounds

The efficiency of the new procedure is compared to Dunnett's procedure in terms of mean heights of confidence lower bounds. The mean heights of confidence lower bounds in (3.5) and (3.6) are $c^{\alpha}_{k,\mathcal{C},\nu}\sqrt{\frac{1}{n_0} + \frac{1}{n_C}}E(S)$, and the mean heights of confidence lower bounds in (3.3) and (3.4) are $d^{\alpha}_{k,\nu}\sqrt{\frac{1}{n_0} + \frac{1}{n}}E(S)$. The ratio of the mean height of Dunnett's confidence lower bounds to that of the new procedure is

$$R^{\alpha}_{k,\nu} = \frac{d^{\alpha}_{k,\nu}\sqrt{\frac{1}{n_0} + \frac{1}{n}}}{c^{\alpha}_{k,\mathcal{C},\nu}\sqrt{\frac{1}{n_0} + \frac{1}{n_C}}} \tag{3.11}$$

The values of $R^{\alpha}_{k,\nu}$ are provided in Table 3.4 for $\alpha = 0.05$, $\nu = \infty$, $n_0 = n_1 = \ldots = n_k$, and $k = 2, 4, 6, 8, 9, 10$ and for the number of combinations in C from 1, the singleton, to 10. In the case of a single treatment versus a control, the mean heights of Dunnett's confidence lower bounds are shorter than those of the new

procedure. But the ratios are very close to 1 and the losses in efficiency are no more than 2.3%.

The mean heights of the new procedure are shorter than those of Dunnett's for treatment combinations of two or more treatments. The ratios of the latter to the former are larger than 1 and the gains in efficiency are at least 12.9%. Further comparisons can be found in the numerical example in Section 5. When the variance σ^2 is known, the ratio of the mean height of Marcus and Talpaz's (1992) simultaneous confidence lower bound to that of the new procedure lies between 1.022 and 1.180 when $k = 2$; between 1.041 and 1.317 when $k = 4$; between 1.064 and 1.393 when $k = 6$. There is no advantage in using Marcus and Talpaz's procedure to compare μ_{ij}, $\mu_{i.}$, $\mu_{.j}$, or μ to the control mean μ_0.

3.4.2 Probabilities of Nonnegative Confidence Lower Bounds

In this subsection, we compare the probabilities of nonnegative confidence lower bounds by Dunnett's procedure with the corresponding ones by the new procedure. The experiment is a 2×3 layout with a control. The row effects are denoted by α_1, α_2 and the column effects are denoted by $\beta_1, \beta_2, \beta_3$ excluding a control. The control mean is assumed to be zero, i.e., $\mu_0 = 0$, and $n_0 = 2n$. Six cases are considered and they are

Case 1: $\alpha_1 = \alpha_2 = 0$; $\beta_1 = \beta_2 = \beta_3 = 2.0$;

Case 2: $\alpha_1 = 0$, $\alpha_2 = 1.0$; $\beta_1 = \beta_2 = \beta_3 = 2.0$;

Case 3: $\alpha_1 = 0$, $\alpha_2 = 2.0$; $\beta_1 = \beta_2 = \beta_3 = 2.0$;

Case 4: $\alpha_1 = \alpha_2 = 0$; $\beta_1 = 1.0$, $\beta_2 = 2.0$, $\beta_3 = 3.0$;

Case 5: $\alpha_1 = 0$, $\alpha_2 = 1.0$; $\beta_1 = 1.0$, $\beta_2 = 2.0$, $\beta_3 = 3.0$;

Case 6: $\alpha_1 = 0, \alpha_2 = 2.0$: $\beta_1 = 1.0, \beta_2 = 2.0, \beta_3 = 3.0$.

For each case, the probabilities of nonnegative confidence lower bounds for Dunnett's procedure and the new procedure are calculated. For simplicity, we assume $\sigma^2/n = 1$. The results are provided in Table 3.5.

These probabilities are the percentages of detecting the differences between the row mean $\mu_{i.}$ and the control mean μ_0, between the column mean $\mu_{.j}$ and the control mean μ_0, and between the grand mean μ and the control mean μ_0. The probabilities of nonnegative confidence lower bounds by the new procedure are always higher than those by Dunnett's. The differences are substantial and they could be as large as 0.416.

3.5 A Numerical Example

The results of an experiment on the effects of applications of sulphur in reducing scab disease of potatoes can be found in Cochran and Cox (1957). The objective of applying sulphur is to increase the acidity of the soil. In addition to untreated plots which serve as a control, 3 amounts of dressing were compared at 300, 600, and 1200lb. per acre. Both a fall and a spring application of each amount were tested, so that there were six distinct treatment groups excluding the control. There were four observations for each treatment group and eight observations for the control group. The thirty-two observations are provided in Table 3.6, where the six treatment groups are labelled as F3, S3, F6, S6, F12 and S12 respectively. The pooled variance estimate is $S^2 = 41.93$ with 27 degrees of freedom. The interaction between the time of application and the amount of dressing were not significant, the corresponding p-value was 0.670, and a two-way additive model is used.

The overall F-statistic for testing $H_0 : \mu_0 = \mu_{11} = \mu_{12} = \mu_{13} = \mu_{21} = \mu_{22} = \mu_{23}$ against all alternatives has a value of 3.83 and its p-value is 0.007.

The likelihood ratio statistic $\bar{E}^2_{01}$ for testing H_0 against the alternative hypothesis $H_1 : \mu_0 \geq \mu_{ij}$ $(i = 1, 2, j = 1, 2, 3)$ with at least one inequality has a value of 0.461 and its p-value is 0.00303. Here the degrees of freedom for the LRT $\bar{E}^2_{01}$ is 25.

In this application, low treatment response is preferred. The statistics (3.1) and (3.2) are given respectively by

$$D_6 = \max_{1 \leq i \leq 2, 1 \leq j \leq 3} \frac{\bar{X}_{0\cdot} - \bar{X}_{ij}}{S\sqrt{\frac{1}{n_0} + \frac{1}{n}}},$$

and

$$G_{6,\mathcal{C}} = \max_{C \in \mathcal{C}} \frac{\bar{X}_{0\cdot} - \bar{X}_C}{S\sqrt{\frac{1}{n_0} + \frac{1}{n_C}}}. \tag{3.12}$$

By using the interpolation method in Hochberg and Tamhane (1987), the critical value for Dunnett's one-sided test for $n_0 = 8$ and $n_{11} = n_{12} = n_{13} = n_{21} = n_{22} = n_{23} = 4$ is $d^{0.05}_{6,27} = 2.48$. The Dunnett's statistic is equal to 4.24 and its exact p-value is 0.000657. The new statistic $G_{k,\mathcal{C}}$ of (3.12) is used to test H_0 against H_1. Its critical value $c^{0.05}_{6,\mathcal{C},27}$ depends on the number of comparisons used in (3.12). That number is normally $(I + 1)(J + 1)$. However, for this particular application we are also interested in comparing the mean response of dressing 600lb. or more per acre with the control mean. Therefore, there are thirteen comparisons in this application and the new statistic has the value of 4.24. The critical value is $c^{0.05}_{6,\mathcal{C},25} = 2.54$ through simulation and its exact p-value is 0.000935.

For these thirteen comparisons, their simultaneous confidence lower bounds by Dunnett's procedure and the new procedure are provided in Table 3.7. In these

interval estimations, we do not assume that the application of sulphur is effective in reducing scab indices. The notations used here are: F stands for Factor A at level one which includes F3, F6, and F12; S stands for Factor A at level two which includes S3, S6, S12; 3 stands for Factor B at level one which includes F3 and S3; 6 stands for Factor B at level two which includes F6 and S6; 12 stands for Factor B at level three which includes F12 and S12; "$\geq$ 6" stands for Factor B at level two and level three which includes F6, S6, F12 and S12; and "$\geq$ 3" stands for the combination of all six treatments. Table 3.7 is divided into two cases. The first case is for the confidence lower bounds of a single treatment against the control $\mu_{ij} - \mu_0$. Both Dunnett's procedure and the new procedure detect the difference between F3 and the control, as well as the difference between F12 and the control. Dunnett's confidence lower bounds are larger than our confidence lower bounds with differences of 0.236 and 0.237. The second case is for the confidence lower bounds of combined treatment against the control $\mu_C - \mu_0$. Both Dunnett's procedure and the new procedure detect the difference between 12 (F12 and S12) and the control as well as the difference between F(F3, F6 and F12) and the control. However, the new procedure also detects the difference between 3 (F3 and S3) and the control, and the difference between "$\geq$ 6" (F6, S6, F12, S12) and the control. It is also found that the application of sulphur ("$\geq$ 3") resulted in a significant reduction in scab index. Dunnett's procedure fails to detect these three differences. Our confidence lower bounds are larger than Dunnett's confidence lower bounds with differences ranging from 1.61 to 3.12.

The ratio of the mean height of Dunnett's confidence lower bound to that of the new method is $R^{0.05}_{6,27,2} = 1.196$ for combined treatments 3, 6, or 12; $R^{0.05}_{6,27,3} = 1.310$

for combined treatments F or S; $R^{0.05}_{6,27,4} = 1.381$ for treatment combinations ("≥ 6"), and $R^{0.05}_{6,27,6} = 1.465$ for all treatment combinations ("≥ 3"). The results demonstrate that the new procedure is more efficient than Dunnett's for comparing combined treatment means with the control mean. The larger the number in the combined treatment group, the higher the relative efficiency.

3.6 Extensions

The use of the statistic $G_{k,C}$ in (3.2) and its corresponding simultaneous confidence lower bounds can be applied to many different designs involving a control. One extension is to a two-way design. Factor A is a time factor and Factor B has J levels and a control of no treatment such as the above example. Each ith level of Factor A at the control level of Factor B represents an independent and identical repetition of a controlled experimental trial. The data gathered represents an $I \times J$ experiment with n observations in each treatment group and a control with nI observations. For 2×2, 2×3, 2×4 and 2×5 experiments with a control of $2n$ observations, one may use the critical values provided in Table 3.2.

Another extension is also to a two-way design. Factor A has I levels and a control and Factor B has J levels and a control. There are $(I+1)(J+1)-1$ treatments. Experimenters are interested in comparing μ_{ij} with the control mean μ_{00} and they may also be interested in comparing $\mu_{i\cdot} = \sum_{j=0}^{J} \frac{\mu_{ij}}{J+1}$, $\mu_{0\cdot} = \sum_{j=1}^{J} \frac{\mu_{0j}}{J}$, $\mu_{\cdot j} = \sum_{i=0}^{I} \frac{\mu_{ij}}{I+1}$, $\mu_{\cdot 0} = \sum_{i=1}^{I} \frac{\mu_{i0}}{I}$, or $\mu_{\cdot\cdot} = (\sum_{i=0}^{I}\sum_{j=0}^{J} - \mu_{00})/(IJ+I+J)$ with the control mean μ_{00}.

A third extension is made for the design of k treatments and a control. Two or more treatment groups may have the same characteristic. They may be the same treatment with different dosage levels or different times of application. Ex-

perimenters may be interested in comparing a set of treatment groups to the control. The critical values $c^{\alpha}_{k,\nu}$ of an all-purpose test statistic G_k corresponding to (3.2) involving all $2^k - 1$ possible comparisons can be found in Peng, Lee, and Liu (1999). This approach is somewhat conservative if one is interested in 13 comparisons instead of 63 comparisons when $k = 6$. The former requires a critical value of $c^{0.05}_{6,25,C} = 2.56$ while the latter requires a critical value of $c^{0.05}_{6,25} = 2.66$ when $n_0 = 8$ and $n_1 = n_2 = n_3 = n_4 = n_5 = n_6 = 4$.

3.7 Conclusion

The new procedure is appealing in that its mean height of simultaneous confidence lower bounds is only slightly larger than that of Dunnett's for the difference of a single treatment mean and the control mean, but it is substantially smaller than that of Dunnett's for the difference of a treatment factor mean $\mu_{i\cdot}$ (or $\mu_{\cdot j}$) and the control mean. As a consequence, the new procedure is significantly more powerful than Dunnett's in detecting the difference between factor means $\mu_{i\cdot}$ or $\mu_{\cdot j}$ and the control mean μ_0. When those comparisons are of interest, the new procedure is recommended. On the other hand, our new procedure as a test statistic is more powerful than the LRT and Dunnett's test along the center direction and pairwise directions.

Table 3.1: Upper Percentage Points for $I \times J$ Experiment With Equal Sample Size.

		k=treatment groups					
df	α	2×1	2×2	2×3	2×4	3×3	2×5
5	.10	1.90	2.29	2.48	2.62	2.67	2.72
	.05	2.48	2.91	3.12	3.28	3.33	3.39
	.01	3.96	4.53	4.80	5.00	5.07	5.16
7	.10	1.80	2.16	2.33	2.45	2.49	2.55
	.05	2.29	2.68	2.86	2.98	3.03	3.08
	.01	3.47	3.92	4.13	4.28	4.34	4.40
10	.10	1.74	2.07	2.23	2.34	2.37	2.42
	.05	2.18	2.52	2.68	2.80	2.83	2.88
	.01	3.16	3.54	3.71	3.83	3.87	3.92
15	.10	1.69	2.00	2.15	2.25	2.29	2.33
	.05	2.10	2.41	2.56	2.66	2.70	2.73
	.01	2.96	3.27	3.41	3.52	3.56	3.61
20	.10	1.67	1.97	2.11	2.22	2.24	2.29
	.05	2.06	2.35	2.50	2.60	2.63	2.67
	.01	2.86	3.16	3.29	3.39	3.42	3.46
25	.10	1.65	1.95	2.09	2.19	2.22	2.26
	.05	2.03	2.33	2.46	2.55	2.59	2.63
	.01	2.80	3.09	3.22	3.31	3.34	3.38
30	.10	1.64	1.94	2.08	2.17	2.20	2.25
	.05	2.02	2.31	2.44	2.53	2.56	2.60
	.01	2.76	3.04	3.17	3.26	3.28	3.32
40	.10	1.63	1.92	2.06	2.15	2.19	2.23
	.05	2.00	2.28	2.41	2.50	2.54	2.57
	.01	2.72	2.99	3.11	3.19	3.23	3.26
60	.10	1.62	1.91	2.04	2.14	2.17	2.20
	.05	1.98	2.26	2.39	2.47	2.50	2.54
	.01	2.68	2.94	3.05	3.13	3.17	3.19
100	.10	1.61	1.90	2.03	2.12	2.15	2.19
	.05	1.97	2.24	2.36	2.45	2.48	2.52
	.01	2.65	2.90	3.01	3.08	3.11	3.15
∞	.10	1.60	1.88	2.01	2.10	2.13	2.16
	.05	1.94	2.21	2.33	2.42	2.44	2.48
	.01	2.59	2.84	2.95	3.02	3.04	3.08

Table 3.2: Upper Percentage Points for $I \times J$ Experiment With $n_0 = 2n_1 = \ldots = 2n_k$.

		k=treatment groups					
df	α	2×1	2×2	2×3	2×4	3×3	2×5
5	.10	1.95	2.41	2.63	2.79	2.85	2.91
	.05	2.53	3.05	3.29	3.46	3.52	3.59
	.01	4.02	4.70	5.02	5.24	5.31	5.42
7	.10	1.85	2.27	2.46	2.60	2.64	2.70
	.05	2.35	2.79	2.99	3.13	3.18	3.25
	.01	3.53	4.04	4.27	4.45	4.51	4.58
10	.10	1.78	2.17	2.35	2.47	2.51	2.56
	.05	2.23	2.62	2.80	2.92	2.96	3.02
	.01	3.21	3.63	3.83	3.96	4.00	4.06
15	.10	1.73	2.09	2.26	2.37	2.41	2.46
	.05	2.14	2.49	2.66	2.77	2.81	2.85
	.01	3.00	3.35	3.51	3.62	3.66	3.71
20	.10	1.71	2.06	2.22	2.33	2.36	2.41
	.05	2.09	2.44	2.59	2.70	2.73	2.78
	.01	2.89	3.23	3.38	3.48	3.51	3.56
25	.10	1.69	2.03	2.19	2.30	2.33	2.38
	.05	2.07	2.40	2.55	2.65	2.69	2.74
	.01	2.83	3.16	3.30	3.39	3.42	3.47
30	.10	1.68	2.02	2.18	2.28	2.31	2.36
	.05	2.05	2.38	2.53	2.63	2.66	2.70
	.01	2.79	3.11	3.24	3.34	3.36	3.41
40	.10	1.67	2.01	2.16	2.26	2.29	2.34
	.05	2.03	2.36	2.50	2.60	2.63	2.67
	.01	2.75	3.05	3.18	3.27	3.30	3.34
60	.10	1.66	1.99	2.14	2.24	2.27	2.31
	.05	2.01	2.33	2.47	2.56	2.59	2.63
	.01	2.70	3.00	3.12	3.20	3.23	3.26
100	.10	1.65	1.98	2.12	2.22	2.25	2.29
	.05	2.00	2.31	2.44	2.54	2.57	2.60
	.01	2.67	2.96	3.07	3.15	3.18	3.21
∞	.10	1.63	1.96	2.10	2.20	2.23	2.27
	.05	1.98	2.28	2.41	2.50	2.52	2.57
	.01	2.62	2.89	3.01	3.08	3.10	3.14

Table 3.3: Power Comparisons.

			Test		
Direction	k	Δ	$G_{k,\mathcal{C}}$	Dunnett's	LRT
Center					
	2	1	24.44	23.60	20.57
		2	60.07	58.46	53.68
		3	88.86	87.76	85.29
	4	1	24.41	22.87	17.13
		2	59.38	56.48	45.75
		3	88.25	86.59	77.93
	6	1	23.34	22.58	15.63
		2	57.63	55.92	39.79
		3	86.97	85.38	71.88
Pairwise					
	2	1	22.21	21.78	20.19
		2	56.92	56.12	53.54
		3	87.22	86.74	85.32
	4	1	20.33	19.20	16.21
		2	52.29	50.93	44.40
		3	84.17	83.89	77.65
	6	1	17.07	18.95	14.14
		2	46.91	48.30	38.68
		3	80.93	80.11	70.67
Edge					
	2	1	15.99	15.87	17.37
		2	42.33	42.81	49.23
		3	74.39	74.93	83.10
	4	1	10.82	10.62	12.26
		2	28.36	29.16	36.50
		3	57.75	59.30	71.76
	6	1	9.04	9.12	9.88
		2	22.82	23.46	30.10
		3	48.82	50.27	62.69

Table 3.4: Ratio of the Mean Height of Dunnett's Confidence Lower Bound to That of the New Procedure. $\alpha = 0.05$, $n_0 = n_1 = \ldots = n_k$ and $\nu = \infty$.

Combination	k=2	k=4	k=6	k=8	k=9	k=10
1	0.990	0.977	0.983	0.988	0.992	0.988
2	1.143	1.129	1.135	1.140	1.145	1.141
3		1.197	1.204	1.209	1.215	1.210
4		1.236	1.243	1.249	1.255	1.250
5			1.269	1.275	1.280	1.275
6			1.287	1.293	1.299	1.293
7				1.306	1.312	1.307
8				1.317	1.322	1.317
9					1.331	1.325
10						1.332

Table 3.5: Probabilities of Nonnegative Confidence Lower Bounds for Dunnett's Procedure and the New Procedure. $n_0 = 2n_1 = \ldots = 2n_k$.

Case	Method	$\mu_{1.} - \mu_0$	$\mu_{2.} - \mu_0$	$\mu_{.1} - \mu_0$	$\mu_{.2} - \mu_0$	$\mu_{.3} - \mu_0$	$\mu_{\cdot} - \mu_0$
Case 1	D*	0.171	0.171	0.193	0.193	0.193	0.144
	N	0.413	0.413	0.341	0.341	0.341	0.516
Case 2	D	0.171	0.558	0.357	0.357	0.357	0.327
	N	0.413	0.810	0.536	0.536	0.536	0.743
Case 3	D	0.171	0.893	0.553	0.553	0.553	0.565
	N	0.413	0.976	0.722	0.722	0.722	0.897
Case 4	D	0.171	0.171	0.031	0.193	0.553	0.144
	N	0.413	0.413	0.079	0.341	0.722	0.516
Case 5	D	0.171	0.558	0.086	0.357	0.737	0.327
	N	0.413	0.810	0.181	0.536	0.862	0.743
Case 6	D	0.171	0.893	0.193	0.553	0.872	0.565
	N	0.413	0.976	0.341	0.722	0.944	0.897

* D for Dunnett's procedure and N for the new procedure

Table 3.6: The Scab Index Data.

	0		F3	S3	F6	S6	F12	S12
	12	30	9	30	16	18	10	17
	10	18	9	7	10	24	4	7
	24	32	16	21	18	12	4	16
	29	26	4	9	18	19	5	17
Totals	181		38	67	62	73	23	57
Means	22.6		9.5	16.8	15.5	18.2	5.8	14.2

Notation: F=fall, S=spring, 0=control.

Table 3.7: The 95% Simultaneous Confidence Lower Bounds for the Scab Index Data.

	Treatment	Confidence Lower Bound	
		Dunnett's Method	New Method
Single Treatment	F3	3.27	3.03
	S3	-4.03	-4.27
	F6	-2.73	-2.97
	S6	-5.43	-5.67
	F12	6.97	6.73
	S12	-1.43	-1.67
Combined Treatment	3	-0.38	1.23
	6	-4.08	-2.47
	12	2.77	4.38
	F	2.50	4.83
	S	-3.63	-1.31
	≥ 6	-0.66	2.05
	≥ 3	-0.57	2.55

Chapter 4

Statistical Inference for Best Treatment versus a Control

4.1 Introduction

In many experiments, the primary goal is to compare several treatment means with a control mean, or a standard. This is often the case in pharmaceutical studies, where the superiority of any proposed new treatment over a standard treatment must be demonstrated before it is accepted. Alternatively, there may be no standard treatment and the main problem may be to establish whether the new treatment has any beneficial effect, in which case a placebo control treatment may be included in the trial as a standard for comparison.

A number of statistical procedures have been proposed to test whether any of k treatments are different from a control, most of which are multiple comparisons procedures. The most important work is by Dunnett (1955). Many generalizations of Dunnett's procedure have been made. For instance, Hoover (1991) extended it to the case where there are several treatments with two or more controls and joint confidence intervals are required between each treatment and each control

simultaneously. Cheung and Holland (1991, 1992) extended it to the case of more than one group of treatments, each group containing several treatments compared with a specified treatment, with the error rate covering all groups and treatment comparisons simultaneously. Peng, Lee, and Liu (2000) generalized it to compare treatment means with a control mean in two-factor experiments. On the other hand, Steel (1959), Fligner and Wolfe (1982), Spurrier (1988), among others, studied distribution-free analogues of Dunnett's procedure. The reader is referred to the book by Miller (1981) and the more recent books by Hochberg and Tamhane (1987), and Hsu (1996) for detailed accounts of some of these developments and extensive references.

The experiment considered in this chapter is a one-way analysis with $k+1$ levels. Let $Y_{ij}, i = 0, 1, \ldots, k, j = 1, \ldots, n_i$, be independent normal variates with unknown means μ_i $(i = 0, 1, \ldots, k)$ and a common but unknown variance σ^2, where μ_0 denotes the control mean and $\mu_1, \ldots, \mu_k$ denote the treatment means. The statistic $S^2 = \sum_{i=0}^{k} \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2/\nu$ is used as an estimator for σ^2, and it is independent of $\bar{\mathbf{Y}} = (\bar{Y}_0, \ldots, \bar{Y}_k)$, where $\nu S^2/\sigma^2 \sim \chi^2_\nu$ and $\nu = \sum_{i=0}^{k} n_i - k - 1 > 0$.

The first problem we study is to make interval inference with the prior knowledge that treatments are at least as effective as the control. This type of prior knowledge may come from past experiences. The parameter space is $\Omega = \{ \boldsymbol{\mu} \in R^{k+1} : \mu_0 \leq \mu_i, \ i = 1, \ldots, k\}$. The null hypothesis is $H_0 : \mu_0 = \mu_1 = \cdots = \mu_k$ and the alternative hypothesis is $H_1 = \Omega - H_0$. The test of H_0 against H_1 has been well developed. The likelihood ratio test statistic S_{01} for testing H_0 against H_1 is known to possess generally superior operating characteristics to those of its competitors, see Robertson, Wright and Dykstra (1988). A variety of other procedures have

also been proposed, most of which are based on one or more contrasts among the sample means. These include the multiple contrast tests of Mukerjee, Robertson, and Wright (1987) which includes Dunnett's procedure and the single-contrast test as special cases, and ad hoc tests proposed by Williams (1971, 1972). Conaway, Pillers, Robertson, and Sconing (1991) used a circular cone to approximate S_{01} and they developed a test which has the advantages of being easier to compute and can be used with unequal sample sizes. Tang and Lin (1997) used an orthant to approximate S_{01}. McDermott (1999) proposed a class of tests based on an improved orthant approximation which can be viewed as generaliztions of the multiple contrast tests of Mukerjee, Robertson, and Wright (1987). In a quite different way, Chakraborti and Hettmansperger (1996) used suitably defined one-sample confidence intervals to test H_0 versus H_1 by utilizing the priori of H_1.

The second problem we study is making inference with no prior knowledge: in other words, some treatments may be inferior to the control. This situation occurs quite often in real life data. The test of the null hypothesis of homogeneity of the k treatments and the control versus the non-homogeneity has been extensively studied; Dunnett's (1955) procedure is the best known one. The null hypothesis in this chapter is $H_0' : \mu_0 \geq \mu_i$ $(i = 1, \ldots, k)$ and the alternative hypothesis is H_1' : at least one $\mu_0 < \mu_i$. This type of hypothesis is different from the classical null hypothesis of homogeneity.

Interval estimation provides a visual perspective unmatched by a point estimate or a test statistic. The problem of confidence intervals under ordered restrictions has not received much attention in the literature. This is primarily due to the general intractablity of these types of problems (page 405 in Robertson, Wright and Dykstra

(1988)). Only few of the aforementioned test procedures can provide simultaneous confidence bounds (SCB). Marcus (1978) developed a set of simultaneous one-sided confidence bounds in the case of known variance. Marcus and Talpetz (1992) further proposed an alternative test statistic and used it to construct a set of SCB, but their procedure is inferior to the Dunnett's procedure in comparing μ_i to the control mean μ_0. Berk and Marcus (1996) summarized the results of SCB for simple order, simple tree order and umbrella order.

When the null hypothesis is rejected in favour of the alternative hypothesis for the above two cases, there exists at least one treatment better than the control. Let $\mu_{best} = \max_{1 \le i \le k} \mu_i$, which is the mean of the best treatment. Since $\mu_{best} - \mu_0$ is the largest difference between any treatment mean and the control mean, the confidence lower bound for $\mu_{best} - \mu_0$ is bounded below by that for any $\mu_i - \mu_0, (i = 1, \ldots, k)$ or their non-negative linear combinations. If this maximized confidence lower bound for $\mu_{best} - \mu_0$ is positive, then μ_{best} is significantly larger than μ_0. The key is to choose a suitable test statistic such that the positiveness of the maximized confidence lower bound for $\mu_{best} - \mu_0$ is equivalent to the rejection of the null hypothesis by the test statistic. This heuristic forms the basis of our method. However, the likelihood ratio test cannot be used to provide confidence intervals. In this chapter, we propose some test procedures and use them to search for the sharpest simultaneous confidence lower bound for $\mu_{best} - \mu_0$ through efficient algorithms. The construction of the simultaneous confidence lower bound of $\mu_{best} - \mu_0$, as discussed in this chapter, is a particularly useful inference method that has not been considered before. Our method is different from multiple comparison with the best (MCB) proposed by Edwards and Hsu (1983). MCB is the procedure in which the mean for each

treatment is compared to the best of the other treatment means without using any prior information. MCB can be used to provide a simultaneous confidence lower bound for $\mu_{best} - \mu_0$, but it is not as efficient as ours, as illustrated in Section 4.4.

The layout of this chapter is as follows. A test procedure and an iterative algorithm to obtain simultaneous confidence lower bounds for $\mu_{best} - \mu_0$ are given in Section 4.2 for the first problem under the parameter space $\Omega = \{ \boldsymbol{\mu} \in R^{k+1} : \mu_0 \leq \mu_i, i = 1, \ldots, k\}$ and in Section 4.3 for the second problem under the null hypothesis $H_0' : \mu_0 \geq \mu_i$ $(i = 1, \ldots, k)$ respectively. In Section 4.4 a numerical example is presented to illustrate the methods. In Section 4.5 a power comparison is conducted to investigate the behaviors of test procedures T in (4.2) and T^a in (4.12) to that of Dunnett's procedure and S_{01}. In Section 4.6 a brief discussion is given. All proofs can be found in Section 4.7.

4.2 Treatments at Least as Good as a Control

In this section we assume that treatments are at least as good as the control. The parameter space in this case is $\Omega = \{ \boldsymbol{\mu} \in R^{k+1} : \mu_0 \leq \mu_i, i = 1, \ldots, k\}$, where the null hypothesis is H_0: $\mu_0 = \mu_1 = \cdots = \mu_k$ and the alternative hypothesis is $H_1 = \Omega - H_0$.

4.2.1 Likelihood Ratio Test

The estimator of the common value of μ_i under H_0 is $\hat{\mu} = \bar{\bar{Y}}$, where $\bar{\bar{Y}} = \frac{\sum_{i=0}^{k} n_i \bar{Y}_i}{\sum_{i=0}^{k} n_i}$. The restricted maximum likelihood estimator of $\boldsymbol{\mu}$ subject to Ω is denoted by $\boldsymbol{\mu}^* = (\mu_0^*, \mu_1^*, \ldots, \mu_k^*)$. It is called the isotonic regression of $\bar{\mathbf{Y}} = (\bar{Y}_0, \ldots, \bar{Y}_k)$ under

Ω and it can be computed as follows. If $\bar{Y}_0 \leq \bar{Y}_i, i = 1, \ldots, k$, then $\mu_i^* = \bar{Y}_i$. Otherwise, arrange $\bar{Y}_1 \leq \bar{Y}_2 \leq \ldots \leq \bar{Y}_k$ in ascending order ($\bar{Y}_0$ excluded). Denote these values by $\bar{Y}_{[1]} \leq \bar{Y}_{[2]} \leq \ldots \leq \bar{Y}_{[k]}$ and let $n_{[i]}$ denote the corresponding sample size for $i = 1, \ldots, k$, where $\bar{Y}_{[0]} = \bar{Y}_0$ and $n_{[0]} = n_0$. Let l be the smallest non-negative integer such that $A_l = \sum_{i=0}^{l} n_{[i]}\bar{Y}_{[i]} / \sum_{i=0}^{l} n_{[i]} < \bar{Y}_{[l+1]}$, then $\mu_0^* = A_l$, and $\mu_i^* = \max(A_l, \bar{Y}_i)$.

The LRT rejects H_0 in favor of H_1 for large values of

$$S_{01} = \sum_{i=0}^{k} n_i(\mu_i^* - \hat{\mu})^2 / (\sum_{i=0}^{k} n_i(\bar{Y}_i - \mu_i^*)^2/\nu + S^2).$$

The null distribution of S_{01} under H_0 is given by

$$P[S_{01} > s] = \sum_{j=2}^{k+1} P(j, k+1; \mathbf{w}) P\Big[F_{j-1,N-j} > \frac{s(N-j)}{\nu(j-1)}\Big] \tag{4.1}$$

for any $s > 0$, where $N = \sum_{i=0}^{k} n_i$, $\mathbf{w} = (n_0, \ldots, n_k)$, $P(j, k+1; \mathbf{w})$ is the level probability under H_0 that $\boldsymbol{\mu}^*$ takes j distinct values. Through transformation $U_i = \bar{Y}_i - \bar{Y}_0, i = 1, \ldots, k$, we have that

$$P(k+1, k+1; \mathbf{w}) = P(U_1 > 0, \ldots, U_k > 0).$$

The above probability is the orthant probability in which $\boldsymbol{U} = (U_1, U_2, \ldots, U_k)$ has a multivariate normal distribution with zero mean and correlation matrix $\boldsymbol{\rho}$ of the form ($\rho_{ii} = 1$)

$$\rho_{ij} = \Big[\frac{w_i w_j}{(w_0 + w_i)(w_0 + w_j)}\Big]^{1/2}, \text{ for } 1 \leq i \neq j \leq k.$$

For the equal weights case, $P(l, k+1; \mathbf{w})$ and the critical values $s_{k,\nu,\alpha}$ for S_{01} can be found in Robertson, Wright and Dykstra (1988).

4.2.2 Multiple Contrast Test Statistic T

When $S_{01} > s_{k,\nu,\alpha}$, one rejects H_0 and concludes that there is at least one treatment mean μ_i significantly larger than μ_0. However, there is no corresponding simultaneous confidence lower bound for $\mu_i - \mu_0$ when $k > 1$. We introduce the following test statistic,

$$T = \max_{\mathbf{c}\in\mathbf{C}} \sum_{i=0}^{k} n_i c_i \bar{Y}_i / S\Big(\sum_{i=0}^{k} n_i c_i^2\Big)^{1/2}. \tag{4.2}$$

where

$$\mathbf{C} = \Big\{\mathbf{c} = (c_0, c_1, \ldots, c_k) : \sum_{i=0}^{k} n_i c_i = 0,\ c_0 \le c_i,\ i = 1, \ldots, k\Big\}.$$

Let $t_{k,\nu,\alpha}$ be the critical value of T, then

$$P_\mu\Big\{\sum_{i=0}^{k} n_i c_i \mu_i \ge \sum_{i=0}^{k} n_i c_i \bar{Y}_i - t_{k,\nu,\alpha} S\Big(\sum_{i=0}^{k} n_i c_i^2\Big)^{1/2},\ \text{for all } \mathbf{c} \in \mathbf{C}\Big\} = 1 - \alpha. \tag{4.3}$$

The left-hand side of (4.3) can be rewritten as

$$\begin{aligned} & P_\mu\Big\{\max_{\mathbf{c}\in\mathbf{C}} \sum_{i=0}^{k} n_i c_i (\bar{Y}_i - \mu_i)/S\Big(\sum_{i=0}^{k} n_i c_i^2\Big)^{1/2} \le t_{k,\nu,\alpha}, \mu \in R^{k+1}\Big\} \\ = \ & P_0\Big\{\max_{\mathbf{c}\in\mathbf{C}} \sum_{i=0}^{k} n_i c_i \bar{Y}_i / S\Big(\sum_{i=0}^{k} n_i c_i^2\Big)^{1/2} \le t_{k,\nu,\alpha}\Big\} \\ = \ & P_0\Big\{\sum_{i=0}^{k} n_i (\mu_i^* - \hat{\mu})^2 / S^2 \le t_{k,\nu,\alpha}^2\Big\} \end{aligned}$$

and the last identity follows a similar argument as in Hogg (1965). It follows that

$$T^2 = \sum_{i=0}^{k} n_i (\mu_i^* - \hat{\mu})^2 / S^2. \tag{4.4}$$

The right hand side of (4.4) is given by Wright (1988) but it was derived for a different purpose. The statistic T^2 is asymptotically equivalent to S_{01}. The null distribution of T under H_0 is given by

$$P[T \ge t] = \sum_{j=2}^{k+1} P(j, k+1; \mathbf{w}) P\Big[F_{j-1,\nu} \ge \frac{t^2}{j-1}\Big] \tag{4.5}$$

for any $t > 0$. The critical value $t_{k,\nu,\alpha}$ is the value t when one equates (4.5) to α.

4.2.3 Confidence Lower Bound for $\mu_{best} - \mu_0$

According to (4.3), the $1 - \alpha$ simultaneous confidence bound for any contrast $\sum_{i=0}^{k} n_i c_i \mu_i$ with $c_i \geq c_0, i = 1, \ldots, k$, is given by

$$l\Big(\sum_{i=0}^{k} n_i c_i \mu_i\Big) = \sum_{i=0}^{k} n_i c_i \bar{Y}_i - t_{k,\nu,\alpha} S \Big(\sum_{i=0}^{k} n_i c_i^2\Big)^{1/2}. \tag{4.6}$$

Specifically, the $1-\alpha$ simultaneous confidence lower bound for the difference between the ith treatment mean μ_i and the control mean μ_0 is given by

$$l(\mu_i - \mu_0) = \bar{Y}_i - \bar{Y}_0 - t_{k,\nu,\alpha} S (n_i^{-1} + n_0^{-1})^{1/2}. \tag{4.7}$$

Let $\mathcal{K} = \Big\{\mathbf{c} : \mathbf{c} \in \mathbf{C},\ \sum_{i=0}^{k} n_i c_i \mu_i \leq \mu_{best} - \mu_0,\ \text{for all}\ \ \mu \in \Omega\Big\}$. The confidence lower bound for $\mu_{best} - \mu_0$ is given by

$$L(\mu_{best} - \mu_0) = \max_{\mathbf{c} \in \mathcal{K}} l\Big(\sum_{i=0}^{k} n_i c_i \mu_i\Big). \tag{4.8}$$

The following lemma gives another description of the set $\mathcal{K}$ and its proof is trivial.

Lemma 4.2.1 For $\mu \in \Omega$, $\sum_{i=0}^{k} n_i c_i \mu_i \leq \mu_{best} - \mu_0$ if and only if $\sum_{c_i > 0} n_i c_i \leq 1$, for all $\mathbf{c} \in \mathbf{C}$.

The following theorem establishes an equivalence relationship between the positiveness of the above optimal lower bound and rejection of H_0 by statistic T: its proof can be found in Section 4.7.

Theorem 4.2.1 When $\mu \in \Omega$, we have that $T > t_{k,\nu,\alpha}$ if and only if $L(\mu_{best} - \mu_0) > 0$.

When the lower bound (4.8) is positive, it indicates that the best treatment mean is significantly larger than the control mean and it also provides the size of the difference. One may use T to test $H_0^\delta : \mu_0 \leq \mu_i (i = 1, \ldots, k). \mu_{best} - \mu_0 \leq \delta$ versus $H_1^\delta : \mu_{best} - \mu_0 > \delta$ if $L(\mu_{best} - \mu_0) > \delta$. The latter indicates at least one treatment is a "good" treatment: here "good" treatment implies that its mean is greater than the mean of the control by a size of δ. One can use the method in Tong (1969) to classify the treatments as "good" relative to the control. In the remainder of this section, we shall restrict our attention to the case $T > t_{k,\nu,\alpha}$ and we shall relabel the treatments so that $\bar{Y}_1 \leq \bar{Y}_2 \leq \ldots \leq \bar{Y}_k$. The following lemma relates the optimal solution $\mathbf{c}^o$ to (4.8) and the MLE μ^*. Its proof is trivial.

Lemma 4.2.2 Suppose $\bar{Y}_1 \leq \bar{Y}_2 \leq \ldots \leq \bar{Y}_k$. Let l be the nonnegative integer such that $\bar{Y}_l \leq \mu_0^* < \bar{Y}_{l+1}$. If $\mathbf{c}^o$ is the optimal solution to (4.8), then $c_0^o = c_1^o = \cdots = c_l^o \leq c_{l+1}^o \leq \cdots \leq c_k^o$. Furthermore, if $\bar{Y}_i = \bar{Y}_{i+1}$, then $c_i^o = c_{i+1}^o$.

When $\bar{Y}_i = \bar{Y}_{i+1}$, we may group them together with a combined sample size $n_i + n_{i+1}$.

According to the algorithm of computing the MLE μ^* and Lemma 4.2.2, it can be easily shown that the maximum problem in (4.8) is equivalent to the following problem:

$$\max \Big\{ \sum_{i=0}^{k} n_i c_i \mu_i^* - t_{k,\nu,\alpha} S \Big(\sum_{i=0}^{k} n_i c_i^2 \Big)^{1/2} \Big\}, \tag{4.9}$$

subject to $\mathbf{c} \in \mathbf{C}$ and $\sum_{c_i > 0} n_i c_i \leq 1$.

The following theorem establishes a necessary and sufficient condition for an optimal solution and its proof is in Section 4.7.

Theorem 4.2.2 Suppose that $T > t_{k,\nu,\alpha}$. The vector $\mathbf{c}^o \in \mathcal{K}$ is an optimal solution to (4.9) if and only if there exist non-negative integers p and q, $l \le p < q \le k$, such that $\mu_p^* < \hat{\mu} < \mu_q^*$, $S_{0p}^2 + S_{qk}^2 < t_{k,\nu,\alpha}^2 S^2$, $c_i^o = -N_{0p}^{-1} + b^{-1}(\mu_i^* - \bar{Y}_{0p})$, $i = 0, \ldots, p$, $c_{p+1}^o = \ldots = c_{q-1}^o = 0$, $c_i^o = N_{qk}^{-1} + b^{-1}(\mu_i^* - \bar{Y}_{qk})$, $i = q, \ldots, k$, and

$$\max\{N_{0p}(\mu_p^* - \bar{Y}_{0p}), N_{qk}(\bar{Y}_{qk} - \mu_q^*)\} < b \le \min\{N_{0(p+1)}(\mu_{p+1}^* - \bar{Y}_{0(p+1)}), N_{(q-1)k}(\bar{Y}_{(q-1)k} - \mu_{q-1}^*)\}. \tag{4.10}$$

where

$$b^2 = (t_{k,\nu,\alpha}^2 S^2 - S_{0p}^2 - S_{qk}^2)/(N_{0p}^{-1} + N_{qk}^{-1}). \tag{4.11}$$

and

$$N_{ab} = \sum_a^b n_i, \quad \bar{Y}_{ab} = \sum_a^b n_i \mu_i^* / N_{ab}, \quad S_{ab}^2 = \sum_a^b n_i (\mu_i^* - \bar{Y}_{ab})^2$$

When $q = p + 1$, the upper bound for b in (4.10) is replaced by $(\bar{Y}_{qk} - \bar{Y}_{0p})/(N_{0p}^{-1} + N_{qk}^{-1})$.

4.2.4 Iterative Algorithm I

There are $\binom{k-l+1}{2}$ possible choices of p and q, $l \le p < q \le k$. From Theorem 4.2.2 the choice of $p < q$ provides the optimal solution if and only if (4.10) holds. For given $\bar{Y}_0, \bar{Y}_1 \le \ldots \le \bar{Y}_k$ and S^2, we shall show that there are $k - l$ possible choices of (p, q) for the optimal solution $\mathbf{c}^o$, depending upon the confidence level $1 - \alpha$.

The following algorithm selects the optimal solution $(p_0, q_0), (p_1, q_1), \ldots, (p_r, q_r)$ from confidence level $1 - p$ to the desired level $1 - \alpha$, where p is the p-value of the statistic T.

(0) Set $i = 0$, $p_0 = \max\{0 \le j < k : \mu_j^* < \hat{\mu}\}$ and $q_0 = \min\{1 \le j \le k : \mu_j^* > \hat{\mu}\}$.

(i) Let $\mathcal{B}_{i+1} = \max\{N_{0p_i}(\mu^*_{p_i} - \bar{Y}_{0p_i}), N_{q_ik}(\bar{Y}_{q_ik} - \mu^*_{q_i})\}$, $t_{k,\nu,\alpha_{i-1}} = \{S^2_{0p_i} + S^2_{q_ik} + (N^{-1}_{0p_i} + N^{-1}_{q_ik})\mathcal{B}^2_{i-1}\}^{1/2}/S$. If $t_{k,\nu,\alpha_{i-1}} < t_{k,\nu,\alpha}$, the optimal solution is $\mathbf{c}^\circ$ with $p = p_i$ and $q = q_i$. Otherwise, go to (ii).

(ii) If $N_{0p_i}(\mu^*_{p_i} - \bar{Y}_{0p_i}) > N_{q_ik}(\bar{Y}_{q_ik} - \mu^*_{q_i})$, then set $p_{i+1} = \max\{j : 0 \le j < p_i, \mu^*_j < \mu^*_{p_i}\}$ and $q_{i+1} = q_i$. Otherwise, set $p_{i-1} = p_i$ and $q_{i+1} = \min\{j : q_i < j \le k, \mu^*_j > \mu^*_{q_i}\}$. Set $i = i + 1$, go to step (i).

The justification of the above algorithm can be found in Section 4.7.

4.3 No Prior Ordering on Treatments And a Control

In this section, we consider the case when the experimenters have no prior knowledge of treatment means and the control mean. Experimenters wish to know whether there are any treatments superior to the control and if so, how much is the difference between the best treatment and the control?

4.3.1 Union-intersection Method

In this case, the null hypothesis is $H'_0 : \mu_0 \ge \mu_i$ $(i = 1, \ldots, k)$ and the alternative hypothesis is H'_1 : at least one $\mu_0 < \mu_i$. The null hypothesis H'_0 can be expressed as infinite intersections $H'_0 = \cap_{\mathbf{c}\in\mathbf{C}^\circ} H'_{0\mathbf{c}}$, here $H'_{0\mathbf{c}} : \sum_0^k n_i c_i \mu_i \le 0$, with $\mathbf{c} \in \mathbf{C}^\circ$ and

$$\mathbf{C}^\circ = \Big\{\mathbf{c} = (c_0, c_1, \ldots, c_k) : \sum_{i=0}^{k} n_i c_i = 0, \quad c_i \ge 0, i = 1, \ldots, k\Big\}.$$

The alternative hypothesis H'_1 can be expressed as infinite unions $H'_1 = \cup_{\mathbf{c}\in\mathbf{C}^\circ} H'_{1\mathbf{c}}$, where $H'_{1\mathbf{c}} : \sum_0^k n_i c_i \mu_i > 0$. The rationale behind this union-intersection method is

simple. If any one of $H'_{0\mathbf{c}}$ is rejected, then H'_0, which is true only if $H'_{0\mathbf{c}}$ is true for every $\mathbf{c} \in \mathbf{C}^o$, must also be rejected. Only if each of the hypotheses $H'_{0\mathbf{c}}$ is accepted as true will the intersection of H'_0 be accepted as true. For each $H'_{0\mathbf{c}}$ versus $H'_{1\mathbf{c}}$, the test statistic used is $T^o_{\mathbf{c}} = \sum_{i=0}^{k} n_i c_i \bar{Y}_i / \{S(\sum_{i=0}^{k} n_i c_i^2)^{1/2}\}$, where $\mathbf{c} \in \mathbf{C}^o$. Suppose the corresponding rejection region for $T^o_{\mathbf{c}}$ has the form $\{y : T^o_{\mathbf{c}} > a\}$, where a is a constant which does not depend on index $\mathbf{c}$. Then the rejection region for the union-intersection test is

$$\cup_{\mathbf{c}\in\mathbf{C}^o}\{y : T^o_{\mathbf{c}} > a\} = \{y : \sup_{\mathbf{c}\in\mathbf{C}^o} T^o_{\mathbf{c}} > a\}.$$

Accordingly, the test statistic for testing H'_0 is $\sup_{\mathbf{c}\in\mathbf{C}^o} T^o_{\mathbf{c}}$ which will be denoted by T^o as in (4.12) below. By the union-intersection method, T^o is used to test H'_0 versus H'_1. Note that T^{o2} is the LRT of H'_0 versus H'_1.

The null hypothesis H'_0 can also be expressed as the finite intersections $H'_0 = \cap_{i=1}^{k} H'_{0i}$, where $H'_{0i} : \mu_0 \geq \mu_i$, then the alternative hypothesis can be expressed correspondingly as the finite unions $H'_1 = \cup_{i=1}^{k} H'_{1i}$, with $H'_{1i} : \mu_0 < \mu_i$. For each H'_{0i} versus H'_{1i}, the test statistic is $D_i = (\bar{Y}_i - \bar{Y}_0)/\{S(n_i^{-1} + n_0^{-1})^{1/2}\}$. Suppose the corresponding rejection region has the form $\{y : D_i > d\}$, where d is a constant which does not depend on index i. Hence the rejection region for the union-intersection test is

$$\cup_i\{y : D_i > d\} = \{y : \max_i D_i > d\}$$

Then the test statistic for testing H'_0 is $\max_i D_i$, which is Dunnett's test statistic $D = \max_{1\leq i\leq k}(\bar{Y}_i - \bar{Y}_0)/\{S(n_i^{-1} + n_0^{-1})^{-1/2}\}$. Notice that $T^o \geq D$. The latter is a special case of the former. When $c_{i0} = -1/n_0, c_{ii} = 1/n_i, c_{ij} = 0, j = 1, \ldots, i-1, i+1, \ldots, k$, then H'_{0i} is in the form of $H'_{0\mathbf{c}}$. Both T^o and D can be used to test H'_0 versus H'_1 and

each has its own advantages. As we shall see in Section 4.5, when there is only one treatment better than the control, Dunnett's test statistic D is the right choice, but T^o is the one to use when there is more than one treatment better than the control.

4.3.2 Multiple Contrast Test Statistic T^o

According to the above union-intersection method, we introduce a "new" test statistic to test H_0' against H_1':

$$T^o = \max_{\mathbf{c}\in\mathbf{C}^o} \sum_{i=0}^{k} n_i c_i \bar{Y}_i / S\Big(\sum_{i=0}^{k} n_i c_i^2\Big)^{1/2}. \tag{4.12}$$

It is easy to see that $T^o \geq 0$.

Denote the critical value of T^o by $t^o_{k,\nu,\alpha}$, then

$$P_\mu\Big\{\sum_{i=0}^{k} n_i c_i \mu_i \geq \sum_{i=0}^{k} n_i c_i \bar{Y}_i - t^o_{k,\nu,\alpha} S\Big(\sum_{i=0}^{k} n_i c_i^2\Big)^{1/2}, \text{for all } \mathbf{c}\in\mathbf{C}^o\Big\} = 1-\alpha. \tag{4.13}$$

Let r be the largest integer for which $\bar{Y}_{r-1} < \bar{Y}_{r,k+1}$, where $\bar{Y}_{r,k+1} = (n_0\bar{Y}_0 + N_{rk}\bar{Y}_{rk})/(n_0+N_{rk})$. Let $\mu_i^o = \hat{\mu}$ for $i = 1,\ldots,r-1$, $\mu_i^o = \bar{Y}_i - \bar{Y}_{r,k+1} + \hat{\mu}$ for $i = 0, r, r+1, \ldots, k$. If $\bar{Y}_0 \geq \bar{Y}_k$, then $\mu_i^o = \hat{\mu}$ for $i = 0, 1, \ldots, k$ and $T^o = 0$.

The left-hand side of (4.13) can be rewritten as

$$\begin{aligned} & P_\mu\Big\{\max_{\mathbf{c}\in\mathbf{C}^o}\sum_{i=0}^{k} n_i c_i(\bar{Y}_i-\mu_i)/S\Big(\sum_{i=0}^{k} n_i c_i^2\Big)^{1/2} \leq t^o_{k,\nu,\alpha}, \mu\in R^{k+1}\Big\} \\ = \; & P_0\Big\{\max_{\mathbf{c}\in\mathbf{C}^o}\sum_{i=0}^{k} n_i c_i \bar{Y}_i/S\Big(\sum_{i=0}^{k} n_i c_i^2\Big)^{1/2} \leq t^o_{k,\nu,\alpha}\Big\} \\ = \; & P_0\Big\{\sum_{i=0}^{k} n_i\big(\mu_i^o-\hat{\mu}\big)^2/S^2 \leq t^{o2}_{k,\nu,\alpha}\Big\} \end{aligned}$$

and the last identity follows a similar argument as in Hogg (1965). Thus, we have

$$T^{o2} = \sum_{i=0}^{k} n_i(\mu_i^o-\hat{\mu})^2/S^2. \tag{4.14}$$

The null hypothesis distribution of T^o under the least favorable configuration H_0 of H_0' is given by

$$\sup_{\mu \in H_0'} P[T^o \geq t] = \sum_{j=1}^{k} P(j, k+1; \mathbf{w}) P\Big[F_{k+1-j,\nu} \geq \frac{t^2}{k+1-j}\Big] \tag{4.15}$$

for any $t > 0$. The statistic T^{o2} has the same distribution as statistic S_{12} in Robertson, Wright and Dykstra (1988). The critical value $t^o_{k,\nu,\alpha}$ of T^o is the square root of the corresponding one of S_{12}. The latter can be found in Table A.9 of Robertson, Wright and Dykstra (1988). Note that $t^o_{k,\nu,\alpha} > d_{k,\nu,\alpha}$, where $d_{k,\nu,\alpha}$ is the critical value of Dunnett's procedure D.

4.3.3 Confidence Lower Bound for $\mu_{best} - \mu_0$

The $1-\alpha$ one-sided simultaneous confidence bound for the contrast $\sum_{i=0}^{k} n_i c_i \mu_i$ with $\mathbf{c} \in \mathbf{C^o}$ is given by

$$l^o\Big(\sum_{i=0}^{k} n_i c_i \mu_i\Big) = \sum_{i=0}^{k} n_i c_i \bar{Y}_i - t^o_{k,\nu,\alpha} S \Big(\sum_{i=0}^{k} n_i c_i^2\Big)^{1/2}. \tag{4.16}$$

Specifically, the $1-\alpha$ one-sided simultaneous confidence bound between the ith treatment mean μ_i and the control mean μ_0 is given by

$$l^o(\mu_i - \mu_0) = \bar{Y}_i - \bar{Y}_0 - t^o_{k,\nu,\alpha} S (n_i^{-1} + n_0^{-1})^{1/2} \tag{4.17}$$

Let $\mathcal{K}^o = \Big\{\mathbf{c} : \mathbf{c} \in \mathbf{C^o},\ \sum_{i=0}^{k} n_i c_i \mu_i \leq \mu_{best} - \mu_0\Big\}$. The simultaneous confidence lower bound for $\mu_{best} - \mu_0$ is given by

$$L^o(\mu_{best} - \mu_0) = \max_{\mathbf{c} \in \mathcal{K}^o} l^o\Big(\sum_{i=0}^{k} n_i c_i \mu_i\Big). \tag{4.18}$$

The following theorem establishes an equivalence relationship between the positiveness of the above optimal lower bound and rejection of H_0' by statistic T^o. Its proof is similar to Theorem 4.2.1, hence is omitted.

Theorem 4.3.1 $T^o > t^o_{k,\nu,\alpha}$ if and only if $L^o(\mu_{best} - \mu_0) > 0$.

One may use T^o to test $H_0'^{\delta} : \mu_0 + \delta \geq \mu_i (i = 1, \ldots, k)$ versus $H_1'^{\delta} : \mu_0 + \delta < \mu_{best}$ if $L^o(\mu_{best} - \mu_0) > \delta$. The positive maximized lower bound indicates the significant difference between the best treatment mean and the control mean as well as the size of this difference. It also suggests that at least one treatment is a "good" treatment. The following theorem establishes a necessary and sufficient condition for an optimal solution to (4.18) and its proof, similar to that of Theorem 4.2.2, is omitted.

Theorem 4.3.2 Suppose $T^o > t^o_{k,\nu,\alpha}$ and $\bar{Y}_1 \leq \bar{Y}_2 \leq \ldots \leq \bar{Y}_k$. Then $\mathbf{c}^o \in \mathcal{K}^o$ is an optimal solution to (4.18) if and only if there exists a positive q, $r \leq q \leq k$ such that $c_0^o = -\frac{1}{n_0}$, $c_1^o = \ldots = c_{q-1}^o = 0$, $c_i^o = N_{qk}^{-1} + b^{-1}(\mu_i^o - \bar{Y}_{qk})$, $i = q, \ldots, k$, and

$$N_{qk}(\bar{Y}_{qk} - \mu_q^o) < b \leq N_{(q-1)k}(\bar{Y}_{(q-1)k} - \mu_{q-1}^o) \tag{4.19}$$

where

$$b^2 = (t^{o2}_{k,\nu,\alpha} S^2 - S^2_{qk})/(n_0^{-1} + N_{qk}^{-1}).$$

and

$$\bar{Y}_{qk} = \sum_q^k n_i \mu_i^o / N_{qk}, \quad S^2_{qk} = \sum_q^k n_i (\mu_i^o - \bar{Y}_{qk})^2$$

When $q = r$, the upper bound for b in (4.19) is replaced by $(\bar{Y}_{qk} - \bar{Y}_0)/(n_0^{-1} + N_{qk}^{-1})$.

4.3.4 Iterative Algorithm II

There are k possible choices of q, $1 \leq q \leq k$. Suppose that $T^o > t^o_{k,\nu,\alpha}$, from Theorem 4.3.2, for given $\bar{Y}_0, \bar{Y}_1 \leq \ldots \leq \bar{Y}_k$ and S^2, there are $k - r + 1$ possible choices of q for the optimal solution $\mathbf{c}^o$, depending upon the confidence level $1 - \alpha$. The optimal solution q can be obtained iteratively in a few steps by the following algorithm.

(0) Set $i = 0$ and $q_0 = r$.

(i) Let $\beta_{i+1} = N_{q_i,k}(\bar{Y}_{q_i,k} - \mu^o_{q_i})$. $t^o_{k,\nu,\alpha_{i+1}} = \{S^2_{q_i,k} + \beta^2_i(n_0^{-1} + N^{-1}_{q_i,k})\}^{1/2}/S$. If $t^o_{k,\nu,\alpha_{i+1}} \leq t^o_{k,\nu,\alpha}$, the optimal solution is $\mathbf{c}^o$ with $q = q_i$. Otherwise, go to (ii).

(ii) Set $q_{i+1} = \min\{j : q_i < j \leq k, \mu^o_j > \mu^o_{q_i}\}$ and set $i = i + 1$. Go to (i).

4.4 A Numerical Example

For the purpose of illustration, we consider the data in Table 4.1. (see Ruberg (1995)). The six treatments means are 23.9, 27.7, 33.4, 74.4, 73.4, and 73.5, respectively, while the control mean is 25.5. The pooled mean square error is $S^2 = 47.16$ with the degrees of freedom $\nu = 35$. After relabelling, the sample means are $\bar{Y}_0 = 25.5, \bar{Y}_1 = 23.9, \bar{Y}_2 = 27.7, \bar{Y}_3 = 33.4, \bar{Y}_4 = 73.4, \bar{Y}_5 = 73.5$, and $\bar{Y}_6 = 74.4$.

4.4.1 Treatments at Least as Good as the Control

The statistic T for testing $H_0 : \mu_0 = \mu_1 = \ldots = \mu_6$ against the alternative hypothesis $H_1 : \mu_0 \leq \mu_i$ $(i = 1, \ldots, 6)$ with at least one strict inequality has a value of 21.70 with p-value 0.000 and $\mu^*_i = \bar{Y}_i$ except that $\mu^*_0 = \mu^*_1 = 24.7$. Since $t_{6,35,.05} = 3.41$, one concludes that μ_{best} is significantly larger than μ_0. The $(1 - \alpha)100\%$ simultaneous confidence lower bound $L(\mu_{best} - \mu_0)$ under H_1 can be computed as follows.

Step 0: $p_0 = 3$, $q_0 = 4$, $\beta_1 = 138.6$, $t_{6,35,\alpha_1} = 6.80$.

Step 1: $p_1 = 2$, $q_1 = 4$, $\beta_2 = 36$, $t_{6,35,\alpha_2} = 1.97 < t_{6,35,.05}$, stop.

For 95% confidence level, one has the simultaneous confidence lower bound $L(\mu_{best} - \mu_0) = 40.53$, where the optimal coefficient is

$$\mathbf{c}^o = (-0.0703, -0.0703, -0.0260, 0.0000, 0.0501, 0.0516, 0.0649).$$

Note that we do not use the information from the third treatment since $c_3 = 0$.

Suppose our confidence level is 50% instead. Then the corresponding critical value is $t_{6,35,.5} = 1.94$. Since $t_{6,35,\alpha_2} = 1.97 > t_{6,35,.5}$, one continues to the next step.

Step 2: $p_2 = 1$, $q_2 = 4$, $\beta_3 = 6.6$, $t_{6,35,\alpha_3} = 0.48 < 1.94$

The 50% simultaneous confidence lower bound $L(\mu_{best} - \mu_0) = 44.18$, where the optimal coefficient is

$$\mathbf{c}^o = (-0.0833, -0.0833, 0.0000, 0.0000, 0.0452, 0.0480, 0.0735).$$

4.4.2 No Prior Knowledge of Treatments and the Control

The test statistic T^o for testing $H_0' : \mu_0 \geq \mu_i$ $(i = 1, \cdots, 6)$ versus $H_1' : \mu_i > \mu_0$ for some i has a value of 14.91 with p-value 0.000 and

$\boldsymbol{\mu}^o = (11.2, 47.4, 47.4, 47.4, 59.1, 59.2, 60.1)$. Since $t^o_{6,35,.05} = 2.52$, one concludes that μ_{best} is significantly larger that μ_0. The $(1-\alpha)100\%$ simultaneous confidence lower bound $L^o(\mu_{best} - \mu_0)$ under H_1' can be computed as follows.

Step 0: $q_0 = r = 4$, $\beta_1 = 6.6$, $t^o_{6,35,\alpha_1} = 0.53 < 2.52$.

For 95% confidence level, one has the simultaneous confidence lower bound $L(\mu_{best} - \mu_0) = 40.13$, where the optimal coefficient is

$$\mathbf{c}^o = (-0.1667, 0.0000, 0.0000, 0.0000, 0.0455, 0.0482, 0.0729).$$

Note that the optimal lower bound does not use the information from the first three treatments.

To test H_0 versus H_1 and H_0' versus H_1', one may use Dunnett's procedure. Since $D = \max_{1\le i\le 6} \frac{\bar{Y}_i - \bar{Y}_0}{S/\sqrt{3}} = 12.3 > d_{6,35,.05} = 2.39$, one rejects H_0 and H_0' at $\alpha = .05$ significance level. The corresponding simultaneous confidence lower bound is 39.42.

Edwards and Hsu (1983) introduced mutliple comparisons with best to construct confidence interval for $\mu_i - \max_{0\le j\le k} \mu_j$ $(i = 0, 1, \ldots, k)$. According to their method, the 95% simultaneous confidence lower bound for $\mu_{best} - \mu_0$ is 38.20.

Hence, Dunnett's procedure D and Edwards and Hsu (1983)'s MCB are not as good as statistics T and T^o to detect the difference between the control and treatments. The statistic T provides the sharpest simultaneous confidence lower bound for $\mu_{best} - \mu_0$ with T^o a close second.

4.5 Power Comparisons

The behavior of the power functions of S_{01}, T, T^o, and Dunnett's procedure D are investigated. The Monte Carlo method is used with 10,000 iterations. For simplicity, we consider equal sample size case with $\nu = 20$. The simulated powers are provided in Table 4.2 and Table 4.3 for the two cases in Section 4.5.1 and Section 4.5.2 respectively, where $\Delta^2 = n\sum_{\mu_i\ge\mu_o}(\mu_i - \bar{\mu})^2/\sigma^2$ is the noncentrality parameter, here $\bar{\mu} = \sum_{\mu_i\ge\mu_o}\mu_i/(k+1)$.

4.5.1 Treatments at Least as Good as the Control

Three cases are considered in Table 4.2: Case 1, the center direction $(-k, 1, \ldots, 1)$, which means that all treatments are effective and their effects are approximately equal; Case 2, pairwise comparison $(-1, 1, 0, \ldots, 0)$, which means that all treatments are effective but one treatment is more effective than the other treatments; and Case 3, the edge direction $(-k/2, \ldots, -k/2, k/2+1, \ldots, k/2+1)$ which consists of one half treatment means with values $-k/2$ and the other half with value $k/2 + 1$; in other words, one half of the treatments are effective while the remaining half treatments are ineffective. When the prior information states that treatments are at least as effective as the control, one may use the above four test statistics. The maximum and minimum powers for S_{01} are conjectured to occur, respectively, at the center of the cone Ω (i.e., Case 1) and at the edges (i.e., Case 3) (See Robertson, Wright and Dykstra (1988)).

The test statistic T^o is shown to be the most powerful one-sided test along the center direction when all treatments are better than the control. For example, at $k = 6$, the power of T^o are 6.65%, 17.59%, 19.68%, and 10.19% higher than those of S_{01}, for $\Delta = 1, 2, 3$, and 4, respectively. T^o is also the most powerful one along pairwise directions $(-1, 1, 0, \ldots, 0)$, ..., or $(-1, 0, \ldots, 0, 1)$ when treatment means are larger than the control mean and one treatment mean is larger than the remaining treatment means. In this case, S_{01} and T are not as powerful as T^o and D. The powers of T^o and D are very close. Therefore, we recommend T^o for Case 1 and Case 2.

However, S_{01} has the highest power along the edge direction among these four

test statistics. The difference in power for S_{01} over D can be as large as 17.95%. The power of T is the second highest and very close to that of S_{01} in this case.

Since statistic T has competitive power performance and it can provide confidence lower bound, statistic T is recommended for statistical inference under Case 3.

4.5.2 No Prior Ordering on Treatments and the Control

Cases 4, 5, 6 are considered in Table 4.3, which are $(-k+1, 1, \ldots, 1, -k)$, $(-1, 1, 0, \ldots, 0, -2)$, and $(-(k-1)/2, \ldots, -(k-1)/2, (k+1)/2, \ldots, (k+1)/2, -k)$, respectively. They have one more extra non-effective treatment than the corresponding Cases 1, 2, 3 respectively. The last treatment is not effective in comparison to the control. These three cases apply to test statistics T^o and D only. The statistic T^o is more powerful than the Dunnett procedure D in Case 4, but in Case 5. The Dunnett procedure D has higher power than that of T^o. In Case 6, when $k = 3$, Dunnett's test statistic D has slightly higher power than that of T^o; howerever, statistic T^o tends to have larger power than that of D as k and Δ increase.

Based on the results in Table 4.3, the following recommendations are made: when there is more than one good treatment, T^o is the optimum choice for testing H_0' versus H_1'; when there is only one good treatment, D is the right choice for for testing H_0' versus H_1'.

4.6 Discussion

Two different sets of hypotheses are considered. For H_0 versus H_1, when every treatment is effective, statistic T^o is recommended, but when some treatments are not effective, statistic T is recommended. For H_0' versus H_1', when there is more than one good treatment, statistic T^o is recommended; when there is only one good treatment, Dunnett's test D is suggested. The major advantage of test statistics T and T^o is that they have a convenient conversion to simultaneous confidence lower bounds for the difference between the best treatment mean and the control mean. Once the optimal confidence lower bound is larger than a threshold (the threshold depending on the experimenter's experience), the next step is to identify the best treatment or select those "good" treatments. There are several methods to achieve this purpose, for example, Gupta's subset selection (1956, 1965), Bechhofer's indifference zone selection (1954), Lam's procedure for selecting good populations (1986), Hsu's multiple comparisons with the best (1996), and among others.

Theorems 4.2.2 and 4.3.2 utilizing the Kuhn-Tucker equivalence theorem are the keys to the optimization problems and the proposed algorithms. This approach can also be applied to other optimization problems such as umbrella restrictions.

4.7 Proofs

4.7.1 Proof of Theorem 4.2.1

If $T > t_{k,\nu,\alpha}$, then there exists a $\mathbf{c_1} \in C$ such that

$$l\Big(\sum_{i=0}^{k} n_i c_{1i}\mu_i\Big) = \sum_{i=0}^{k} n_i c_{1i}\bar{Y}_i - t_{k,\nu,\alpha} S\Big(\sum_{i=0}^{k} n_i c_{1i}^2\Big)^{1/2} > 0.$$

Without loss of generality, one may assume that $\sum_{i=0}^{k} n_i|c_{1i}| = 2$. It is trivial that $\sum_{i=0}^{k} n_i c_{1i}\mu_i \leq \mu_{best} - \mu_0$ when $\mu_i \geq \mu_0, i = 1, \ldots, k$. Therefore, $L(\mu_{best} - \mu_0) \geq l\Big(\sum_{i=0}^{k} n_i c_{1i}\mu_i\Big) > 0$.

On the other hand, if $L(\mu_{best} - \mu_0) > 0$, then there exists a $\mathbf{c_2} \in \mathcal{K}$ such that

$$L(\mu_{best} - \mu_0) = l\Big(\sum_{i=0}^{k} n_i c_{2i}\mu_i\Big) > 0.$$

This leads to

$$T = \max_{\mathbf{c}\in\mathbf{C}} \sum_{i=0}^{k} n_i c_i \bar{Y}_i / S\Big(\sum_{i=0}^{k} n_i c_i^2\Big)^{1/2} \geq \sum_{i=0}^{k} n_i c_{2i} \bar{Y}_i / S\Big(\sum_{i=0}^{k} n_i c_{2i}^2\Big)^{1/2} > t_{k,\nu,\alpha}.$$

4.7.2 Proof of Theorem 4.2.2

According to Lemma 4.2.1 and Lemma 4.2.2, when $\bar{Y}_1 \leq \bar{Y}_2 \leq \ldots \leq \bar{Y}_k$, the constraint in (4.8) can be replaced by $\sum_{i=0}^{k} n_i c_i = 0, \sum_{i=j}^{k} n_i c_i \leq 1, j = 1, \ldots, k$. It is trivial that $\sum_{i=0}^{k} n_i c_i \mu_i^* - t_{k,\nu,\alpha} S(\sum_{i=0}^{k} n_i c_i^2)^{1/2}$ is a concave function of $c_0, c_1, \ldots, c_k$. Let

$$\phi(c, v, \lambda) = \sum_{i=0}^{k} n_i c_i \mu_i^* - t_{k,\nu,\alpha} S\Big(\sum_{i=0}^{k} n_i c_i^2\Big)^{1/2} + \sum_{j=1}^{k} v_j\Big(1 - \sum_{r=j}^{k} n_r c_r\Big) - \lambda \sum_{i=0}^{k} n_i c_i.$$

Let $\frac{\partial \phi}{\partial \mathbf{c}^o}$ denote the partial derivatives evaluated at the point $\mathbf{c}^o, \mathbf{v}^o$, and λ^o. By the Kuhn-Tucker equivalence theorem (Kuhn and Tucker (1951)), $\mathbf{c}^o$ is the optimal solution if and only if

(i) $\frac{\partial \phi}{\partial c_i^o} = n_i \mu_i^* - n_i c_i^o b - n_i \sum_{j=1}^{i} v_j^o - \lambda^o n_i = 0, (i = 0, \ldots, k)$,

where $b = t_{k,\nu,\alpha} S / (\sum_{i=0}^{k} n_i c_i^{o2})^{1/2}$,

(ii) $\sum_{r=j}^{k} n_r c_r^o \leq 1$ $(j = 1, \ldots, k)$, $(\frac{\partial \phi}{\partial \mathbf{v}^o})' \mathbf{v}^o = \mathbf{0}$, $\mathbf{v}^o \geq \mathbf{0}$ and $\frac{\partial \phi}{\partial \lambda^o} = 0$.

Suppose $\mathbf{c}^o$ is the optimal solution. By Lemma 4.2.2, $\mathbf{c}_i^o$ is monotone nondecreasing, and there exists $l \leq p < q$ such that

$$c_0^o = c_1^o = \ldots = c_l^o < c_{l+1}^o \leq \ldots \leq c_p^o < c_{p+1}^o = \ldots = c_{q-1}^o = 0 < c_q^o \leq \ldots \leq c_k^o,$$

with

$$\sum_{i=0}^{p} n_i c_i^o = -1, \text{ and } \quad \sum_{i=q}^{k} n_i c_i^o = 1.$$

From (ii),

$$v_i^o = 0, i = 1, \ldots, p, q+1, \ldots, k.$$

From (i),

$$c_i^o = b^{-1}\Big(\mu_i^* - \sum_{j=1}^{i} v_j^o - \lambda^o\Big). \tag{4.20}$$

Adding the first $p+1$ equations in (4.20) and using $\sum_{i=0}^{p} n_i c_i^o = -1$, we obtain

$$\lambda^o = \bar{\Upsilon}_{0p} + b\bar{N}_{0p}^{-1}.$$

Substituting λ^o into (i), then we have

$$c_i^o = -\bar{N}_{0p}^{-1} + b^{-1}(\mu_i^* - \bar{\Upsilon}_{0p}), i = 0, 1, \ldots, p.$$

Let $V = \sum_{i=p+1}^{q} v_i^o$, using $\sum_{i=q}^{k} n_i c_i^o = 1$, then $V = \bar{\Upsilon}_{qk} - b\bar{N}_{qk}^{-1} - \lambda^o$, and

$$c_i^o = \bar{N}_{qk}^{-1} + b^{-1}(\mu_i^* - \bar{\Upsilon}_{qk}), i = q, \ldots, k.$$

It follows that $b = t_{k,\nu,\alpha} S / (\sum_{i=0}^{k} n_i c_i^{o2})^{1/2}$ can be written as

$$b^2 = \frac{t_{k,\nu,\alpha}^2 S^2 - S_{0p}^2 - S_{qk}^2}{\bar{N}_{0p}^{-1} + \bar{N}_{qk}^{-1}}.$$

In order to prove that (4.10) is true, there are two cases that need to be considered, namely, $q > p+1$ and $q = p+1$. For the case $q > p+1$, from (4.20),

$v^o_{p+1} = \mu^*_{p+1} - \bar{\Upsilon}_{0p} - bN^{-1}_{0p}$, $v^o_{p+2} = \mu^*_{p+2} - \mu^*_{p+1} \geq 0, \ldots, v^o_{q-1} = \mu^*_{q-1} - \mu^*_{q-2} \geq 0$, $v^o_q = \bar{\Upsilon}_{qk} - \mu^*_{q-1} - bN^{-1}_{qk}$. By (ii), $v^o_{p+1} \geq 0$ and $v^o_q \geq 0$, then

$$\begin{aligned} b &\leq \min\{N_{0p}(\mu^*_{p+1} - \bar{\Upsilon}_{0p}), N_{qk}(\bar{\Upsilon}_{qk} - \mu^*_{q-1})\} \\ &= \min\{N_{0(p+1)}(\mu^*_{p+1} - \bar{\Upsilon}_{0(p+1)}), N_{(q-1)k}(\bar{\Upsilon}_{q-1k} - \mu^*_{q-1})\}. \end{aligned}$$

For the case $q = p + 1$, since

$$\begin{aligned} v^o_q &= \bar{\Upsilon}_{qk} - bN^{-1}_{qk} - \bar{\Upsilon}_{0p} - bN^{-1}_{0p} \\ &= \bar{\Upsilon}_{qk} - \bar{\Upsilon}_{0p} - b(N^{-1}_{0p} + N^{-1}_{qk}) \geq 0. \end{aligned}$$

thus,

$$\begin{aligned} b &\leq (\bar{\Upsilon}_{qk} - \bar{\Upsilon}_{0p})/(N^{-1}_{0p} + N^{-1}_{qk}) \\ &= (\bar{\Upsilon}_{qk} - \bar{\Upsilon}_{0p})/(N^{-1}_{0(p+1)} + N^{-1}_{qk}) \end{aligned}$$

On the other hand, $c^o_p < 0$ and $c^o_q > 0$, then

$$b > \max\{N_{0p}(\mu^*_p - \bar{\Upsilon}_{0p}), N_{qk}(\bar{\Upsilon}_{qk} - \mu^*_q)\}.$$

Hence, (4.10) follows.

Next, we prove that $\mu^*_p < \hat{\mu} < \mu^*_q$ is true. According to the algorithm of computing μ^* in section 4.2.1 and the assumption that $\bar{\Upsilon}_1 \leq \bar{\Upsilon}_2 \leq \ldots \leq \bar{\Upsilon}_k$, we have that

$$\mu^*_0 = \ldots = \mu^*_l < \mu^*_{l+1} = \bar{\Upsilon}_{l+1} \leq \ldots \leq \mu^*_p = \bar{\Upsilon}_p \leq \mu^*_q = \bar{\Upsilon}_q \leq \ldots \leq \mu^*_k = \bar{\Upsilon}_k.$$

For the case $q = p + 1$, using $c^o_p < 0$, $c^o_q > 0$ and $b \leq \frac{\bar{\Upsilon}_{qk} - \bar{\Upsilon}_{0p}}{N^{-1}_{0p} + N^{-1}_{qk}}$, we can easily derive $\mu^*_p < \hat{\mu} < \mu^*_q$.

For the case $q > p+1$, we have proved that $b \le N_{(q-1)k}(\bar{Y}_{(q-1)k} - \mu^*_{q-1})$. From $c^o_p < 0$,

$$N_{0p}(\mu^*_p - \bar{Y}_{0p} < N_{(q-1)k}(\bar{Y}_{(q-1)k} - \mu^*_{q-1}),$$

in other words,

$$N_{0p}\mu^*_p + N_{q-1k}\mu^*_{q-1} < N_{0p}\bar{Y}_{0p} + N_{(q-1)k}\bar{Y}_{(q-1)k}.$$

Notice that $\mu^*_p \le \mu^*_{p+1} = \bar{Y}_{p+1} \le \ldots \le \mu^*_k = \bar{Y}_k$. from the above ineqality, we obtain that $\mu^*_p < \hat{\mu}$ in the case of $q > p+1$: likewise for $\hat{\mu} < \mu^*_q$.

4.7.3 Justification of Iterative Algorithm I

At Step 0, let $p = p_0, q = q_0, \beta_0 = (\bar{Y}_{qk} - \bar{Y}_{0p})/(N^{-1}_{0p} + N^{-1}_{qk})$ and let $t_{k,\nu,\alpha_0} = T$. Then α_0 is the p-value. It is trivial that

$$t^2_{k,\nu,\alpha_0} = \{S^2_{0p} + S^2_{qk} + (N^{-1}_{0p} + N^{-1}_{qk})\beta^2_0\}/S^2.$$

When $t_{k,\nu,\alpha_0} \ge t_{k,\nu,\alpha} > t_{k,\nu,\alpha_1}$, one has that $\alpha_0 \le \alpha < \alpha_1$, and $\beta_0 \ge b(\alpha) > \beta_1$ with $b(\alpha)$ given by (4.11). It follows that (4.10) holds at Step 0.

Suppose that $p_1 = p-1$ and $q_1 = q$. Then $\beta_1 = N_{0p}(\mu^*_p - \bar{Y}_{0p})$. It is trivial that

$$\begin{aligned} t^2_{k,\nu,\alpha_1}S^2 &= S^2_{0p} + S^2_{qk} + (N^{-1}_{0p} + N^{-1}_{qk})\beta^2_1 \\ &= S^2_{0p-1} + N_{0p-1}(\bar{Y}_{0p-1} - \bar{Y}_{0p})^2 + n_p(\mu^*_p - \bar{Y}_{0p})^2 \\ &\quad + S^2_{qk} + (N^{-1}_{0p} + N^{-1}_{qk})N^2_{0p}(\mu^*_p - \bar{Y}_{0p})^2 \\ &= S^2_{0p-1} + S^2_{qk} + (N^{-1}_{0p-1} + N^{-1}_{qk})\beta^2_1. \end{aligned}$$

Therefore, $t_{k,\nu,\alpha} \le t_{k,\nu,\alpha_1}$ implies $\alpha \ge \alpha_1$ and $\beta_1 \ge b(\alpha)$. Similarly for the case $p_1 = p$ and $q_1 = q+1$. By induction, one obtains the desired p_i and q_i such that (4.10) holds for the $b(\alpha)$ of a given level $1-\alpha$.

Table 4.1: Dose Response Data.

Group	Sample size	Mean response	SD response
P	6	25.5	2.6
A	6	23.9	4.0
B	6	27.7	3.3
C	6	33.4	2.3
D	6	74.4	14.6
E	6	73.4	7.6
F	6	73.5	4.5

Table 4.2: Simulated Powers for Four Test Statistics with $\alpha = 0.05, \nu = 20$.

			Test			
Direction	k	Δ	S_{01}	T	T^o	D
Case 1						
	2	1	19.75	19.23	22.96	22.28
		2	50.83	49.96	56.62	55.14
		3	80.74	80.07	85.71	83.90
		4	96.53	96.35	98.03	97.31
	4	1	16.26	15.00	21.58	21.14
		2	41.74	39.57	55.58	53.07
		3	71.13	68.60	83.59	81.09
		4	91.53	90.36	96.89	95.64
	6	1	14.69	13.08	21.34	19.96
		2	35.14	31.70	52.73	49.79
		3	63.28	59.54	82.96	79.25
		4	86.29	83.69	96.48	95.01
Case 2						
	2	1	18.74	18.28	20.51	20.04
		2	50.25	49.42	53.41	53.24
		3	81.08	80.52	83.58	83.46
		4	96.29	96.15	97.11	97.07
	4	1	14.97	14.32	18.11	17.76
		2	40.80	38.82	48.31	47.62
		3	70.86	68.64	78.58	78.23
		4	91.38	90.37	95.49	95.41
	6	1	13.19	11.98	17.15	16.86
		2	33.13	30.17	43.76	42.29
		3	62.27	58.64	75.06	74.66
		4	85.86	83.56	93.92	94.02
Case 3						
	2	1	16.39	16.23	15.10	15.33
		2	44.95	44.96	38.85	39.75
		3	78.42	78.77	69.85	70.97
		4	95.73	95.78	90.92	91.52
	4	1	12.78	12.57	12.05	12.32
		2	33.09	32.41	29.01	29.18
		3	64.19	64.04	52.61	52.59
		4	88.49	88.47	77.00	76.40
	6	1	10.93	10.71	11.23	10.92
		2	26.24	25.91	22.10	22.04
		3	54.26	53.65	42.13	41.05
		4	82.28	81.93	66.26	64.33

Table 4.3: Simulated Powers for Test Statistics T^a and D with $\alpha = 0.05, \nu = 20$.

Direction	k	Δ	Test T^a	Test D
Case 4				
	3	1	17.96	17.89
		2	49.12	48.48
		3	80.26	78.94
		4	96.40	95.74
	5	1	19.29	18.45
		2	51.32	49.36
		3	80.86	78.21
		4	95.91	94.58
	7	1	19.58	18.89
		2	50.22	47.67
		3	80.79	77.62
		4	95.84	94.37
Case 5				
	3	1	15.93	16.23
		2	46.13	46.93
		3	78.26	79.02
		4	95.78	95.95
	5	1	15.71	15.76
		2	44.48	44.19
		3	75.50	75.51
		4	94.26	94.57
	7	1	15.70	15.72
		2	40.88	40.66
		3	72.75	72.89
		4	92.88	93.27
Case 6				
	3	1	11.65	12.06
		2	32.09	34.28
		3	62.69	65.13
		4	87.31	88.80
	5	1	10.31	10.63
		2	25.60	26.24
		3	48.47	49.10
		4	73.66	73.57
	7	1	10.25	10.17
		2	20.35	20.54
		3	39.50	39.04
		4	63.52	62.36

Chapter 5

Cone Order Monotonicity of Tests for Treatments versus a Control

5.1 Introduction

A problem frequently encountered in the practice of statistics is comparing several treatment means with a control mean, or a standard. This problem has been received considerable attention in statistical literature over the past fifty years, of which the best known is Dunnett's (1955) multiple comparison procedure.

Let $Y_{ij}, i = 0, 1, \ldots, k, j = 1, \ldots, n_i$ be independent normal variates with unknown means μ_i $(i = 0, 1, \ldots, k)$ and a common but unknown variance σ^2, where μ_0 denotes the control mean and $\mu_1, \ldots, \mu_k$ denote the treatment means. The statistic $S^2 = \sum_{i=0}^{k} \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2/\nu$ is used as an estimator for σ^2, and it is independent of $\bar{\mathbf{Y}} = (\bar{Y}_0, \ldots, \bar{Y}_k)$, where $\nu S^2/\sigma^2 \sim \chi^2_\nu$ and $\nu = \sum_{i=0}^{k} n_i - k - 1 > 0$. When comparing treatments versus a control, in many situations, experimenters may have the prior knowledge that each treatment mean is at least as large as the control mean, or each treatment mean is at least as large as the grand mean. These types of prior knowledge may come from the past experiences. For example, the same treatments

and the control may have been studied in previous experiments as part of an ongoing investigation, and therefore the above prior information might be available on all of them. If we assume that all of the treatments are at least as good as the control, then the parameter space is $\Omega = \{\mu \in R^{k+1} : \mu_0 \leq \mu_i \ (i = 1, \ldots, k)\} = \{\mu : A\mu \geq 0\}$, where A is a $(k-1) \times k$ matrix whose rows are pairwise contrasts with

$$A = \begin{pmatrix} -1 & 1 & 0 & 0 & \ldots & 0 \\ -1 & 0 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -1 & 0 & 0 & 0 & \ldots & 1 \end{pmatrix}.$$

Robertson, Wright and Dykstra (1988) refers to the binary relationship as the simple tree order. Usually the null hypothesis is $H_0 : \mu_0 = \mu_1 = \cdots = \mu_k$ and the alternative hypothesis is $H_1 = \Omega - H_0$. The likelihood ratio test statistic rejects H_0 in favor of H_1 for large values of

$$S_{01} = \sum_{i=0}^{k} n_i(\mu_i^* - \hat{\mu})^2 / \Big(\sum_{i=0}^{k} n_i(\bar{Y}_i - \mu_i^*)^2/\nu + S^2 \Big),$$

where $\hat{\mu} = \sum_{i=0}^{k} n_i \bar{Y}_i / \sum_{i=o}^{k} n_i$ and $\boldsymbol{\mu}^* = (\mu_0^*, ..., \mu_k^*)$ is the MLE under Ω, i.e., $\boldsymbol{\mu}^*$ minimizes

$$\sum_{i=0}^{k} n_i(\bar{Y}_i - \mu_i)^2$$

subject to the restriction $\boldsymbol{\mu} \in \Omega$. Lee (1988) recognized a shortcoming of the MLE $\boldsymbol{\mu}^*$ which has been introduced in Chapter 1. The LRT S_{01} for testing H_0 against H_1 is known to possess generally superior operating characteristics to those of its competitors, see Robertson, Wright and Dykstra (1988). A variety of other procedures have also been proposed, most of which are based on one or more contrasts among the sample means, for example, the multiple contrast test of Mukerjee, Robertson,

and Wright (1987) which includes Dunnett's procedure and the single-contrast test as special cases. Tang and Lin (1997) used an orthant to approximate S_{01}.

However, the LRT S_{01} lacks some practical monotonicity property, as observed by Cohen, Kemperman, and Sackrowitz (2000). Any closed convex cone C can induce a quasiordering $\prec$ as follows: $\mathbf{x} \prec_C \mathbf{y}$ if and only if $\mathbf{y} - \mathbf{x} \in C$, where $\mathbf{x}, \mathbf{y} \in R^{k+1}$. If $\mathbf{x}$ and $\mathbf{y}$ are sample points such that $\mathbf{y} - \mathbf{x} \in C$, then our intuition tell us that $\mathbf{y}$ exhibits a greater level of agreement with cone C than does $\mathbf{x}$. A test ϕ is said to be cone order monotone with respect to the cone C (COM[C]) if whenever $\mathbf{y} - \mathbf{x} \in C$, $\phi(\mathbf{x}) \leq \phi(\mathbf{y})$. Tests which are not COM[C] are said to be reverse. The cone order monotone property has appeared in Robertson and Wegman (1978).

The parameter space Ω is a closed convex cone, see Rockafellar (1972). Cohen, Kemperman, and Sackrowitz (2000) noted that the LRT S_{01} is not COM[Ω]. A simple example of the reversal phenomenon can be seen in the following. Let $k = 2$, $n_0 = n_1 = n_2 = 1$, and σ^2 is known. In this case S_{01} is $\bar{\chi}^2_{01}$ and we have that $\bar{\chi}^2_{01} = 24$ at $\mathbf{x} = (0, 0, 6)$ and $\bar{\chi}^2_{01} = 18$ at $\mathbf{y} = (0, 3, 6)$. Therefore, the LRT S_{01} is not COM[Ω]. Our instincts tell us that the sample point $\mathbf{y}$ is making stronger statement than $\mathbf{x}$ and should be reflected in the inference procedures. A lack of COM for a test procedure may be counter-intuitive and undesirable.

The reason that S_{01} is not COM[Ω] is due to the fact that the angles between the corners of the cone Ω are obtuse. Since likelihood inference is the most common approach for order restricted models, it is of interest to find a likelihood inference which is cone order monotone for comparison of treatments with a control. In this chapter, we consider a new parameter space Ω^o, which is of practical importance; here $\Omega^o = \{\mu : \mu_0 \leq \bar{\mu} \leq \mu_i, i = 1, \ldots, k\}$, where $\bar{\mu} = \sum_{i=0}^{k} n_i \mu_i / \sum_{i=0}^{k} n_i$ is the grand

mean. When $n_0 = n_1 = \ldots = n_k$, Ω^o can be rewritten as $\Omega^o = \{\mu : A\mu \geq 0\}$, where

$$A = \begin{pmatrix} -1 & k & -1 & -1 & \ldots & -1 \\ -1 & -1 & k & -1 & \ldots & -1 \\ -1 & -1 & -1 & k & \ldots & -1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -1 & -1 & -1 & -1 & \ldots & k \end{pmatrix}.$$

Obviously, each row of A has $k+1$ nonzero elements and the result in Cohen, Kemperman and Sackrowitz (1994) cannot be used here. We develop a likelihood ratio statistic S_{01}^o and its modified test statistic MLRT T_{01}^o to test H_0 versus H_1^o : $\Omega^o - H_0$.

In the parameter space Ω^o all treatment means are at least as large as the grand mean, and Ω^o does not have any obtuse angles between the edges. We shall show that the statistics S_{01}^o and T_{01}^o are COM[Ω^o].

Cohen and Sackrowitz (1998) offered a new test procedure which is COM[Ω], but as Cohen and Sackrowitz (2000) mentioned, the shortcoming of this new test procedure is that its distribution is not clear and its critical values must be obtained through simulation. In this chapter, we propose a multiple contrast test procedure T^o which is COM[Ω] and COM[Ω^o]; moreover, its distribution and critical values are available in the literature.

The outline of this chapter is as follows. In Section 5.2 we introduce the LRT S_{01}^o. In Section 5.3 we give the MLRT T_{01}^o and the multiple contrast test procedure T^o. In Section 5.4 we discuss the behavior of the power functions of four test statistics S_{01}, S_{01}^o, T^o and Dunnett's test statistic D.

5.2 Cone Order Monotonicity of LRT S_{01}^o

Let $\boldsymbol{\mu}^o = (\mu_0^o, ..., \mu_k^o)$ be the MLE under Ω^o, i.e., $\boldsymbol{\mu}^o$ minimizes

$$\sum_{i=0}^{k} n_i(\bar{Y}_i - \mu_i)^2$$

subject to the restriction $\boldsymbol{\mu} \in \Omega^o$. The solution $\boldsymbol{\mu}^o$ is a projection of $\bar{\mathbf{Y}}$ onto Ω^o with weights $\mathbf{w} = (n_0, \ldots, n_k)$ and is denoted by $\boldsymbol{\mu}^o = P_{\mathbf{w}}(\bar{\mathbf{Y}}|\Omega^o)$. It can be computed as follows. Without loss of generality, one may assume that $\bar{Y}_1 \le \bar{Y}_2 \le \ldots \le \bar{Y}_k$. Let r be the largest positive integer for which $\bar{Y}_{r-1} < (\sum_{i=r}^{k} n_i \bar{Y}_i + n_0 \bar{Y}_0)/(\sum_{i=r}^{k} n_i + n_0)$. Denote the right hand side of this inequality by $\bar{Y}_{r,k+1}$. Then $\mu_i^o = \hat{\mu}$ for $i = 1, \ldots, r-1$. $\mu_i^o = \bar{Y}_i - \bar{Y}_{r,k+1} + \hat{\mu}$ for $i = 0, r, r+1, \ldots, k$. If $\bar{Y}_0 \ge \bar{Y}_{r,k+1}$, then $\mu_i^o = \hat{\mu}$ for $i = 0, 1, \ldots, k$. Let us use the scab index data on page 97 of Cochran and Cox (1957) to illustrate how to use the algorithm. The sample means are labeled as $\bar{Y}_0 = 22.6, \bar{Y}_1 = 18.2, \bar{Y}_2 = 16.8, \bar{Y}_3 = 15.5, \bar{Y}_4 = 14.2, \bar{Y}_5 = 9.5, \bar{Y}_6 = 5.8$. This is the case where $\Omega^o = \{\mu : \mu_0 \ge \bar{\mu} \ge \mu_i, i = 1, \ldots, k\}$. According to the algorithm, $r = 5$, where $\bar{Y}_{r-1} > \bar{Y}_{r,k+1} = 15.125$. The expression for $\boldsymbol{\mu}^o$ remains the same and it is $\boldsymbol{\mu}^o$ $= (23.31, 15.65, 15.65, 15.65, 15.65, 14.91, 10.21, 6.51)$.

One rejects H_0 in favor of H_1^o for large values of

$$S_{01}^o = \sum_{i=0}^{k} n_i(\mu_i^o - \hat{\mu})^2 / \Big(\sum_{i=0}^{k} n_i(\bar{Y}_i - \mu_i^o)^2/\nu + S^2 \Big) \tag{5.1}$$

The null distribution of S_{01}^o under H_0 is given by

$$P[S_{01}^o > c] = \sum_{l=1}^{k} P(l, k+1; \mathbf{w}) P\Big[F_{k+1-l,\nu+l-1} > \frac{c(\nu+l-1)}{\nu(k+1-l)} \Big] \tag{5.2}$$

for $c > 0$, where $\mathbf{w} = (n_0, \ldots, n_k)$. $P(l, k+1; \mathbf{w})$ is the level probability that the MLE $\boldsymbol{\mu}^*$ under Ω takes l distinct values when H_0 is true. The values of c and

$P(l, k+1; \mathbf{w})$ are available respectively in Table A.11 and Table A.13 of Robertson, Wright and Dykstra (1988).

The dual $\mathcal{K}^*$ of a closed convex cone $\mathcal{K}$ is defined as follows:

$$\mathcal{K}^* = \{f : \sum_{i=0}^{k} w_i f_i g_i \leq 0, \text{ for all } g \in \mathcal{K}\}.$$

where $\mathbf{w} = (n_0, \ldots, n_k)$. Let $P_{\mathbf{w}}(.|\mathcal{K})$ denote the projection onto $\mathcal{K}$ with weights $\mathbf{w}$ and $||.||_{\mathbf{w}}$ be the usual weighted norm. The following two lemmas are used to prove Theorem 5.2.1.

Lemma 5.2.1 Let $\mathcal{L}$ be a linear space, and $\mathcal{L} \subset \mathcal{K}$, here $\mathcal{K}$ is a closed convex cone. Then $P_{\mathbf{w}}(\mathbf{y} + \mathbf{c}|\mathcal{K}) = P_{\mathbf{w}}(\mathbf{y}|\mathcal{K}) + \mathbf{c}$ if $\mathbf{c} \in \mathcal{L}$ and for any $\mathbf{y} \in R^{k+1}$.

Lemma 5.2.2 Let $\mathcal{K}^*$ be the dual of $\mathcal{K}$. If $\boldsymbol{\delta} \in \mathcal{K}^*$, then for any $\mathbf{y} \in R^{k+1}$, $||P_{\mathbf{w}}(\mathbf{y} + \boldsymbol{\delta}\ |\mathcal{K})||_{\mathbf{w}} \leq ||P_{\mathbf{w}}(\mathbf{y}|\mathcal{K})||_{\mathbf{w}}$.

The proofs for these two lemmas are trivial.

Theorem 5.2.1 The LRT S_{01}^o is COM$[\Omega^o]$.

Proof. Let $\bar{Y}_i \sim N(0, \frac{\sigma^2}{n_i})$, $\bar{Y}_i^{\mu_1} = \bar{Y}_i + \mu_{1i}$, $\bar{Y}_i^{\mu_2} = \bar{Y}_i + \mu_{2i}$, $\boldsymbol{\mu}_1^o = P_{\mathbf{w}}(\bar{Y}^{\mu_1}|\Omega^o)$, $\boldsymbol{\mu}_2^o = P_{\mathbf{w}}(\bar{Y}^{\mu_2}|\Omega^o)$, where $\boldsymbol{\lambda} = \boldsymbol{\mu}_2 - \boldsymbol{\mu}_1 \in \Omega^o$. Then $\bar{Y}^{\mu_2} = \bar{Y}^{\mu_1} + \lambda$. Denote $\hat{\sigma}^{o2}(\mu_1) = \sum_{i=0}^{k} n_i(\bar{Y}_i^{\mu_1} - \mu_{1i}^o)^2/\nu + S^2$ and similarly for $\hat{\sigma}^{o2}(\mu_2)$. Since Ω^o is a closed convex cone, $\lambda + P_{\mathbf{w}}(\bar{Y}^{\mu_1}|\Omega^o) \in \Omega^o$, then

$$\begin{aligned} ||\bar{Y}^{\mu_2} - P_w(\bar{Y}^{\mu_2}|\Omega^o)||_{\mathbf{w}} &= ||\bar{Y}^{\mu_1} + \lambda - P_{\mathbf{w}}((\bar{Y}^{\mu_1} + \lambda)|\Omega^o)||_{\mathbf{w}} \\ &\leq ||\bar{Y}^{\mu_1} + \lambda - [P_{\mathbf{w}}(\bar{Y}^{\mu_1}|\Omega^o) + \lambda]||_{\mathbf{w}} \\ &= ||\bar{Y}^{\mu_1} - P_{\mathbf{w}}(\bar{Y}^{\mu_1}|\Omega^o)||_{\mathbf{w}}. \end{aligned}$$

The above inequality comes from the definition of projection and it implies that

$$\hat{\sigma}^{o2}(\mu_1) \geq \hat{\sigma}^{o2}(\mu_2).$$

From Lemma 5.2.1,

$$\begin{aligned} P_{\mathbf{w}}(\bar{\Upsilon}^{\cdot\mu_1} - P_{\mathbf{w}}(\bar{\Upsilon}^{\cdot\mu_1}|H_0) + \lambda^{(1)}|\Omega^o) &= P_{\mathbf{w}}(\bar{\Upsilon}^{\cdot\mu_1} + \lambda - P_{\mathbf{w}}(\bar{\Upsilon}^{\cdot\mu_1} + \lambda|H_0)|\Omega^o) \\ &= P_{\mathbf{w}}(\bar{\Upsilon}^{\cdot\mu_1} + \lambda|\Omega^o) - P_{\mathbf{w}}(\bar{\Upsilon}^{\cdot\mu_1} + \lambda|H_0) \\ &= P_{\mathbf{w}}(\bar{\Upsilon}^{\cdot\mu_2}|\Omega^o) - P_{\mathbf{w}}(\bar{\Upsilon}^{\cdot\mu_2}|H_0) \end{aligned}$$

where $\lambda^{(1)} = \lambda - P_{\mathbf{w}}(\lambda|H_0)$. Applying Lemma 5.2.2 with $\delta = -\lambda^{(1)}$, then

$$\begin{aligned} ||P_{\mathbf{w}}(\bar{\Upsilon}^{\cdot\mu_2}|\Omega^o) - P_{\mathbf{w}}(\bar{\Upsilon}^{\cdot\mu_2}|H_0)||_{\mathbf{w}} &= ||P_{\mathbf{w}}(\bar{\Upsilon}^{\cdot\mu_1} - P_{\mathbf{w}}(\bar{\Upsilon}^{\cdot\mu_1}|H_0) - \delta|\Omega^o)||_{\mathbf{w}} \\ &\geq ||P_{\mathbf{w}}(\bar{\Upsilon}^{\cdot\mu_1} - P_{\mathbf{w}}(\bar{\Upsilon}^{\cdot\mu_1}|H_0)|\Omega^o)||_{\mathbf{w}} \\ &= ||P_{\mathbf{w}}(\bar{\Upsilon}^{\cdot\mu_1}|\Omega^o) - P_{\mathbf{w}}(\bar{\Upsilon}^{\cdot\mu_1}|H_0)||_{\mathbf{w}}. \end{aligned}$$

Hence, we have that

$$\begin{aligned} \sum_{i=0}^{k} n_i(\mu_{1i}^o - \hat{\mu}_1)^2 &= ||P_{\mathbf{w}}(\bar{\Upsilon}^{\cdot\mu_1}|\Omega^o) - P_{\mathbf{w}}(\bar{\Upsilon}^{\cdot\mu_1}|H_0)||_{\mathbf{w}}^2 \\ &\leq ||P_{\mathbf{w}}(\bar{\Upsilon}^{\cdot\mu_2}|\Omega^o) - P_{\mathbf{w}}(\bar{\Upsilon}^{\cdot\mu_2}|H_0)||_{\mathbf{w}} \\ &= \sum_{i=0}^{k} n_i(\mu_{2i}^o - \hat{\mu}_2)^2 \end{aligned}$$

Therefore,

$$S_{01}^o(\mu_1) = \frac{\sum_{i=0}^{k} n_i(\mu_{1i}^o - \hat{\mu}_1)^2}{\hat{\sigma}^{o2}(\mu_1)} \leq S_{01}^o(\mu_2).$$

This completes the proof.

5.3 Cone Order Monotonicity of Test Statistic T^o

From the LRT S^o_{01}, a more straightforward approach to testing H_0 versus H^o_1 is to replace the denominator of S^o_{01} by S^2. One rejects H_0 for large values of

$$T^o_{01} = \sum_{i=0}^{k} n_i(\mu^o_i - \hat{\mu})^2/S^2. \tag{5.3}$$

We call T^o_{01} the modified likelihood ratio test (MLRT). The null distribution of T^o_{01} is as follows:

$$P[T^o_{01} \geq c] = \sum_{l=1}^{k} P(l, k+1; \mathbf{w}) P\Big[F_{k+1-l, N-k-1} \geq \frac{c}{k+1-l}\Big] \tag{5.4}$$

for any $c > 0$. The LRT S^o_{01} and the MLRT T^o_{01} are asymptotically equivalent. The statistic T^o_{01} has the same distribution as statistic S_{12} in Robertson, Wright and Dykstra (1988). The critical values of T^o_{01} are the same as the corresponding ones of S_{12}. The latter can be found in their Table A.9.

Next we consider a multiple contrast test statistic T^o as follows:

$$T^o = \max_{\mathbf{c} \in \mathcal{C}^o} \sum_{i=0}^{k} n_i c_i \bar{Y}_i / S\Big(\sum_{i=0}^{k} n_i c_i^2\Big)^{1/2}. \tag{5.5}$$

where the contrast cone $\mathcal{C}^o$ which corresponds to Ω^o is defined as follows.

$$\mathcal{C}^o = \Big\{\mathbf{c} = (c_0, c_1, \ldots, c_k) : \sum_{i=0}^{k} n_i c_i = 0. \quad c_i \geq 0. \, i = 1, \ldots, k\Big\}.$$

It is easy to see that $T^o \geq 0$.

Following the argument in Hogg (1965), one can show that

$$T^{o2} = \sum_{i=0}^{k} n_i(\mu^o_i - \hat{\mu})^2/S^2 = T^o_{01}. \tag{5.6}$$

The critical values of T^o is the square root of the corresponding ones of S_{12}. Moreover, T^o is cone order monotone for Ω and Ω^o and can be used to make interval inference.

Theorem 5.3.1 The statistic T^o is COM[Ω] and COM[Ω^o].

Proof. Let $\bar{Y}_i \sim N\left(0, \frac{\sigma^2}{n_i}\right)$, $\bar{Y}_i^{\mu_1} = \bar{Y}_i + \mu_{1i}$, $\bar{Y}_i^{\mu_2} = \bar{Y}_i + \mu_{2i}$, and $\mu_2 - \mu_1 \in \Omega$ (or Ω^o). Then,

$$T^o(\mu_2) = \max_{\mathbf{c}\in\mathcal{C}^o} \frac{\sum_{i=0}^k n_i c_i \bar{Y}_i^{\mu_2}}{S\left(\sum_{i=0}^k n_i c_i^2\right)^{1/2}} = \max_{\mathbf{c}\in\mathcal{C}^o} \frac{\sum_{i=0}^k n_i c_i \bar{Y}_i^{\mu_1} + \sum_{i=0}^k n_i c_i(\mu_{2i} - \mu_{1i})}{S\left(\sum_{i=0}^k n_i c_i^2\right)^{1/2}}.$$

Since $\mu_2 - \mu_1 \in \Omega$ (or Ω^o), then $\mu_{2i} - \mu_{1i} \geq \mu_{20} - \mu_{10}$, $i = 1, \ldots, k$. Notice that $\mathbf{c} \in \mathcal{C}^o$, $\sum_{i=0}^k n_i c_i(\mu_{2i} - \mu_{1i}) \geq 0$. Thus, $T^o(\mu_2) \geq T^o(\mu_1)$. Hence, T^o is COM[Ω] (or COM[Ω^o]).

5.4 Power Comparisons

A simulation study is conducted to compare the behavior of the power functions of S_{01}, S_{01}^o, T^o, and Dunnett's test statistic D. For simplicity, we consider equal sample size case with $\nu = 60$. Three cases are considered. They are the center direction $(-k, 1, \ldots, 1)$, the edge direction $(-1, k, -1, \ldots, -1)$, and the direction of pairwise comparison $(-1, 1, 0, \ldots, 0)$ which lies in the middle of the center direction and the edge direction. Let $\Delta^2 = n\sum_{i=0}^k(\mu_i - \bar{\mu})^2/\sigma^2$ be the noncentrality parameter, here $\bar{\mu} = \sum_{i=0}^k \mu_i/(k+1)$. Within each direction we consider $k = 2, 4, 6$ and $\Delta = 1, 2, 3, 4$. For each configuration, 10,000 multivariate normal random vector with the identity variance-covariance matrix are generated. Table 5.1 gives the powers for the above four test statistics for these cases. These three cases satisfy the hypothesis that treatment means are at least as large as the control mean. Case 1 and Case 2 also satisfy the hypothesis that treatment means are at least as large as the grand mean.

The LRT S_{01}^{o} is shown to be the most powerful along the center direction when all treatments are better than the control, i.e., every treatment is effective, and along the pairwise directions $(-1, 1, 0, \ldots, 0), \ldots,$ or $(-1, 0, \ldots, 0, 1)$ when treatment means are larger than the control mean and one treatment mean is larger than the remaining treatment means. The test statistic T^{o} also has larger powers than that of the LRT S_{01} in these two cases. And statistic T^{o} outperforms Dunnett's procedure D in Case 1 and in Case 2 when $\Delta = 1, 2, 3$. The power of statistic T^{o} is very close to LRT S_{01}^{o}. However, along the edge direction, S_{01} outperforms the other three statistics.

Based on the power results in Table 5.1, we can make the following recommendation, when treatment means are at least as good as the grand mean, one may use LRT S_{01}^{o} or T^{o}. Although LRT S_{01} is not cone order monotone, when some treatments are not effective, one should use LRT S_{01}.

Table 5.1: Simulated Powers with $\alpha = 0.05, \nu = 60$.

Direction	k	Δ	Test			
			S_{01}	S_{01}^a	T^a	D
Case 1						
	2	1	21.15	24.93	24.70	24.22
		2	53.23	59.59	59.38	57.98
		3	84.81	88.83	88.63	87.50
		4	97.50	98.44	98.38	98.04
	4	1	17.17	24.12	23.85	22.06
		2	43.46	58.53	57.67	54.96
		3	75.18	87.43	86.72	84.27
		4	93.92	97.84	97.77	97.01
	6	1	14.82	23.51	22.82	21.84
		2	37.45	57.36	56.46	53.84
		3	68.34	87.74	86.86	83.77
		4	90.57	98.01	97.87	97.11
Case 2						
	2	1	20.15	22.24	22.08	21.81
		2	52.33	55.35	55.31	55.21
		3	84.39	86.66	86.68	86.57
		4	97.33	97.99	97.97	98.00
	4	1	16.07	19.45	19.34	18.66
		2	41.96	49.81	49.91	48.80
		3	75.14	81.87	81.91	81.57
		4	93.61	96.41	96.27	96.49
	6	1	13.58	18.30	18.08	17.46
		2	36.41	46.91	46.57	45.72
		3	67.09	79.48	79.21	78.71
		4	91.14	95.88	95.83	95.94
Case 3						
	2	1	16.75	15.54	15.68	15.85
		2	47.40	39.74	40.52	41.77
		3	81.63	71.61	73.12	74.43
		4	96.75	91.39	92.44	93.10
	4	1	12.11	10.37	10.64	10.97
		2	34.36	25.76	26.74	28.27
		3	68.54	51.18	54.29	57.37
		4	91.50	76.27	79.70	82.31
	6	1	9.77	8.16	8.53	8.96
		2	28.12	19.20	20.74	23.23
		3	58.30	40.41	43.96	48.04
		4	86.70	67.03	72.28	76.40

Chapter 6

On the Test for Equivalence of Treatments with respect to a Control

6.1 Introduction

A problem frequently encountered in pharmaceutical and other fields is the comparison of $k+1$ populations $\Pi_0, \Pi_1, \ldots \Pi_k$, where Π_0 is a standard or control population to be compared with the remaining k treatment populations. This problem has received considerable attention in statistics over the past fifty years and the best known method is Dunnett's multiple comparison procedure (1955). Specifically, assume that we have independent observations Y_{ij} from $k+1$ normal populations with unknown means μ_i $(i = 0, 1, \ldots, k, j = 1, \ldots, n_i)$ and a common but unknown variance σ^2, where μ_0 denotes the control mean and $\mu_1, \ldots, \mu_k$ denote the treatment means. We may assume a larger treatment effect μ_i implies a better treatment. Let $\mu_{best} = \max_{1 \le i \le k} \mu_i$ represent the mean of the best treatment. We shall consider the case of equal sample size n for each treatment but

allowing for a different number of observations n_0 under the control in this chapter. Let $S^2 = \sum_{i=0}^{k} \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2/\nu$. It is independent of $\bar{\mathbf{Y}} = (\bar{Y}_0, \ldots, \bar{Y}_k)$ and $\nu S^2/\sigma^2 \sim \chi^2_\nu$ and $\nu = \sum_{i=0}^{k} n_i - (k+1) = n_0 + kn - (k+1) > 0$, where sample means $\bar{Y}_i = \sum_{j=1}^{n_i} Y_{ij}/n_i$. It is well known that, for large enough sample size, the null hypothesis of the equality of treatment means and the control mean $\mu_0 = \mu_1 = \ldots = \mu_k$ will almost always be rejected as has been pointed out by many researchers (see Berger (1985)) if the underlying distribution is continuous, since irrational values cannot be recorded with perfect precision. The point null hypothesis is unrealistic in applications. Bofinger and Bofinger (1993) studied the test of a null hypothesis that $\mu_i - \mu_0$ exceeds a preassigned constant for some treatment with the alternative of all the differences $\mu_i - \mu_0$ being no more than the constant. Bofinger and Bofinger (1995) developed stepwise tests for this type of hypotheses. Giani and Straßburger (1994) also considered the problem using the two-sided Dunnett's procedure.

In many situations there is a prior knowledge that the treatments are at least as effective as the control. This type of prior knowledge may come from past experiences or it may arise in the experiments where, for example, a higher dose level of a drug has larger effect on the patients. The prior knowledge can be expressed as $\mu_0 \leq \mu_i, i = 1, \cdots, k$, which is typically called the simple tree order in order restricted inference. Robertson and Wegman (1978) proposed a likelihood ratio test to test the null hypothesis as $\mu_0 \leq \mu_i$ $(i = 1, \cdots, k)$, which is not a point null hypothesis but a restricted hypothesis such as H_1 versus H_2 in Chapter 2. However, in some applications, although we have the prior knowledge that the treatments are at least as effective as the control, all the treatments may not be significantly better than the control; in other words, the best treatment mean, $\mu_{best} = \max_{1 \leq i \leq k} \mu_i$

may not be substantially better than the control mean, μ_0, for practical purposes. Thus, it is of interest to establish a practical equivalence of all the treatments with respect to the control μ_0 under the prior knowledge $\mu_0 \leq \mu_i$. Establishing this equivalence is the purpose of this chapter. As an example, among the drugs available for HIV, some may be too expensive to be affordable to people in underdeveloped nations whose funds for medical programs are limited and where treatment costs are prohibitive. Because to date no drug has been very effective in fighting HIV, it is of interest to choose cheaper drugs with comparable treatment effect. This idea leads to the consideration of interval hypothesis $H_0^I : (\mu_{best} - \mu_0)/\sigma \leq \delta$ versus $H_1^I : (\mu_{best} - \mu_0)/\sigma > \delta$, where $\delta \geq 0$ is a prespecified critical threshold value, possibly given by drug evaluation guidelines. The constant δ can be interpreted as the amount of variability about which we do not care, and the null hypothesis can be explained as saying that there is little difference or there is practical equality among treatments and the control. For each treatment the threshold value δ specifies the largest effect difference from the control that is not worth detecting for practical purposes. For the interval hypothesis in k populations without a control, one can refer to Bau, Chen, and Xiong (1993) and Chen and Lam (1991).

In Section 6.2, a test statistic is proposed and its critical value computation is discussed. The distribution of the test statistic depends on all mean differences between treatments and the control. The least favourable configuration (LFC) of these parameters which maximizes the significance level occurs at $(\mu_0, \mu_0+\sigma\delta, \ldots, \mu_0+\sigma\delta)$ under $\mu_i \geq \mu_o$ $(i = 1, \ldots, k)$. Since the distribution under LFC in H_0^I involves double integrals, the double integrals is evaluated by a 64-point Gaussian-Legendre quadrature over 4 by 4 grids. For the special case of $\delta = 0$, the percentage points agree

with existing tables (e.g. Bechhofer and Dunnett (1988)) to three decimal places.

We also construct lower confidence bounds for $d(\mu)$ in Section 6.3. Such confidence bounds can be used to assist the experimenter to choose δ when it is difficult to specify δ in advance.

6.2 Testing Equivalence of Treatment Means and the Control Mean under $\mu_i \geq \mu_0$

The null hypothesis and the alternative hypothesis are respectively

$$H_0^I : (\mu_{best} - \mu_0)/\sigma \leq \delta \text{ and } H_1^I : (\mu_{best} - \mu_0)/\sigma > \delta.$$

where $\delta \geq 0$ is a preassigned constant. Based on Dunnett's (1955) multiple comparison procedure, the test statistic considered is:

$$D = \frac{\max_{1 \leq i \leq k} \bar{Y}_i - \bar{Y}_0}{S\sqrt{\frac{1}{n_0} + \frac{1}{n}}}.$$

The null hypothesis is rejected if $D > c$, where the critical value $c > 0$ is chosen such that the nominal level $\alpha \in (0, 1)$ will be kept. To evaluate the critical value c for which the maximum Type I error probability just attains the level α, it is necessary to determine the supremum of the power $P_{(\mu,\sigma)}(c)$ over H_0^I, i.e. $\sup_{H_0^I} P_{(\mu,\sigma)}(c) = \alpha \in (0, 1)$, where

$$P_{(\mu,\sigma)}(c) = P\Big\{ \max_{1 \leq i \leq k} \bar{Y}_i - \bar{Y}_0 > cS\sqrt{\frac{1}{n_0} + \frac{1}{n}} \Big\}.$$

The supremum of the probability of rejecting H_0^I is calculated as follows:

$$P_{(\mu,\sigma)}(c) = P\Big\{ \max_{1 \leq i \leq k} \bar{Y}_i - \bar{Y}_0 > cS\sqrt{\frac{1}{n_0} + \frac{1}{n}} \Big\}$$

$$
\begin{aligned}
&= 1 - P\Big\{ \max_{1\le i\le k} \bar{Y}_i - \bar{Y}_0 \le cS\sqrt{\frac{1}{n_0}+\frac{1}{n}}\Big\} \\
&= 1 - P\Big\{ \max_{1\le i\le k} \bar{Y}_i \le \bar{Y}_0 + cS\sqrt{\frac{1}{n_0}+\frac{1}{n}}\Big\} \\
&= 1 - P\Big\{ \bar{Y}_i \le \bar{Y}_0 + cS\sqrt{\frac{1}{n_0}+\frac{1}{n}}, i = 1,\ldots,k\Big\} \qquad (6.1)\\
&= 1 - P\Big\{ Z_i + \sqrt{n}\delta_i \le Z_0\sqrt{\frac{n}{n_0}} + \frac{cS}{\sigma}\sqrt{1+\frac{n}{n_0}}, i = 1,\ldots,k\Big\} \\
&= 1 - \int_0^\infty \int_{-\infty}^\infty \prod_{i=1}^k \Phi\Big(z\sqrt{\frac{n}{n_0}} + uc\sqrt{1+\frac{n}{n_0}} - \sqrt{n}\delta_i\Big)\phi(z)q_\nu(u)dzdu.
\end{aligned}
$$

where $\delta_i = (\mu_i - \mu_0)/\sigma, i = 1,\ldots,k$. The $\Big\{Z_i = \sqrt{n_i}(\bar{Y}_i - \mu_i)/\sigma\Big\}$ $(i = 0, 1,\ldots,k)$ are independent standard normal random variables with p.d.f $\phi(\cdot)$ and c.d.f $\Phi(\cdot)$, the random variable $U = S/\sigma$ is distributed as $\chi_\nu/\sqrt{\nu}$ with the probability density function $q_\nu(\cdot)$ and degrees of freedom $\nu = n_0 + kn - (k+1) > 0$. Under additional conditions on the means, the follwoing result can be stated.

Theorem 6.2.1 When $\mu_i \ge \mu_0$ $(i = 1,\ldots,k)$, the least favorable configuration for maximum Type I error occurs at $\mu^* = (\mu_0, \mu_0 + \sigma\delta, \ldots, \mu_0 + \sigma\delta)$.

Proof: From H_0^I and the prior knowledge $\mu_0 \le \mu_i, i = 1,\ldots,k$. $0 \le \delta_i \le \delta$. Notice that $\Phi(\cdot)$ is increasing. On the condition that $S = s$, the minimum of $\prod_{i=1}^k \Phi\Big(z\sqrt{\frac{n}{n_0}} + uc\sqrt{1+\frac{n}{n_0}} - \sqrt{n}\delta_i\Big)$ is attained at $\delta_i = \delta$: in other words, the maximum Type I error probability occurs at $\mu^* = (\mu_0, \mu_0 + \sigma\delta, \ldots, \mu_0 + \sigma\delta)$. Since μ^* is independent of the value u assumed by the random variable U, μ^* is the least favorable configuration for $P_{(\mu,\sigma)}(c)$ over $(\mu,\sigma) \in H_0^I$. This completes the proof.

Let $P_\delta(c) = \sup_{(\mu,\sigma)\in H_0^I} P_{(\mu,\sigma)}(c)$, then we have

$$
P_\delta(c) = 1 - \int_0^\infty \int_{-\infty}^\infty \Phi^k\Big(z\sqrt{\frac{n}{n_0}} + uc\sqrt{1+\frac{n}{n_0}} - \sqrt{n}\delta\Big)\phi(z)q_\nu(u)dzdu. \qquad (6.2)
$$

Note that in the special case where $\delta = 0$, (6.2) reduces to

$$\begin{aligned} P_0(c) &= 1 - \int_0^\infty \int_{-\infty}^\infty \Phi^k\Big(z\sqrt{\frac{n}{n_0}} + uc\sqrt{1+\frac{n}{n_0}}\Big)\phi(z)q_\nu(u)dzdu \qquad (6.3) \\ &= P\Big\{D > c | H_0^I : \mu_0 = \mu_1 = \ldots = \mu_k\Big\}, \end{aligned}$$

where c is the corresponding critical value of Dunnett's procedure in the classical null hypothesis of homogeneity of the k treatment means and the control mean .

We now want to obtain the critical value c, where c is the solution to the equation

$$P_\delta(c) = \alpha. \qquad (6.4)$$

Since the distribution of the test statistic D at the LFC involves double integrals, Gaussian quadrature and Newton-Raphson's iteration method are employed to obtain the percentage point c of the test statistic D in (6.2). Moreover, since the double integrals involves infinite integral limits, it is necessary to truncate the infinite limits to a finite one for ease of calculation and avoiding underflow and/or overflow. In this chapter, we truncate the normal variable z at $-\bar{\tau}$ and $\bar{\tau}$ with trucation error $\leq 2 \times 10^{-9}$ (because $\Phi(\bar{\tau}) > 1 - 10^{-9}$) and the $\chi_\nu/\sqrt{\nu}$ variable u at a point b such that $\int_b^\infty q_\nu(u)du \leq 10^{-9}$ so that the total truncation error is less than $2(k+1) \times 10^{-9} + (k+1) \times 10^{-9}$. In the numerical integration we use a 64-point Gaussian-Legendre quadrature over each of 4 by 4 subintegrals to evaluate the double integrals. At special case with $\delta = 0$, these percentage points agree with existing tables to at least three decimal places. (see Bechhofer and Dunnett (1988)).

The Newton-Raphson iteration is used to solve Equation (6.4) for c, for the nth iteration

$$c_n = c_{n-1} - \frac{P_\delta(c_{n-1}) - \alpha}{P'_\delta(c_{n-1})},$$

where $P'_\delta(c)$ is the partial derivative of $P_\delta(c)$ with respect to c and is given by

$$P'_\delta(c) = -k\sqrt{1+\frac{n}{n_0}}\int_0^\infty\int_{-\infty}^\infty \Phi^{k-1}\Big(z\sqrt{\frac{n}{n_0}}+uc\sqrt{1+\frac{n}{n_0}}-\sqrt{n}\delta\Big)$$
$$.\phi\Big(z\sqrt{\frac{n}{n_0}}+uc\sqrt{1+\frac{n}{n_0}}-\sqrt{n}\delta\Big)\phi(z)uq_\nu(u)dzdu \qquad (6.5)$$

which is negative since u, $\Phi(\cdot)$, $\phi(\cdot)$ and $q_\nu(\cdot)$ are nonnegative. The solution is unique since $P_\delta(c)$ is monotonically decreasing in c.

In (6.2), the double integral is partitioned over 4 by 4 subrectangles where the z variable is partitioned into $(-7, -0.6745, 0, 0.6745, 7)$ and the U variable into $(a_1, a_2, a_3, a_4, a_5)$, where a_1 is the 10^{-9}th percentile, a_2 is the 25th percentile, a_3 is the median, a_4 is the 75th percentile, and a_5 is the $1. - 10^{-9}$th percentile of U. A 64-point Gaussian-Legendre quadrature is used to evaluate over each of 16 subintegrals and then the results are summed to obtain the overall probability integral in (6.2). The same process is used to obatin the value of the probability integral in (6.5).

Table 6.1 gives the percentage points with equal sample size case, i.e., $n_0 = n$, for $k = 2(1)10$, $n = 2(2)10, 15, 20, 30(10)80$, $\alpha = .05$, and $\delta = .10$ and $1/3$ ($\delta = .10$ and $1/3$ are presented here. Fortran 77 source code is available which can be used to make more tables). Table 6.2 is similar to Table 6.1 except that the allocation $\frac{n_0}{n} = \sqrt{k}$ is considered, this type of allocation is called square root allocation rule which was shown to be nearly optimal in Dunnett (1955) and other papers.

When the common variance is known or when the degrees of freedom ν are very large, S converges to σ. Thus, H_0^I is rejected if $D > c$, where c is the solution to

$$1-\alpha = \int_{-\infty}^\infty \Phi^k\Big(z\sqrt{\frac{n}{n_0}}+c\sqrt{1+\frac{n}{n_0}}-\sqrt{n}\delta\Big)\phi(z)dz. \qquad (6.6)$$

The percentage points are given in Table 6.3 and Table 6.4 for $\alpha = .05$, $\sqrt{n}\delta = 1(1)11$, and $k = 2(1)10$ for equal sample size case $n_0 = n$ and $n_0/n = \sqrt{k}$, respectively.

Example 1. Let us assume that a psychologist wishes to test differences in I.Q. scores between a control group and 4 treatment groups with equal sample size $n = 15$. According to her previous research, she believes that the mean I.Q. scores for each treatment group is at least as high as that of the control group. Because of lack of precision of measurement, she regards difference of 1/3 standard unit to be irrelevant. Assume that $D = \frac{\max_{1\le i\le 4} \bar{Y}_i - \bar{Y}_0}{S/\sqrt{n}} = 2.83$. The critical value at $\alpha = 0.05$ for the point null hypothesis $\mu_0 = \mu_1 = \mu_2 = \mu_3 = \mu_4$ is 2.20 from Table E.2 in Hsu (1996). If she uses the point null hypothesis, she has to reject the point null hypothesis. From Table 6.1, the critical value for the interval null hypothesis is 3.053, she does not reject the interval null hypothesis and concludes that the differences bewteen the control and treatments are no more than 1/3 standard unit.

6.3 Confidence Lower Bound

An experimenter may not be able to specify the size of δ in the null hypothesis. In such situations, the tables can be used to construct a confidence lower bound for the standardized range $d(\mu) = (\mu_{best} - \mu_0)/\sigma$ instead. The confidence lower bound is obtained as follows. For simplicity, we assume that σ is known and $n_0 = n$. Denote

$$L(c, \delta; k, n) = \inf_{(\mu_{best}-\mu_0)/\sigma=\delta} P_\mu\Big\{D \le c\Big\} = 1 - \sup_{(\mu_{best}-\mu_0)/\sigma=\delta} P_\mu\Big\{D \ge c\Big\},$$

which increases from 0 to 1 as c increases from 0 to ∞. Given $\alpha > 0$, we can find $c = c_{k,n}(\delta)$ such that $L(c, \delta; k, n) = 1 - \alpha$. It follows that

$$P_\mu\Big\{D \leq c_{k,n}(\delta)\Big\} \geq 1 - \alpha.$$

Using (6.2) and the monotonicity of $\Phi(\cdot)$, one can easily show that $c_{k,n}(\delta)$ increases strictly from $c_{k,n}(0)$ to ∞ as δ increases from 0 to ∞ for a given α, n and k,

$$P_\mu\Big\{c_{k,n}^{-1}(D) \leq \delta\Big\} \geq 1 - \alpha$$

for all μ with $(\mu_{best} - \mu_0)/\sigma = \delta$. Where $c_{k,n}^{-1}(y)$ is the inverse function of $c_{k,n}$ and is defined to be zero when $y < q$ and q is the $(1 - \alpha)$ percentage point of the test statistic D of $k + 1$ independently and identically distributed random variables $N(\mu, \sigma^2)$ when $\delta = 0$. The $(1 - \alpha)$ percentage point lower confidence bound for $d(\mu)$ is given by $c_{k,n}^{-1}(D)$.

As an illustration of the use of Table 6.3, suppose we want to compute a 95% confidence lower bound for δ. Let k= 5 and assume that $D = \frac{\max_{1 \leq i \leq k} \bar{Y}_i - \bar{Y}_0}{\sigma/\sqrt{n}} = 8.244$. From Table 6.3, $\alpha = .05$, k=5, $d(\mu) = 8.0/\sqrt{n}$ for $c = 7.891$ and $9.0/\sqrt{n}$ for $c = 8.598$. By a numerical algorithm, the 95% confidence lower bound for δ would be $8.5/\sqrt{n}$. When σ is unknown and $n_0/n = \sqrt{k}$, the confidence lower bound for $d(\mu)$ can be similarly obtained from Table 6.1, Table 6.2 and Table 6.4.

6.4 Discussion

As a preliminary step in the problem of comparing k treatments with a control, it is often desirable to test the null hypothesis $H_0 : \mu_0 = \mu_1 = \ldots = \mu_k$ versus the

alternative hypothesis $H_1 : \mu_0 \leq \mu_i$ $(i = 1, \ldots, k)$ with at least one strict inequality. Testing H_0 versus H_1 can be done by the Dunnett's procedure or the likelihood ratio test statistic $\bar{\chi}^2$ or $\bar{E}_{01}^2$ which has been introduced in Section 2.5 of Chapter 2. Once H_0 is rejected, one may then consider to test H_0^I versus H_1^I to measure the dispersion among treatment means and the control mean. There is a large body of work on statistical methods for assessing equivalence of several treatments with a control. When the treatments are too expensive for most of the population, which is often the case in many developing countries, the equivalence under the simple tree prior knowledge is advisable to study. When $k = 2$, the critical regions of the likelihood ratio test and Dunnett's test for H_0 versus H_1 and the new test proposed in Section 6.2 for H_0^I versus H_1^I are presented in Figure 6.1.

NOTE TO USERS

Page(s) not included in the original manuscript are unavailable from the author or university. The manuscript was microfilmed as received.

104

This reproduction is the best copy available.

UMI

Table 6.2: Percentage Points with $n_0/n = \sqrt{k}, \alpha = .05$.

n/k	2	3	4	5	6	7	8	9	10
					$\delta = 0.1$				
2	2.869	2.843	2.856	2.877	2.899	2.921	2.942	2.961	2.980
4	2.337	2.460	2.548	2.616	2.672	2.719	2.759	2.795	2.827
6	2.269	2.413	2.513	2.590	2.652	2.704	2.748	2.787	2.822
8	2.256	2.408	2.513	2.594	2.658	2.712	2.759	2.799	2.835
10	2.258	2.414	2.523	2.606	2.672	2.728	2.775	2.817	2.854
15	2.283	2.446	2.559	2.645	2.714	2.772	2.821	2.864	2.902
20	2.314	2.482	2.598	2.685	2.756	2.815	2.865	2.908	2.947
30	2.378	2.550	2.669	2.760	2.832	2.892	2.944	2.988	3.028
40	2.436	2.612	2.733	2.825	2.899	2.960	3.013	3.058	3.099
50	2.489	2.668	2.791	2.885	2.959	3.022	3.075	3.121	3.162
60	2.538	2.719	2.844	2.939	3.014	3.077	3.131	3.178	3.219
70	2.583	2.767	2.893	2.989	3.066	3.129	3.183	3.231	3.272
80	2.626	2.812	2.939	3.036	3.113	3.177	3.232	3.280	3.322
					$\delta = 1/3$				
2	3.329	3.256	3.247	3.255	3.269	3.285	3.302	3.318	3.334
4	2.780	2.900	2.987	3.056	3.111	3.159	3.200	3.236	3.268
6	2.770	2.919	3.024	3.104	3.169	3.223	3.269	3.310	3.346
8	2.813	2.976	3.089	3.175	3.244	3.302	3.352	3.395	3.433
10	2.869	3.040	3.158	3.248	3.321	3.381	3.432	3.477	3.517
15	3.012	3.197	3.324	3.420	3.497	3.561	3.616	3.664	3.706
20	3.146	3.340	3.473	3.574	3.654	3.721	3.778	3.827	3.871
30	3.384	3.591	3.733	3.840	3.925	3.996	4.056	4.108	4.154
40	3.590	3.807	3.956	4.068	4.157	4.231	4.294	4.348	4.396
50	3.774	4.000	4.155	4.271	4.363	4.439	4.504	4.560	4.610
60	3.942	4.176	4.335	4.455	4.550	4.629	4.695	4.753	4.804
70	4.098	4.338	4.502	4.625	4.723	4.803	4.871	4.931	4.983
80	4.243	4.489	4.657	4.783	4.883	4.966	5.036	5.096	5.149

Table 6.3: Percentage Points for $n_0 = n$ with Variance Known.

$\sqrt{n}\delta$	k								
	2	3	4	5	6	7	8	9	10
				$\alpha = .01$					
1.0	3.265	3.392	3.479	3.544	3.596	3.640	3.677	3.709	3.738
2.0	3.972	4.099	4.186	4.251	4.304	4.347	4.384	4.416	4.445
3.0	4.679	4.806	4.893	4.958	5.011	5.054	5.091	5.123	5.152
4.0	5.386	5.513	5.600	5.665	5.718	5.761	5.798	5.831	5.859
5.0	6.093	6.220	6.307	6.373	6.425	6.468	6.505	6.538	6.566
6.0	6.800	6.928	7.014	7.080	7.132	7.175	7.212	7.245	7.273
7.0	7.508	7.635	7.721	7.787	7.839	7.882	7.920	7.952	7.980
8.0	8.215	8.342	8.428	8.494	8.546	8.590	8.627	8.659	8.687
9.0	8.922	9.049	9.136	9.201	9.253	9.297	9.334	9.366	9.395
10.0	9.629	9.756	9.843	9.908	9.960	10.004	10.041	10.073	10.102
11.0	10.336	10.463	10.550	10.615	10.667	10.711	10.748	10.780	10.809
				$\alpha = .05$					
1.0	2.623	2.769	2.867	2.941	2.999	3.048	3.089	3.124	3.156
2.0	3.331	3.476	3.575	3.648	3.706	3.755	3.796	3.831	3.863
3.0	4.038	4.183	4.282	4.355	4.414	4.462	4.503	4.538	4.570
4.0	4.745	4.891	4.989	5.062	5.121	5.169	5.210	5.245	5.277
5.0	5.452	5.598	5.696	5.769	5.828	5.876	5.917	5.953	5.984
6.0	6.159	6.305	6.403	6.476	6.535	6.583	6.624	6.660	6.691
7.0	6.866	7.012	7.110	7.184	7.242	7.290	7.331	7.367	7.398
8.0	7.573	7.719	7.817	7.891	7.949	7.997	8.038	8.074	8.105
9.0	8.280	8.426	8.524	8.598	8.656	8.704	8.745	8.781	8.812
10.0	8.987	9.133	9.231	9.305	9.363	9.412	9.453	9.488	9.519
11.0	9.695	9.840	9.939	10.012	10.070	10.119	10.160	10.195	10.227
				$\alpha = .10$					
1.0	2.284	2.441	2.545	2.623	2.685	2.736	2.779	2.816	2.849
2.0	2.991	3.148	3.252	3.330	3.392	3.443	3.486	3.523	3.556
3.0	3.698	3.855	3.960	4.038	4.099	4.150	4.193	4.231	4.263
4.0	4.405	4.562	4.667	4.745	4.806	4.857	4.900	4.938	4.971
5.0	5.113	5.269	5.374	5.452	5.513	5.564	5.607	5.645	6.566
6.0	5.820	5.976	6.081	6.159	6.221	6.271	6.315	6.352	6.385
7.0	6.527	6.683	6.788	6.866	6.928	6.979	7.022	7.059	7.092
8.0	7.234	7.390	7.495	7.573	7.635	7.686	7.729	7.766	7.799
9.0	7.941	8.097	8.202	8.280	8.342	8.393	8.436	8.473	8.506
10.0	8.648	8.805	8.909	8.987	9.049	9.100	9.143	9.180	9.213
11.0	9.355	9.512	9.616	9.694	9.756	9.807	9.850	9.887	9.920

Table 6.4: Percentage Points for $n_0/n = \sqrt{k}$ with Variance Known.

$\sqrt{n}\delta$	k								
	2	3	4	5	6	7	8	9	10
				$\alpha = .01$					
1.0	3.329	3.495	3.609	3.695	3.764	3.821	3.870	3.912	3.950
2.0	4.095	4.291	4.425	4.526	4.607	4.673	4.729	4.778	4.821
3.0	4.860	5.087	5.242	5.357	5.449	5.525	5.589	5.644	5.693
4.0	5.625	5.884	6.058	6.189	6.292	6.377	6.448	6.510	6.564
5.0	6.391	6.680	6.875	7.020	7.135	7.229	7.308	7.376	7.436
6.0	7.156	7.476	7.691	7.851	7.977	8.080	8.167	8.242	8.308
7.0	7.921	8.272	8.508	8.683	8.820	8.932	9.027	9.108	9.179
8.0	8.687	9.069	9.324	9.514	9.663	9.784	9.886	9.974	10.051
9.0	9.452	9.865	10.141	10.345	10.505	10.636	10.746	10.840	10.923
10.0	10.218	10.661	10.957	11.176	11.348	11.488	11.606	11.706	11.794
11.0	10.983	11.457	11.774	12.008	12.191	12.340	12.465	12.572	12.666
				$\alpha = .05$					
1.0	2.693	2.884	3.015	3.114	3.194	3.259	3.315	3.364	3.407
2.0	3.458	3.680	3.832	3.946	4.036	4.111	4.175	4.230	4.279
3.0	4.223	4.476	4.648	4.777	4.879	4.963	5.035	5.096	5.151
4.0	4.989	5.272	5.465	5.608	5.722	5.815	5.894	5.962	6.022
5.0	5.754	6.069	6.281	6.439	6.564	6.667	6.754	6.828	6.894
6.0	6.520	6.865	7.098	7.271	7.407	7.519	7.613	7.694	7.765
7.0	7.285	7.661	7.914	8.102	8.250	8.371	8.473	8.560	8.637
8.0	8.050	8.457	8.731	8.933	9.092	9.223	9.332	9.426	9.509
9.0	8.816	9.254	9.547	9.764	9.935	10.074	10.192	10.292	10.380
10.0	9.581	10.050	10.364	10.596	10.778	10.926	11.051	11.158	11.252
11.0	10.346	10.846	11.180	11.427	11.620	11.778	11.911	12.024	12.124
				$\alpha = .10$					
1.0	2.357	2.563	2.705	2.812	2.898	2.969	3.029	3.082	3.129
2.0	3.122	3.359	3.522	3.643	3.741	3.821	3.889	3.948	4.000
3.0	3.888	4.156	4.338	4.475	4.583	4.673	4.749	4.814	4.872
4.0	4.653	4.952	5.155	5.306	5.426	5.525	5.608	5.680	5.743
5.0	5.418	5.748	5.971	6.137	6.269	6.376	6.468	6.546	6.615
6.0	6.184	6.544	6.788	6.968	7.111	7.228	7.327	7.412	7.487
7.0	6.949	7.341	7.604	7.800	7.954	8.080	8.187	8.278	8.358
8.0	7.714	8.137	8.421	8.631	8.797	8.932	9.046	9.144	9.230
9.0	8.480	8.933	9.237	9.462	9.639	9.784	9.906	10.010	10.102
10.0	9.245	9.729	10.054	10.294	10.482	10.636	10.765	10.876	10.973
11.0	10.011	10.526	10.870	11.125	11.325	11.488	11.625	11.742	11.845

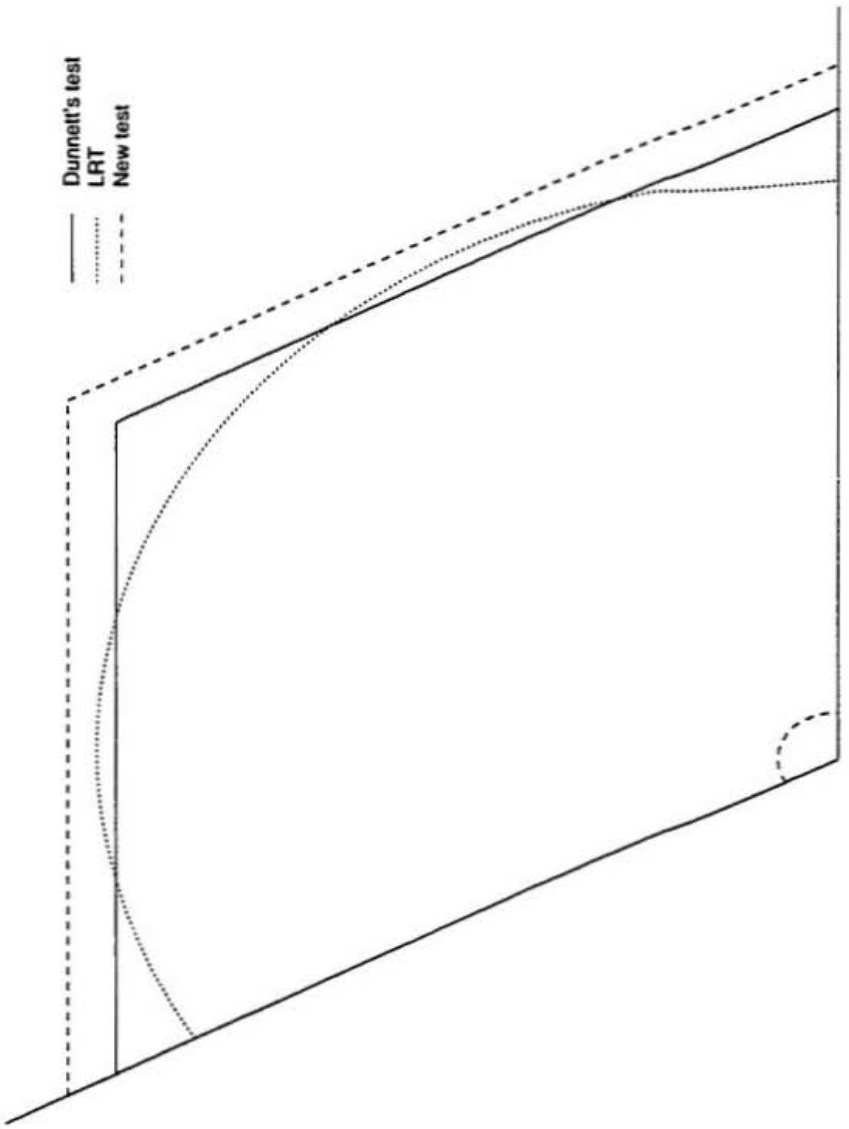

Figure 6.1: Critical Regions for k=2

Chapter 7

Identifying the Minimum Effective Dose

7.1 Introduction

In toxicological and biopharmaceutical studies to investigate the effect of a compound several increasing dose levels are usually compared with a control. The control may be a neutral control (a placebo) or an active control (a standard drug known to be effective). Therefore, a dose-response experiment is often conducted in a one-way layout in which the doses of the compound under consideration are allocated to seperate groups of subjects. There are different concerns in these studies. In toxicological studies, the main concern is the safety of the toxin under consideration and the goal is to estimate the highest dose that shows no significant difference from the control. This highest dose is generally called the no statistical signifcance for trend (NOSTASOT, see Tukey, Ciminera, and Heyse (1985)) or no observed adverse event level dose (NOAEL, see Ryan (1992)). In biopharmaceutical studies, however, the primary goal is to assess whether there is indeed a dose-response effect which means that at least one treatment mean is greater than that of the control,

posed by Tukey, Ciminera, and Heyse (1985) for monotone dose-response effect and modified it for dose-response effect with a reversal at higher doses. Tamhane, Hochberg, and Dunnett (1996) proposed a number of stepwise testing procedures for identifying MED and studied familywise error rate (FWE), bias in identifying MED, and power performance for step-up and step-down multiple comparison procedures using Monte Carlo simulation. They found that step-down procedures generally dominated step-up procedures and recommended some step-down procedures over others. Dunnett and Tamhane (1998) used the step-down tests of Bartholomew and Hayter's procedures for dose finding. More recently, Hsu and Berger (1999) considered the stepwise confidnece sets method, called the DR method, which is based on pairwise t tests without monotone dose-response assumption. They partitioned the parameter space with the idea that exactly one member of the partition contains the true parameter.

We describe the necessory notation first. A typical dose-response study has a (placebo or active) control group indexed as 0 and k treatment groups indexed as $1, \ldots, k$ with increasing dose levels of a compound, with n_i subjects randomly assigned to group $i, i = 0, \cdots, k$. For patient j at dose level i, let Y_{ij} be the dose response. We assume that all observations Y_{ij} are mutually independent with $Y_{ij} \sim N(\mu_i, \sigma^2), i = 0, \ldots, k$ and $j = 1, 2, \ldots, n_i$. The statistic $S^2 = \sum_{i=0}^{k} \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2/\nu$ is used as an estimator for σ^2, and it is independent of the sample means $\bar{\mathbf{Y}} = (\bar{Y}_0, \ldots, \bar{Y}_k)$, where $\nu S^2/\sigma^2 \sim \chi^2_\nu$ and $\nu = \sum_{i=0}^{k} n_i - (k+1) > 0$. Usually the dose-response curve is expected to be continuous. Accordingly, MED should be defined as the minimum dose such that the mean response at that dose is clinically significantly

better than the mean response of the neutral controls; that is

$$MED = \min\{i : \mu_i > \mu_0 + \delta\}, \tag{7.1}$$

where δ defines a clinically significant difference preassigned by an experimenter. Suppose that the control group is an active control group receiving a drug which is known to be effective, then the MED can be defined by (7.1) with δ either positive or 0.

In drug studies, increasing dose levels are frequently expected to produce stronger or at least equal treatment effects. Sometimes the dose response increases monotonically at high dose levels, but, at low levels there might be a negative response, as evidenced in the example in Section 7.4. In that data, the dose response effect at dose 1 dips compared with the effect at placebo. In general, it is possible that $\mu_0 \geq \mu_i$ for some values of i, particularly in the active control setting. In drug development, dose response studies generally are planned to have many active doses and perhaps some potentially ineffective doses to help to characterize the MED. Therefore we may assume that there exists a $q \geq 1$ such that

$$\mu_q \leq \mu_{q+1} \leq \cdots \leq \mu_k, \tag{7.2}$$

where the value of q is preassigned by an experimenter based on past experience. Bauer (1997) pointed out that only the classical many-one pairwise comparison of different doses of the drug with a control will in general guarantee strong family-wise error control without the assumption of monotonicity among means. However, pairwise contrasts do not utilize any prior knowledge about the shape of the dose response function, and hence are not very efficient. The motivation for Williams'

approach (1971, 1972) is presumably to take advantage of the power of isotonic regression when dose responses are monotonically ordered. although the DR method does not explicitly depend on this assumption. It does actually take advantage of this assumption when it is true to much the same extent that Williams's test does (see page 473 in Hsu and Berger (1999)).

Whereas dose response patterns subject to downturn at higher dose levels could occur in practice, this chapter has no intention of discussing this case. One can consult related references like Simpson and Margolin (1986) and Rom, Costello, and Connell (1994) for this case.

The test for overall drug effect on mean response can be assessed with the null hypothesis of $H_0 : \mu_k - \mu_0 \leq \delta$ versus the alternative hypothesis of $H_a : \mu_k - \mu_0 > \delta$. In this chapter, we propose a multiple contrast test which retains the power of the likelihood ratio test. Then, we use it to derive a stepwise method to identify the minimum effective dose under the partial monotone dose-response assumption in (7.2). In Section 7.2, we introduce the multiple contrast test $T_{q,k}$. In Section 7.3, we present the stepwise method. In Section 7.4, we illustrate the method on a dose-response data. In Section 7.5, we conduct some power comparisons. Section 7.6 contains our final conclusions.

7.2 A Multiple Contrast Test Statistic $T_{q,k}$ for Testing Dose-response

The contrast test is very popular in dose response detection. As pointed out by Tukey, Ciminera and Heyse (1985), it offers a semi-nonparametric procedure in

which the treatment effects are left "fully saturated"; that is, they are not restricted by any particular model, and at the same time one can model other design factors to increase the efficiency of the analysis. The contrast test will detect certain expected dose response features without forcing those expected features into the analysis model. Furthermore, analysts can easily interpret contrasts and can clearly present them to clients. Dunnett's procedure (1955), Williams' test (1971, 1972), the regression tests of Tukey, Ciminera and Heyse (1985), Ruberg's basin contrasts (1989), stepwise testing procedures in Tamhane, Hochberg and Dunnett (1996) and Hsu and Berger (1999) are all contrast procedures to study dose-response.

However, the aforementioned contrast procedures do not fully employ the knowledge of the dose ordering and the expected dose shape, and the corresponding powers may be quite low in some directions. Robertson, Wright and Dykstra (1988, page 189) pointed out that the likelihood ratio test statistic may be expressed as the maximum of an infinite number of contrast statistics with the contrast coefficient obtained from the data. Based on this idea, the generalized multiple contrast test statistic $T_{q,k}$ considered in this section has an appealing power performance compared to the LRT.

The layout of this section is as follows. In Section 7.2.1, we introduce the multiple contrast test statistic $T_{q,k}$; in Section 7.2.2, we present an algorithm to compute the restricted maximum likelihood estimates; in Section 7.2.3, we show how to calculate the mixing coefficients in the null distribution of the statistic $T_{q,k}$; in Section 7.2.4, we derive the simultaneous confidence lower bound for $\mu_k - \mu_0$ when the dose-response effectiveness is presented; in Section 7.2.5, we compute the exact power of $T_{1,k}$ when $k = 2$ and 3.

7.2.1 A Multiple Contrast Test Statistic $T_{q,k}$

The dose-response effectiveness can be tested through the null hypothesis $H_0 : \mu_k - \mu_0 \leq \delta$ versus the alternative hypothesis $H_a : \mu_k - \mu_0 > \delta$, where δ defines a clinically significant difference. By incorporating the assumption that $\mu_q \leq \mu_{q+1} \leq \ldots \leq \mu_k$, one rejects H_0 in favour of H_a for large values of

$$T_{q,k} = \max_{\mathbf{c} \in C_q} \frac{\sum_{i=q}^{k} n_i c_i \bar{Y}_i - (\bar{Y}_0 + \delta)}{S\sqrt{\sum_{i=q}^{k} n_i c_i^2 + \frac{1}{n_0}}}. \tag{7.3}$$

where

$$\mathbf{C_{q,k}} = \Big\{ \mathbf{c} = (c_0, c_1, \ldots, c_k) : \sum_{i=q}^{k} n_i c_i = 1, c_1 = \ldots = c_{q-1} = 0, c_0 = -\frac{1}{n_0}, 0 \leq c_q \leq \ldots \leq c_k \Big\}.$$

Without loss of generality we assume $q = 1$ and $\delta = 0$. For simplicity we use T_k to denote $T_{1,k}$ and $\mathbf{C}$ to denote $\mathbf{C_{1,k}}$. Let $t_{k,\alpha,\nu}$ be the critical value of T_k, then

$$P_\mu \Big\{ \sum_{i=0}^{k} n_i c_i \mu_i \geq \sum_{i=0}^{k} n_i c_i \bar{Y}_i - t_{k,\alpha,\nu} S (\sum_{i=0}^{k} n_i c_i^2)^{1/2}, \text{ for all } \mathbf{c} \in \mathbf{C} \Big\} = 1 - \alpha. \tag{7.4}$$

Let $\mathcal{L}=\{\mu : \mu_0 = \mu_1 = \cdots = \mu_k\}$ and $\Omega = \{ \boldsymbol{\mu} \in R^{k+1} : \mu_0 \leq \bar{\mu} \leq \mu_1 \leq \cdots \leq \mu_k\}$, where $\bar{\mu} = \frac{\sum_{j=0}^{k} n_j \mu_j}{\sum_{j=0}^{k} n_j}$. Then $\Omega = \mathcal{C} \ominus \mathcal{L}$. Let $\boldsymbol{\mu}^*$ be the MLE of $\boldsymbol{\mu}$ under Ω which will be discussed in the next subsection. Using the definition of isotonic regression, one may show that

$$\begin{aligned} \max_{\mathbf{c} \in \mathbf{C}} \frac{\sum_{i=0}^{k} n_i c_i \bar{Y}_i}{\sqrt{\sum_{i=0}^{k} n_i c_i^2}} &= \max_{\mathbf{c} \in \mathbf{C}} \frac{\sum_{i=0}^{k} n_i c_i \mu_i^*}{\sqrt{\sum_{i=0}^{k} n_i c_i^2}} \\ &= \sqrt{\sum_{i=0}^{k} n_i (\mu_i^* - \bar{\bar{Y}})^2}, \end{aligned}$$

where $\bar{\bar{Y}} = \sum_{i=0}^{k} n_i \bar{Y}_i / \sum_{i=0}^{k} n_i$. From the above equation,

$$T_k^2 = \sum_{i=0}^{k} n_i (\mu_i^* - \bar{\bar{Y}})^2 / S^2. \tag{7.5}$$

When σ^2 is known, $\sum_{i=0}^{k} n_i(\mu_i^* - \bar{\bar{Y}})^2/\sigma^2$ is the likelihood ratio test statistic for testing $H_0' : \mu_0 = \mu_1 = \ldots = \mu_k$ versus $H_a' : \Omega - H_0'$. When σ^2 is unknown, we call $\sum_{i=0}^{k} n_i(\mu_i^* - \bar{\bar{Y}})^2/S^2$ the modified likelihood ratio test statistic for testing H_0' versus H_a'.

By union-intersection principle, the null distribution of T_k under the least favorable configuration $\mathcal{L}$ of H_0 when $n_0 = n_1 = \cdots = n_k$ is given by

$$\sup_{\mu \in H_0} P[T_k \geq t] = \sum_{l=2}^{k-1} P(l, k+1) P\Big[F_{l-1,\nu} \geq \frac{t^2}{l-1}\Big] \tag{7.6}$$

for any $t > 0$, where the mixing coefficient $P(l, k+1)$ is called the level probability of $\boldsymbol{\mu}^*$ which will be discussed in Section 7.2.3. When k is large, the precision of numerical computation for finding $P(l, k+1)$ is questionable, one may obtain the critical values $t_{k,\alpha,\nu}$ by simulation. The simulated critical values based on 1,000,000 iterations are provided in Table 7.1 for $\alpha = .10,\ .05$ and $.01$, $k = 2, 3, 4, 5, 6, 7, 8, 9, 10$, and the degrees of freedom $\nu = 5, 10, 15, 20, 25, 30, 40, 50, 60$, and ∞. For $k = 2, 3, 4$, the simulated critical values $t_{k,\nu,\alpha}$ match those obtained by solving (7.6).

7.2.2 MLE $\boldsymbol{\mu}^*$

For any $\mathbf{x} \in R^{k+1}$ and $\mathbf{y} \in R^{k-1}$, we define an inner product $< \mathbf{x}, \mathbf{y} >= \sum_{i=0}^{k+1} n_i x_i y_i$. Then Ω is a closed, convex cone. So is $\mathbf{C_q}$ (i.e., $\mathbf{C_q}$ is closed in the topology; $\mathbf{x} \in C_q$ and $\mathbf{y} \in \mathbf{C_q}$ imply that $r\mathbf{x} + (1-r)\mathbf{y} \in \mathbf{C_q}$ for all $0 \leq r \leq 1$; $\mathbf{x} \in \mathbf{C_q}$ and $\lambda \geq 0$ imply that $\lambda\mathbf{x} \in \mathbf{C_q}$).

The restricted maximum likelihood estimator $\boldsymbol{\mu}^*$ of $\boldsymbol{\mu}$ subject to Ω minimizes $\sum_{i=0}^{k} n_i(\bar{Y}_i - \mu_i)^2$. The solution $\boldsymbol{\mu}^*$ is a projection of $\bar{\mathbf{Y}}$ onto Ω. According to Theorem 8.2.7 of Robertson, Wright and Dykstra (1988), $\boldsymbol{\mu}^*$ is the projection if

and only if $\boldsymbol{\mu}^* \in \Omega$,

$$< \mathbf{x} - \mu^*, \mu^* >= 0 \tag{7.7}$$

and

$$< \mathbf{x} - \mu^*, \mathbf{y} > \leq 0 \tag{7.8}$$

for any $\mathbf{y} \in \Omega$. The MLE $\boldsymbol{\mu}^*$ can be computed as follows.

For a given data $\bar{\mathbf{Y}} = (\bar{Y}_0, \ldots, \bar{Y}_k)$, $\mu_i^* = \bar{Y}_i, i = 1, \ldots, q-1$. Without loss of generality, one may assume that $q = 1$ and $\bar{\bar{Y}} = 0$.

Step 1 For $(\bar{Y}_1, \ldots, \bar{Y}_k)$ (excluding the control), using the pool adjacent violation algorithm (PAVA) (see Section 2.4.1) to obtain the isotonic regression with respect to $\mu_1 \leq \mu_2 \leq \ldots \leq \mu_k$, the result is denoted by $(\bar{Y}'_1, \ldots, \bar{Y}'_k)$. If $\bar{Y}_0 \geq \bar{Y}'_k$, then $\mu_i^* = \bar{\bar{Y}}, i = 0, 1, \cdots, k$. Otherwise, go to the next step.

Step 2 For $(\bar{Y}_0, \bar{Y}'_1, \ldots, \bar{Y}'_k)$ (including the control), ignoring those nonpositive treatments, the new data is $(\bar{Y}_0, \bar{Y}'_{i_1}, \cdots, \bar{Y}'_k)$, here $0 \leq \bar{Y}'_{i_1} \leq \ldots \leq \bar{Y}'_k$. Computing the average of the updated data, denoted by $\text{av} = (\sum_{i=i_1}^{k} n_i \bar{Y}'_i + n_0 \bar{Y}_0)/(\sum_{i=i_1}^{k} n_i + n_0)$. If $\text{av} = 0$, stop and $\boldsymbol{\mu}^*$ is

$$(\bar{Y}_0, 0, \ldots, 0, \bar{Y}'_{i_1}, \ldots, \bar{Y}'_k).$$

Otherwise, go to Step 3.

Step 3 Subtracting av from each component, then ignoring those nonpositive treatment components, if there are any. The new data will have the following form:

$$(\bar{Y}_0 - \text{av}, \bar{Y}'_{i_2} - \text{av}, \ldots, \bar{Y}'_k - \text{av}).$$

go to Step 2.

In order to illustrate how to use the above algorithm, let us consider the following example. Suppose $\bar{Y}_0 = -5, \bar{Y}_1 = -1, \bar{Y}_2 = 2, \bar{Y}_3 = 0, \bar{Y}_4 = -2, \bar{Y}_5 = 2$ and $\bar{Y}_6 = 11$. then $\bar{\bar{Y}} = 1$, after subtracting 1 from $\bar{\mathbf{Y}}$. we consider $\bar{\mathbf{Y}} = (-6, -2, 1, -1, -3, 1, 10)$.

Step 1 Through PAVA. the isotonic regression of $(-2, 1, -1, -3, 1, 10)$ subject to $\mu_1 \leq \mu_2 \leq \mu_3 \leq \mu_4 \leq \mu_5 \leq \mu_6$ is $(-2, -1, -1, -1, 1, 10)$.

Step 2 Ignoring those nonpositive components in $(-2, -1, -1, -1, 1, 10)$. then the up-dated data is $(-6, 1, 10)$ and its average av $= 5/3 > 0$.

Step 3 Subtracting $5/3$ from $(-6, 1, 10)$, then ignoring those nonpositive treatment components, thus, the new data is $(-8, 8)$. Its average av $= 0$. Stop. Therefore, we have that $\boldsymbol{\mu}^* = 1 + (-8, 0, 0, 0, 0, 0, 8) = (-7, 1, 1, 1, 1, 1, 9)$.

One can use (7.7) and (7.8) to verify that $(-7, 1, 1, 1, 1, 1, 9)$ is the solution.

7.2.3 The Level Probability $P(l, m+1)$

When $q > 1$, we ignore the first $q-1$ treatments and consider the remaining $m = k-q+1$ treatments. For simplicity, we only consider equal sizes $n_0 = n_1 = \ldots = n_m$ in this subsection. Let

$$\mathbf{C} = \Big\{\mathbf{c} = (c_0, c_1, \ldots, c_m) : \sum_{i=1}^{m} n_i c_i = 1, c_0 = -\frac{1}{n_0}, 0 \leq c_1 \leq \ldots \leq c_m\Big\}.$$

The cone $\mathbf{C}$ is generated by the following m generators $\mathbf{g}_i$, where

$$\mathbf{g}_i = (-1, \underbrace{0, \ldots, 0}_{i-1 \text{ terms}}, \underbrace{1/(m-i+1), \ldots, 1/(m-i+1)}_{m-i+1 \text{ terms}}).$$

For these m generators $\mathbf{g_i}$, there exist corresponding m constraints $\mathbf{f_i}$ such that $< \mathbf{f_i}, \mathbf{g_j} > = \sum_{l=0}^{m} f_{il} g_{jl} = 0$, and $< \mathbf{f_i}, \mathbf{g_i} > = \sum_{l=0}^{m} f_{il} g_{il} < 0$, where $1 \leq i \neq j \leq m$. Any point $\mathbf{y} \in \mathbf{R^{m+1}}$ with $\sum_{i=0}^{m} y_i = 0$ is the nonnegative combination of these generators and constraints. For instance, when $m = 2$, $\mathbf{C}$ is generated by $\mathbf{g_1} = (-1, 1/2, 1/2)$ and $\mathbf{g_2} = (-1, 0, 1)$, the corresponding two constraints are $\mathbf{f_1} = (1, -2, 1)$ and $\mathbf{f_2} = (0, 1, -1)$. When $m = 3$, $\mathbf{C}$ is generated by $\mathbf{g_1} = (-1, 1/3, 1/3, 1/3)$, $\mathbf{g_2} = (-1, 0, 1/2, 1/2)$ and $\mathbf{g_3} = (-1, 0, 0, 1)$, while the corresponding three constraints are $\mathbf{f_1} = (1, -3, 1, 1)$, $\mathbf{f_2} = (0, 1, -1, 0)$ and $\mathbf{f_3} = (0, 0, 1, -1)$.

The level probability $P(l, m+1)$ of $\boldsymbol{\mu}^*$ is the probability that $\boldsymbol{\mu}^*$ can be represented by the nonnegative combination of $l-1$ distinct generators g_i and the $m+1-l$ constraints f_j of cone $\mathbf{C}$ under $H_0' : \mu_0 = \mu_1 = \cdots = \mu_m$. The level probability $P(l, m+1)$ satisfies the following two equations

$$\sum_{l=1}^{m+1} P(l, m+1) = 1, \quad \sum_{l:even} P(l, m+1) = \sum_{l:odd} P(l, m+1) = \frac{1}{2}.$$

When $m = 2$, the space $R^{2+1} \cap \{\mathbf{y} : \sum_{i=0}^{2} y_i = 0\}$ can be decomposed into four convex cones, one cone is Ω which involves two generators, two cones involve one generator and one constraint, the remaining one does not involve any generator but two constraints which is the dual of Ω. Note That $\bar{\mathbf{Y}} \in \Omega$ if and only if $\bar{Y}_0 - 2\bar{Y}_1 + \bar{Y}_2 \leq 0$ and $\bar{Y}_1 - \bar{Y}_2 \leq 0$. Let $U_1 = -\bar{Y}_0 + 2\bar{Y}_1 - \bar{Y}_2 \geq 0, U_2 = -\bar{Y}_1 + \bar{Y}_2 \geq 0$, then (U_1, U_2) has a bivariate normal distribution of zero means and correlation coefficient $\rho_{12} = -\sqrt{3}/2$. Therefore, we have the level probability

$$\begin{aligned} P(3,3) &= P\{U_1 > 0, U_2 > 0\} \\ &= 1/4 + \frac{1}{2\pi} \sin^{-1} \rho_{12} \end{aligned}$$

$$= 1/12.$$

where the formulas for the orthant probabilities can be found on page 75 of Robertson, Wright and Dykstra (1988). Likewise for the other three cones, we can obtain the level probabilities for $m = 2$ as follows. $P(1,3) = 5/12$, $P(2,3) = 1/2$, and $P(3,3) = 1/12$. When $m = 3$, the space $R^{3+1} \cap \{\mathbf{y} : \sum_{i=0}^{3} y_i = 0\}$ can be decomposed into eight convex cones. Each cone is generated by three vectors consisting of generators g_i and constraints f_j. Cone Ω is generated by all three generators g_1, g_2, and g_3; three cones involve two generators and one constraint; three cones involve one generator and two constraints; the remaining one does not involve any generators. For example, the event $(l = 3)$ can be decomposed into three disjoint events or three disjoint blocks which are denoted by B_1, B_2, and B_3, respectively. Say B_1 is generated by $\mathbf{g_1} = (-1, 1/3, 1/3, 1/3)$, $\mathbf{g_2} = (-1, 0, 1/2, 1/2)$, and $\mathbf{f_3} = (0, 0, 1, -1)$; B_2 is generated by $\mathbf{g_1} = (-1, 1/3, 1/3, 1/3)$, $\mathbf{g_3} = (-1, 0, 0, 1)$, and $\mathbf{f_2} = (0, 1, -1, 0)$; B_3 is generated by $\mathbf{g_2} = (-1, 0, 1/2, 1/2)$, $\mathbf{g_3} = (-1, 0, 0, 1)$, and $\mathbf{f_1} = (1, -3, 1, 1)$. B_1 can also be equivalently represented by the following three inequalities: $< \bar{\mathbf{Y}}, \mathbf{g_1'} >< \mathbf{0}$, $< \bar{\mathbf{Y}}, \mathbf{g_2'} >< \mathbf{0}$, and $< \bar{\mathbf{Y}}, \mathbf{f_1'} > < 0$, where $\mathbf{g_1'} = (1, -3, 1, 1)$, $\mathbf{g_2'} = (0, 2, -1, -1)$, and $\mathbf{f_1'} = (0, 0, -1, 1)$. It follows that

$$\begin{aligned} P(B_1) &= P\Big\{< \bar{\mathbf{Y}}, \mathbf{g_1'} >< \mathbf{0}, < \bar{\mathbf{Y}}, \mathbf{g_2'} >< \mathbf{0}, < \bar{\mathbf{Y}}, \mathbf{f_1'} > < 0\Big\} \\ &= \frac{1}{2} P\Big\{< \bar{\mathbf{Y}}, \mathbf{g_1'} >< \mathbf{0}, < \bar{\mathbf{Y}}, \mathbf{g_2'} >< \mathbf{0}\Big\} \\ &= \frac{1}{2}\Big(1/4 + \frac{1}{2\pi}\sin^{-1}(-\sqrt{8}/3)\Big) \\ &= 0.027043. \end{aligned}$$

Similarly, $P(B_2) = 0.048979$, and $P(B_3) = 0.041667$. Accordingly, it follows that $P(l = 3) = P(B_1) + P(B_2) + P(B_3) = 0.11769$. Other level probabilities can be

computed similarly. The level probabilities are presented in Table 7.2 for $m = 2, 3, 4, 5$. For $m \geq 4$ no closed-form is available, numerical computation is used.

7.2.4 Confidence Lower Bound for $\mu_k - \mu_0$

In this section, without loss of generality, we assume $q = 1$ and $\delta = 0$. Once the dose-response effectiveness is shown by rejecting H_0, we are interested in the sharpest confidence lower bound for $\mu_k - \mu_0$. The $1 - \alpha$ one-sided simultaneous confidence bound for the contrast $\sum_{i=0}^{k} n_i c_i \mu_i$ with $\mathbf{c} \in \mathbf{C}$ is given by

$$l\Big(\sum_{i=0}^{k} n_i c_i \mu_i\Big) = \sum_{i=0}^{k} n_i c_i \bar{Y}_i - t_{k,\alpha,\nu} S \Big(\sum_{i=0}^{k} n_i c_i^2\Big)^{1/2}. \tag{7.9}$$

Let $\mathcal{K} = \{\mathbf{c} : \mathbf{c} \in \mathbf{C},\ \sum_{i=0}^{k} n_i c_i \mu_i \leq \mu_k - \mu_0, \text{ where } \mu_1 \leq \ldots \leq \mu_k\}$. The simultanoues confidence lower bound for $\mu_k - \mu_0$ is given by

$$L(\mu_k - \mu_0) = \max_{\mathbf{c} \in \mathcal{K}} l\Big(\sum_{i=0}^{k} n_i c_i \mu_i\Big). \tag{7.10}$$

The following lemma gives another description of the set $\mathcal{K}$ and its proof is trivial.

Lemma 7.2.1 Suppose that H_0 is rejected. For $\mu_1 \leq \ldots \leq \mu_k$, $\sum_{i=0}^{k} n_i c_i \mu_i \leq \mu_{best} - \mu_0$ if and only if $\sum_{j=i}^{k} n_j c_j \leq 1, i = 1, \ldots, k$, for all $\mathbf{c} \in \mathbf{C}$.

The following theorem establishes an equivalence relationship between the positiveness of the above optimal lower bound and rejection of H_0 by statistic T_k.

Theorem 7.2.1 Assuming $\mu_q \leq \ldots \leq \mu_k$, $T_{q,k} > t_{k-q+1,\alpha,\nu}$ if and only if $L(\mu_k - \mu_0) > 0$.

Proof. Without loss of generality assuming $q = 1$. If $T_k > t_{k,\alpha,\nu}$, then there exists a $\mathbf{c_1} \in \mathbf{C}$ such that

$$l\Big(\sum_{i=0}^{k} n_i c_{1i}\mu_i\Big) = \sum_{i=0}^{k} n_i c_{1i}\bar{Y}_i - t_{k,\nu,\alpha} S\Big(\sum_{i=0}^{k} n_i c_{1i}^2\Big)^{1/2} > 0.$$

It is trivial that $\sum_{i=0}^{k} n_i c_{1i}\mu_i \leq \mu_k - \mu_0$ under the assumption $\mu_1 \leq \ldots \leq \mu_k$. Therefore, $L(\mu_k - \mu_0) \geq l(\sum_{i=0}^{k} n_i c_{1i}\mu_i) > 0$.

On the other hand, if $L(\mu_k - \mu_0) > 0$, then there exists a $\mathbf{c_2} \in \mathcal{K}$ such that

$$L(\mu_k - \mu_0) = l\Big(\sum_{i=0}^{k} n_i c_{2i}\mu_i\Big) > 0.$$

This leads to

$$T_k = \max_{\mathbf{c}\in\mathbf{C}} \sum_{i=0}^{k} n_i c_i \bar{Y}_i / S\Big(\sum_{i=0}^{k} n_i c_i^2\Big)^{1/2} \geq \sum_{i=0}^{k} n_i c_{2i}\bar{Y}_i / S\Big(\sum_{i=0}^{k} n_i c_{2i}^2\Big)^{1/2} > t_{k,\alpha,\nu}.$$

□

According to the algorithm of computing MLE $\boldsymbol{\mu}^*$ and through the contradiction method, it is easy to prove the following lemma.

Lemma 7.2.2 *Let $\boldsymbol{\mu}^*$ be the MLE of $\boldsymbol{\mu}$ subject to Ω. If $\mu_i^* = \mu_{i+1}^*$ for some $i, 0 \leq i \leq k-1$, then $c_i^o = c_{i+1}^o$. Here $\mathbf{c}^o$ is the optimal solution to the maximization problem*

$$\max\Big\{\sum_{i=0}^{k} n_i c_i \mu_i^* - t_{k,\alpha,\nu} S\Big(\sum_{i=0}^{k} n_i c_i^2\Big)^{1/2}\Big\}, \tag{7.11}$$

subject to $\mathbf{c} \in \mathcal{K}$.

In order to obtain the optimal solution to the maximization problem (7.10), we will solve the maximization problem (7.11) first, then we shall prove that the maximization problem (7.10) is equivalent to the maximization problem (7.11). A necessary and sufficient condition for an optimal solution to (7.11) (or (7.10)) is given below.

Theorem 7.2.2 Suppose $T_k > t_{k,\alpha,\nu}$. Then $\mathbf{c}^o \in \mathcal{K}$ is an optimal solution to (7.11) if and only if there exists a positive p, $1 \le p \le k$ such that $c_0^o = -\frac{1}{n_0}$, $c_1^o = \cdots = c_{p-1}^o = 0$, $c_i^o = N_{pk}^{-1} + b^{-1}(\mu_i^* - \bar{\mu}^*_{pk})$, $i = p, \cdots, k$, where

$$b^2 = (t_{k,\alpha,\nu}^2 S^2 - S_{pk}^2)/(n_0^{-1} + N_{pk}^{-1}).$$

and

$$N_{pk} = \sum_{i=p}^{k} n_i, \quad \bar{\mu}^*_{pk} = \sum_{i=p}^{k} n_i \mu_i^* / N_{pk}, \quad S_{pk}^2 = \sum_{i=p}^{k} n_i (\mu_i^* - \bar{\mu}^*_{pk})^2.$$

Proof. It is trivial that $\sum_{i=0}^{k} n_i c_i \mu_i^* - t_{k,\alpha,\nu} S \left(\sum_{i=0}^{k} n_i c_i^2 \right)^{1/2}$ is a concave function of $c_0, c_1, \ldots, c_k$. Let

$$\begin{aligned} \phi(c, u, v, \lambda) \;=\; & \sum_{i=0}^{k} n_i c_i \mu_i^* - t_{k,\alpha,\nu} S \Big(\sum_{i=0}^{k} n_i c_i^2 \Big)^{1/2} + \sum_{l=1}^{k-1} u_l (c_{l+1} - c_l) \\ & + u_0 c_1 + \sum_{m=0}^{k-1} v_m \Big(1 - \sum_{j=m+1}^{k} n_j c_j \Big) + \lambda \Big(1 - \sum_{i=1}^{k} n_i c_i \Big). \end{aligned}$$

Let $\frac{\partial \phi}{\partial \mathbf{c}^o}$ denote the partial derivatives evaluated at the point $\mathbf{c}^o$, $\mathbf{u}^o = (u_0^o, \cdots, u_{k-1}^o)$, $\mathbf{v}^o = (v_0^o, \cdots, v_{k-1}^o)$, and λ^o. By the Kuhn-Tucker equivalence theorem (Kuhn and Tucker (1951)), $\mathbf{c}^o$ is the optimal solution if and only if

(i) $\frac{\partial \phi}{\partial c_i^o} = n_i \mu_i^* - n_i c_i^o b + u_{i-1} - u_i - \sum_{j=0}^{i-1} n_i v_j - n_i \lambda = 0, i = 1, \cdots, k$, where $b = t_{k,\alpha,\nu} S / (\sum_{i=0}^{k} n_i c_i^{o2})^{1/2}$;

(ii) $c_1^o \ge 0$, $c_{i+1}^o - c_i^o \ge 0, u_i^o (c_{i+1}^o - c_i^o) = 0, i = 1, \cdots, k-1, u_0^o c_1^o = 0, \mathbf{u}^o \ge \mathbf{0}$; $\sum_{r=j}^{k} n_r c_r^o \le 1, j = 1, \ldots, k, v_j (1 - \sum_{r=j}^{k} n_r c_r^o) = 0$, $\mathbf{v}^o \ge \mathbf{0}$ and $\frac{\partial \phi}{\partial \lambda^o} = 0$.

Suppose $\mathbf{c}^o$ is the optimal solution. For convenience, let $u_k = 0$. Without loss of generality, one may assume that there exists a p $(1 \le p \le k)$ such that

$$c_1^o = \ldots = c_{p-1}^o = 0 < c_p^0 \le \ldots \le c_k^o,$$

with $\sum_{i=p}^{k} n_i c_i^o = 1$. Then from (ii), $u_{p-1}^o = 0$. Let $u_j^o = 0, V^* = \sum_{i=1}^{p-1} v_j^o$ (If $p = 1$, $V^* = v_0^o$), and $v_j^o = 0, j = p, \ldots, k$. From (i), for $j = p, \ldots, k$,

$$c_j^o = b^{-1}(\mu_j^* - V^* - \lambda^o). \tag{7.12}$$

Adding these $k - p + 1$ equations in (7.12) and using $\sum_{j=p}^{k} n_j c_j^o = 1$, one has

$$\lambda^o + V^* = \bar{\mu}^*_{pk} + bN_{pk}^{-1}.$$

Substituting $\lambda^o + V^*$ back in (7.12) , then it follows that

$$c_j^o = N_{pk}^{-1} + b^{-1}(\mu_j^* - \bar{\mu}^*_{pk}), j = p, \ldots, k.$$

Accordingly, $b = t_{k,\alpha,\nu} S / (\sum_{i=0}^{k} n_i c_i^{o2})^{1/2}$ can be written as

$$b^2 = \frac{t_{k,\alpha,\nu}^2 S^2 - S_{pk}^2}{n_0^{-1} + N_{pk}^{-1}}.$$

Therefore, the necessory part is shown. The sufficient part is trivial. □

Now we prove that the maximization problem (7.10) is equivalent to the maximization problem (7.11).

Theorem 7.2.3 *The maximum problem in* (7.10) *is equivalent to the maximization problem* (7.11).

Proof. For convenience, let $f(\mathbf{c})$ denote equation (7.10) and let $g(\mathbf{c})$ denote equation (7.11). Let $\mathbf{c}^o$ be the optimal solution to (7.11), then, $g(\mathbf{c}) \leq g(\mathbf{c}^o)$ for any $\mathbf{c} \in \mathcal{K}$. Without loss of generality we assume $\bar{Y}_1 \leq \ldots \leq \bar{Y}_k$. According to the algorithm of finding $\boldsymbol{\mu}^*$ in Section 7.2.2, $\boldsymbol{\mu}^* = (\bar{Y}_0 - a, \bar{\bar{Y}}, \cdots, \bar{\bar{Y}}, \bar{Y}_i - a, \cdots, \bar{Y}_k - a)$, here a is a

constant. Then by Theorem 7.2.2 and Lemma 7.2.2.

$$\sum_{j=0}^{k} n_j c_j^o \mu_j^* = n_0 c_0^o \mu_0^* + \sum_{j=p}^{k} n_j c_j^o \mu_j^* = n_0 c_0^o (\bar{Y}_0 - a) + \sum_{j=p}^{k} n_j c_j^o (\bar{Y}_j - a) = \sum_{j=0}^{k} n_j c_j^o \bar{Y}_j.$$

It follows that $f(\mathbf{c}^o) = g(\mathbf{c}^o) \geq g(\mathbf{c})$. Since for any $\mathbf{c} \in \mathbf{C}$, $\mathbf{c} \in \Omega$, by (7.8). $\sum_{j=0}^{k} n_j c_j^o \bar{Y}_j \leq \sum_{j=0}^{k} n_j c_j^o \mu_j^*$. It follows that $g(\mathbf{c}) \geq f(\mathbf{c})$. This completes the proof. □

The optimal solution $\mathbf{c}^o$ can be obtained iteratively in a few steps by the following algorithm.

(0) Set $i = 0$ and $p_0 = \min_{1 \leq i \leq k}\{i : \mu_i^* > \bar{\bar{Y}}\}$.

(i) Let $\beta_{i+1} = N_{p_i k}(\bar{\mu}^*_{p_i k} - \mu^*_{p_i})$. $t_{k,\alpha_{i+1},\nu} = \left[S^2_{p_i k} + \beta_i^2 (n_0^{-1} + N_{p_i k}^{-1})\right]^{1/2} / S$. If $t_{k,\alpha_{i+1},\nu} \leq t_{k,\alpha,\nu}$, the optimal solution is $\mathbf{c}^o$ with $p = p_i$. Otherwise, go to (ii).

(ii) Set $p_{i+1} = \min_{1 \leq i \leq k}\{i : \mu_i^* > \mu^*_{p_i}\}$. Go to (i).

7.2.5 The Power Function of T_k When $k = 2$ and $k = 3$

For simplicity, we assume $\sigma^2/n = 1$. We consider $k = 2$ first, let $\bar{Y}_i \sim N(\mu_i, 1)$ for $i = 0, 1, 2$ with $\bar{Y}_0, \bar{Y}_1, \bar{Y}_2$ independent. Since the event $(l = 3)$ implies that $\mu_i^* = \bar{Y}_i$ $(i = 0, 1, 2)$, then for any constant $a > 0$, by (7.5), $P\{T_2 > a, l = 3\}$ can be computed as follows. This probability is

$$\begin{aligned} P\{T_2 > a, l = 3\} &= P\Big\{\frac{\sum_{i=0}^{2} n(\mu_i^* - \bar{\bar{Y}})^2}{\sigma^2} > a^2, l = 3\Big\} \\ &= P\Big\{\bar{Y}_2 - \bar{Y}_1 > 0, 2\bar{Y}_1 - \bar{Y}_0 - \bar{Y}_2 > 0, \sum_{i=0}^{2} (\bar{Y}_i - \bar{\bar{Y}})^2 \geq a^2\Big\}. \end{aligned}$$

By making the orthogonal transformation $V_1 = (\bar{Y}_2 - \bar{Y}_1)/\sqrt{2}$ and $V_2 = (\bar{Y}_1 + \bar{Y}_2 - 2\bar{Y}_0)/\sqrt{6}$, through simple computation with $\sum_{i=0}^{2}(\bar{Y}_i - \bar{\bar{Y}})^2 = V_1^2 + V_2^2$, the probability can be written as

$$P\Big\{V_1 > 0, \frac{\sqrt{6}}{2}V_2 - \frac{3\sqrt{2}}{2}V_1 > 0, V_1^2 + V_2^2 \geq a^2\Big\}.$$

Applying the polar transformation $V_1 = R\cos\theta, V_2 = R\sin\theta$, the conditions $V_1 > 0$, and $\frac{\sqrt{6}}{2}V_2 - \frac{3\sqrt{2}}{2}V_1 > 0$ are equivalent to $\cos\theta > 0, \frac{\sqrt{6}}{2}\sin\theta - \frac{3\sqrt{2}}{2}\cos\theta > 0$, that is, $\pi/3 < \theta < \pi/2$. Because V_1 and V_2 are independent normal variables with means $\lambda_1 = \frac{\mu_2 - \mu_1}{\sqrt{2}}$ and $\lambda_2 = \frac{\mu_1 - \mu_2 - 2\mu_0}{\sqrt{6}}$ and unit variances, with $\lambda_1 = \Delta\sin\beta$ and $\lambda_2 = \Delta\cos\beta$, where $\Delta^2 = \sum_{i=0}^{2}(\mu_i - \bar{\mu})^2$ with $\bar{\mu} = (\mu_0 + \mu_1 + \mu_2)/3$, the above probability is

$$\frac{\exp\Big\{-\Delta^2/2\Big\}}{2\pi}\int_{\pi/3+\beta}^{\pi/2+\beta}\int_{a}^{\infty} r\exp\Big\{-\frac{1}{2}r^2 + r\Delta\sin\theta\Big\}drd\theta$$
$$= \frac{\exp\Big\{-\Delta^2/2\Big\}}{2\pi}\int_{\pi/3+\beta}^{\pi/2+\beta} w(\Delta\sin\theta, a)d\theta \qquad (7.13)$$

where $w(x,a) = [x\Phi(x-a) + \phi(x-a)]/\phi(x)$ with $\Phi(x)$ and $\phi(x)$ being the cumulative distribution function and probability density function of a standard normal variable, respectively. Next, $P\Big\{T_2 > a, l = 2\Big\}$ is considered. The event $(l = 2)$ is the union of $(\bar{Y}_1 - \bar{Y}_2 \geq 0, \bar{Y}_1 + \bar{Y}_2 - 2\bar{Y}_0 \geq 0)$ and $(\bar{Y}_0 - 2\bar{Y}_1 + \bar{Y}_2 \geq 0, \bar{Y}_2 - \bar{Y}_0 \geq 0)$. For the case $(\bar{Y}_1 - \bar{Y}_2 \geq 0, \bar{Y}_1 + \bar{Y}_2 - 2\bar{Y}_0 \geq 0)$, $\boldsymbol{\mu}^* = (\bar{Y}_0, (\bar{Y}_1 + \bar{Y}_2)/2, (\bar{Y}_1 + \bar{Y}_2)/2)$. With the same transformation as above, $T_2^2 = V_2^2$, and

$$P\Big\{\bar{Y}_1 - \bar{Y}_2 \geq 0, \bar{Y}_1 + \bar{Y}_2 - 2\bar{Y}_0 \geq 0, T_2 > a\Big\}$$
$$= P\Big\{V_1 \leq 0, V_2 \geq 0, V_2^2 \geq a^2\Big\} = \Phi(-\Delta\sin\beta)\Phi(\Delta\cos\beta - a). \qquad (7.14)$$

Similarly,

$$P\Big\{\bar{Y}_2 - \bar{Y}_0 \geq 0, \bar{Y}_0 + \bar{Y}_2 - 2\bar{Y}_1 \geq 0, T_2 > a\Big\}$$

$$= \Phi\Big(-a + \Delta\cos(\beta - \pi/6)\Big)\Phi(\Delta\sin(\beta - \pi/6)). \qquad (7.15)$$

For the event $(l = 1)$, $T_2 = 0$. In summary, when $k = 2$ and $n/\sigma^2 = 1$, the power function of T_2 is

$$\begin{aligned}\pi(\mu) &= \frac{\exp\{-\Delta^2/2\}}{2\pi}\int_{\pi/3+\beta}^{\pi/2+\beta} \psi(\Delta\sin\theta, a)d\theta \\ &\quad +\Phi(-\Delta\sin\beta)\Phi(\Delta\cos\beta - a) \\ &\quad +\Phi\Big(-a + \Delta\cos(\beta - \pi/6)\Big)\Phi\Big(\Delta\sin(\beta - \pi/6)\Big). \qquad (7.16)\end{aligned}$$

Using the method in Lee (1987), one can show that the power function of T_2 is increasing on $[-\frac{11\pi}{12}, \frac{\pi}{12}]$ and decreasing on $[\frac{\pi}{12}, \frac{13\pi}{12}]$. Furthermore, the minimum power of T_2 is located at the two boundary points $\beta = 0$ and $\beta = \frac{\pi}{6}$.

Now we consider $k = 3$, let $\bar{Y}_i \sim N(\mu_i, 1)$ for $i = 0, 1, 2, 3$ with $\bar{Y}_0, \bar{Y}_1, \bar{Y}_2, \bar{Y}_3$ independent. The determination of power when $k = 3$ presents a number of difficulties which do not arise in an acute form when $k = 2$.

Since the event $(l = 4)$ implies that $\mu_i^* = \bar{Y}_i$ $(i = 0, 1, 2, 3)$, then for any constant $a > 0$, $P\{T_3 > a, l = 4\}$ can be computed as follows.

$$P(T_3 > a, l = 4) = P(\sum_{i=0}^{3}(\mu_i^* - \bar{\bar{Y}})^2 > a^2, l = 4)$$

which is equivalent to

$$P\{\bar{Y}_2 - \bar{Y}_1 > 0, \bar{Y}_3 - \bar{Y}_2 > 0, 3\bar{Y}_1 - \bar{Y}_0 - \bar{Y}_2 - \bar{Y}_3 > 0, \sum_{i=0}^{3}(\bar{Y}_i - \bar{\bar{Y}})^2 \geq a^2\}.$$

Let $V_3 = (\bar{Y}_3 - \bar{Y}_2)/\sqrt{2}$, $V_4 = (\bar{Y}_2 + \bar{Y}_3 - 2\bar{Y}_0)/\sqrt{6}$, and $V_5 = (3\bar{Y}_1 - \bar{Y}_0 - \bar{Y}_2 - \bar{Y}_3)/\sqrt{12}$. Then V_3, V_4, and V_5 are independent normal variates with unit variances and means $\lambda_3 = \frac{(\mu_3 - \mu_2)}{\sqrt{2}}$, $\lambda_4 = \frac{(\mu_2 + \mu_3 - 2\mu_0)}{\sqrt{6}}$, and $\lambda_5 = \frac{(3\mu_1 - \mu_0 - \mu_2 - \mu_3)}{\sqrt{12}}$. Through simple

computation with $\sum_{i=0}^{3}(\bar{Y}_i - \bar{\bar{Y}})^2 = V_3^2 + V_4^2 + V_5^2$, the probability for the event $(l = 4)$ can be written as

$$P\Big\{V_3 > 0, V_5 > 0, -\frac{\sqrt{3}}{2}V_3 + \frac{1}{2}V_4 - \sqrt{2}V_5 > 0, V_3^2 + V_4^2 + V_5^2 \geq a^2\Big\}.$$

Changing to spherical coordinates, $V_3 = R\cos\theta\sin\eta$, $V_4 = R\sin\theta\sin\eta$, and $V_5 = R\cos\eta$ with $0 \leq \theta \leq 2\pi, 0 \leq \eta \leq \pi$, the conditions $V_3 > 0, V_5 > 0$, and $-\frac{\sqrt{3}}{2}V_3 + \frac{1}{2}V_4 - \sqrt{2}V_5 > 0$ are equivalent to $\frac{\pi}{3} < \theta < \frac{\pi}{2}$ and $\arctan\left(\frac{\sqrt{2}}{\sin(\theta - \frac{\pi}{3})}\right) < \eta < \frac{\pi}{2}$. Therefore, the probability for $(l = 4)$ is

$$\frac{\exp\left\{-\Delta^2/2\right\}}{(2\pi)^{3/2}} \int_a^\infty \int_{\pi/3}^{\pi/2} \int_{\arctan\left(\frac{\sqrt{2}}{\sin(\theta - \frac{\pi}{3})}\right)}^{\pi/2} r^2 \sin\eta \exp(-r^2/2 + \Lambda_{41} r) dr d\theta d\eta \tag{7.17}$$

where $\Lambda_{41} = \lambda_3\cos\theta\sin\eta + \lambda_4\sin\theta\sin\eta + \lambda_5\cos\eta$. The determination of $P\Big\{T_3 > a, l = 3\Big\}$ involves the evaluation of three probabilities, since the event $(l = 3)$ is the union of three cases. For example, in the case $(\bar{Y}_2 - \bar{Y}_3 > 0, \bar{Y}_2 + \bar{Y}_3 - 2\bar{Y}_1 > 0, 3\bar{Y}_1 - \bar{Y}_0 - \bar{Y}_2 - \bar{Y}_3 > 0)$, the MLE $\boldsymbol{\mu}^* = \left(\bar{Y}_0, \bar{Y}_1, \frac{\bar{Y}_2 + \bar{Y}_3}{2}, \frac{\bar{Y}_2 + \bar{Y}_3}{2}\right)$, and then

$$T_3^2 = \frac{12V_5^2}{16} + \frac{(3\bar{Y}_0 - \bar{Y}_1 - \bar{Y}_2 - \bar{Y}_3)^2}{16} + \frac{2(\bar{Y}_2 + \bar{Y}_3 - \bar{Y}_0 - \bar{Y}_1)^2}{16}.$$

The probability corresponding to this case is

$$P\Big\{\bar{Y}_3 - \bar{Y}_2 < 0, \bar{Y}_2 + \bar{Y}_3 - 2\bar{Y}_1 > 0, 3\bar{Y}_1 - \bar{Y}_0 - \bar{Y}_2 - \bar{Y}_3 > 0, T_3 > a\Big\}.$$

Since the distribution of $\bar{Y}_3 - \bar{Y}_2$ is independent of the distribution of V_5, $3\bar{Y}_0 - \bar{Y}_1 - \bar{Y}_2 - \bar{Y}_3$, $\bar{Y}_2 + \bar{Y}_3 - \bar{Y}_0 - \bar{Y}_1$, and $\bar{Y}_2 + \bar{Y}_3 - 2\bar{Y}_1$, and hence of T_3, the above probability may be written as

$$P\Big\{\bar{Y}_3 - \bar{Y}_2 < 0\Big)P(\bar{Y}_2 + \bar{Y}_3 - 2\bar{Y}_1 > 0, 3\bar{Y}_1 - \bar{Y}_0 - \bar{Y}_2 - \bar{Y}_3 > 0, T_3 > a\Big\}.$$

The second term is now found by a straightforward extension of the method used for $P\{T_2 > a, l = 3\}$. The resulting expression for $P\{T_3 > a, l = 3\}$ thus becomes

$$\begin{aligned} P\{T_3 > a, l = 3\} &= \Phi(-\lambda_3)\frac{\exp\{-(\lambda_4^2+\lambda_5^2)/2\}}{2\pi}\int_0^{\arctan\left(\frac{\sqrt{2}}{4}\right)} \psi(\lambda_{31}, a)d\theta \\ &+ \Phi(-\lambda_1)\frac{\exp\{-(\lambda_6^2+\lambda_7^2)/2\}}{2\pi}\int_0^{\arctan\left(\frac{1}{\sqrt{2}}\right)} \psi(\lambda_{32}, a)d\theta \\ &+ \Phi(-\lambda_5)\frac{\exp\{-(\lambda_6^2+\lambda_8^2)/2\}}{2\pi}\int_0^{\frac{\pi}{6}} \psi(\lambda_{33}, a)d\theta \end{aligned} \tag{7.18}$$

where $\lambda_6 = \frac{(\mu_3-\mu_0)}{\sqrt{2}}$, $\lambda_7 = \frac{(\mu_1-\mu_2-\mu_0-\mu_3)}{2}$, $\lambda_8 = \frac{(2\mu_2-\mu_0-\mu_3)}{\sqrt{6}}$, $\lambda_{31} = \lambda_4\cos\theta + \lambda_5\sin\theta$, $\lambda_{32} = \lambda_6\cos\theta + \lambda_7\sin\theta$, $\lambda_{33} = \lambda_6\cos\theta + \lambda_8\sin\theta$. The three contributions to $P\{T_3 > a, l = 2\}$ are readily found by similar methods. Thus for the first term we have that

$$P\{\tilde{Y}_1 + \tilde{Y}_2 + \tilde{Y}_3 - 3\tilde{Y}_0 > 0, \tilde{Y}_2 + \tilde{Y}_3 - 2\tilde{Y}_1 < 0, 2\tilde{Y}_3 - \tilde{Y}_1 - \tilde{Y}_2 < 0, T_3 > a\},$$

where $\boldsymbol{\mu}^* = \left(\tilde{Y}_0, \frac{\tilde{Y}_1+\tilde{Y}_2+\tilde{Y}_3}{3}, \frac{\tilde{Y}_1+\tilde{Y}_2-\tilde{Y}_3}{3}, \frac{\tilde{Y}_1+\tilde{Y}_2+\tilde{Y}_3}{3}\right)$ and $T_3^2 = \frac{(3\tilde{Y}_0-\tilde{Y}_1-\tilde{Y}_2-\tilde{Y}_3)^2}{12}$. We therefore have that

$$\begin{aligned} & P\{-2\tilde{Y}_1 + \tilde{Y}_2 + \tilde{Y}_3 < 0, -\tilde{Y}_1 - \tilde{Y}_2 + 2\tilde{Y}_3 < 0, \tilde{Y}_1 + \tilde{Y}_2 + \tilde{Y}_3 - 3\tilde{Y}_0 > \sqrt{12}a\} \\ = \; & P\{-2\tilde{Y}_1 + \tilde{Y}_2 + \tilde{Y}_3 < 0, -\tilde{Y}_1 - \tilde{Y}_2 + 2\tilde{Y}_3 < 0\}P\{\tilde{Y}_1 + \tilde{Y}_2 + \tilde{Y}_3 - 3\tilde{Y}_0 > \sqrt{12}a\} \\ = \; & \Phi(\lambda_9, \lambda_{10}, 1/2)\Phi(-a + \lambda_{11}). \end{aligned}$$

where $\lambda_9 = \frac{-2\mu_1+\mu_2+\mu_3}{\sqrt{6}}$, $\lambda_{10} = \frac{2\mu_3-\mu_1-\mu_2}{\sqrt{6}}$, $\lambda_{11} = \frac{\mu_1+\mu_2+\mu_3-3\mu_0}{\sqrt{12}}$, and

$$\Phi(x, y; \rho) = \frac{1}{2\pi\sqrt{(1-\rho^2)}}\int_x^\infty\int_y^\infty \exp\{-\frac{1}{2(1-\rho^2)}(u^2 - 2\rho uv + v^2)\}dudv.$$

The total probability for $P\{T_3 > a, l = 2\}$ is then found to be

$$\begin{aligned} P\{T_3 > a, l = 2\} &= \Phi(\lambda_9, \lambda_{10}, 1/2)\Phi(-a + \lambda_{11}) + \Phi(-\lambda_5)\Phi(-\lambda_3)\Phi(-a + \lambda_4) \\ &+ \Phi(\lambda_7, \lambda_8, \sqrt{6}/3)\Phi(-a + \lambda_6). \end{aligned} \tag{7.19}$$

The power function is now obtained by adding (7.17), (7.18), and (7.19).

If σ^2 is unknown, one can obtain the power by conditioning on S^2 and applying the results for σ^2 known.

When $k > 3$ the above method to obtain exact value of the powers becomes very complex and no longer practical. One may simply use Monte-Carlo methods to obtain sufficiently precise estimates of power.

7.3 The Proposed Method

We propose the following procedure to find the MED by making use of the prior knowledge of (7.2). We denote $t_{\alpha,\nu}$ as the upper $100(1-\alpha)$ percentile of the t distribution with degrees of freedom ν.

Step 1: Only the treatment means $\bar{Y}_q, \ldots, \bar{Y}_k$ and the control mean $\bar{Y}_0$ will be used to compute T_{k-q+1}. If $T_{k-q+1} > t_{k-q+1,\alpha,\nu}$, then claim $\mu_k > \mu_0 + \delta$ and go to Step 2; else claim that there is no non-zero dose level which is significantly better than the control and

$$\mu_k - \mu_0 > \max_{c \in C_{q,k}} \Big\{ \sum_{i=q}^{k} n_i c_i \mu_i^* - \mu_0^* - t_{k-q+1,\alpha,\nu} S \sqrt{\frac{1}{n_0} + \sum_{i=q}^{k} n_i c_i^2} \Big\},$$

then stop.

Step 2: Treatment means $\bar{Y}_q, \ldots, \bar{Y}_{k-1}$ and the control mean $\bar{Y}_0$ will be used to compute T_{k-q}. If $T_{k-q} > t_{k-q,\alpha,\nu}$, then claim $\mu_{k-1} > \mu_0 + \delta$ and go to Step 3; else claim $\hat{MED} = k$ and

$$\mu_{k-1} - \mu_0 > \max_{c \in C_{q,(k-1)}} \Big\{ \sum_{i=q}^{k-1} n_i c_i \mu_i^* - \mu_0^* - t_{k-q,\alpha,\nu} S \sqrt{\frac{1}{n_0} + \sum_{i=q}^{k-1} n_i c_i^2} \Big\},$$

then stop.

$\vdots$

Step $k-q$: Treatment means $\bar{Y}_q$ and $\bar{Y}_{q+1}$ and the control mean $\bar{Y}_0$ will be used to compute T_2. If $T_2 > t_{2,\alpha,\nu}$, then claim $\mu_{q+1} > \mu_0 + \delta$ and go to Step $k-q+1$: else claim $\hat{MED} = q+2$ and

$$\mu_{q+1} - \mu_0 > \max_{c \in \mathbf{C}_{q,(q-1)}} \Big\{ \sum_{i=q}^{q+1} n_i c_i \mu_i^* - \mu_0^* - t_{2,\alpha,\nu} S \sqrt{\frac{1}{n_0} + \sum_{i=q}^{q+1} n_i c_i^2} \Big\}.$$

then stop.

Step $k-q+1$: If $\bar{Y}_q - (\bar{Y}_0 + \delta) - t_{\alpha,\nu} S\sqrt{1/n_0 + 1/n_q} > 0$, then claim $\mu_q > \mu_0 + \delta$ and go to Step $k-q+2$: else claim that $\hat{MED} = q+1$ and

$$\mu_q - \mu_0 > \bar{Y}_q - \bar{Y}_0 - t_{\alpha,\nu} S\sqrt{1/n_0 + 1/n_q}.$$

then stop.

$\vdots$

Step k : If $\bar{Y}_1 - (\bar{Y}_0 + \delta) - t_{\alpha,\nu} S\sqrt{1/n_0 + 1/n_1} > 0$, then claim $\mu_1 > \mu_0 + \delta$ and go to Step $k+1$; else claim that $\hat{MED} = 2$ and

$$\mu_1 - \mu_0 > \bar{Y}_1 - \bar{Y}_0 - t_{\alpha,\nu} S\sqrt{1/n_0 + 1/n_1},$$

then stop.

Step $k+1$: If $\bar{Y}_1 - (\bar{Y}_0 + \delta) - t_{\alpha,\nu} S\sqrt{1/n_0 + 1/n_1} > 0$, then claim $\mu_1 > \mu_0 + \delta$ with $\hat{MED} = 1$ and $\min_{1 \le i \le k} \mu_i - \mu_0 = \min_{1 \le i \le q} \mu_i - \mu_0 > \min_{i=1,\ldots,q} \Big\{ \bar{Y}_i - \bar{Y}_0 - t_{\alpha,\nu} S\sqrt{1/n_0 + 1/n_i} \Big\}$, then stop.

Let step M $(1 \leq M \leq k+1)$ be the step at which the stepwise method stops. If $M > 1$, then the stepwise method declares doses $k - M + 2, \ldots, k$ to be efficacious. If $M < k + 1$, then the stepwise method fails to declare doses $1, \ldots, k - M + 1$ to be efficacios. If $M = k + 1$, then the stepwise method gives a lower bound on how efficacious every dose is. This lower bound is greater than δ. The DR method is the special case of our method when $q = k$.

Bauer (1997) showed that for balanced design contrasts like the Hermert contrast (Ruberg (1989)) or the reverse Hermert contrast do not control the probability that a noneffective dose will be erroneously identified as the MED if nonmonotonicity at lower doses occurs. In our stepwise method, just as in the DR method, no α-adjustment is needed, but the familywise error rate is controlled due to the multiple contrast statistic T_k.

7.4 A Numerical Example

In order to illustrate the proposed method in Section 7.3, let us consider the data in Table 7.3, taken from Ruberg (1995). This is an experiment of dose response studies done in laboratory animals. There are ten groups with six animals per group, group 1 with dosage level 0 is the control (placebo), the remaining 9 active groups are the treatments. The mean responses are $\bar{Y}_0 = 25.5, \bar{Y}_1 = 23.9, \bar{Y}_2 = 27.7, \bar{Y}_3 = 33.4, \bar{Y}_4 = 40.5, \bar{Y}_5 = 57.9, \bar{Y}_6 = 74.4, \bar{Y}_7 = 73.4, \bar{Y}_8 = 73.5, \bar{Y}_9 = 76.2$. The pooled mean square error is $S^2 = 60.087$ with the degrees of freedom $\nu = 50$. Table 7.4 shows the different MEDs inferred by the three methods, where $MPGN$ is the stepdown fashion of Dunnett's method in Hsu & Berger (1999). Table 7.5 presents the 95%

step-down confidence lower bounds on $\mu_i - \mu_0$ $(i = 1, \ldots, 9)$ by the three methods. These lower bounds can be used to specify the size of δ. From Table 7.4 and Table 7.5, when q is very close to the true MED, our method tends to identify more doses to be effective; even if q is not close to the true MED, our method is still as good as the DR method.

7.5 Power Comparison of Methods for Dose-response Studies

In order to compare the behavior of the proposed method with the DR method and Dunnett's method, a power study is conducted. Throughout this section for simplicity we assume that $\sigma^2/n = 1$ and $\mu_0 = 0$. We first consider the exact power for $k = 2$ and $k = 3$ with the configuration $\mu_1 = \mu_2$ and $\mu_1 = \mu_2 = \mu_3$ respectively in section 7.5.1. Then we conduct simulation studies in section 7.5.2.

7.5.1 Exact Power for $k = 2$ and $k = 3$

Many dose-responses tend to be sigmoidal. They increase slowly over small doses. Therefore, we consider the configuration of $\mu_1 = \mu_2$ and $\mu_1 = \mu_2 = \mu_3$, respectively. As it will be seen in section 7.5.2, the powers for detecting the true MED by Dunnett's method is inferior to our method and the DR method, we only consider our method and the DR method in this case. According to (7.16) with $\beta = 0$ for $k = 2$, the power for $T_{1,2}$ is as follows,

$$\pi(\mu) = \frac{\exp\left\{-\Delta^2/2\right\}}{2\pi} \int_{\pi/3}^{\pi/2} \psi(\Delta \sin\theta, t_{2,\nu,\alpha}) d\theta$$

$$+\Phi(\Delta - t_{2,\nu,\alpha})/2 + \Phi\Big(-t_{2,\nu,\alpha} + \frac{\sqrt{3}\Delta}{2}\Big)\Phi(-\Delta/2). \qquad (7.20)$$

The power of the DR method is $\Phi\Big(-z_\alpha + \frac{\sqrt{3}\Delta}{2}\Big)$, where z_α is the $100(1-\alpha)$ percentile of standard normal distribution. For $k = 3$, the power of the DR method is $\Phi(-z_\alpha + \frac{2\Delta}{\sqrt{6}})$. One can obtain the power of $T_{1,3}$ by substituting $\mu_0 = 0, \mu_1 = \mu_2 = \mu_3 = \frac{2\Delta}{\sqrt{6}}$ into those expressions in section 7.2.5. Figure 7.1 is the exact power for these two methods when $k = 2$ and 3. The power difference between $T_{1,2}$ and DR method has the following analytical expression

$$\frac{\exp\Big\{-\Delta^2/2\Big\}}{2\pi}\int_{\pi/3}^{\pi/2} v(\Delta\sin\theta, t_{2,\nu,\alpha})d\theta + \Phi\Big(\Delta - t_{2,\nu,\alpha}\Big)/2$$
$$+\Phi\Big(-t_{2,\nu,\alpha} + \frac{\sqrt{3}\Delta}{2}\Big)\Phi(\Delta/2) - \Phi\Big(-z_\alpha + \frac{\sqrt{3}\Delta}{2}\Big). \qquad (7.21)$$

Figure 7.2 describes the power differences between $T_{1,k}$ and the DR method for $k = 2$ and 3. From Figure 7.1 and Figure 7.2, it is clear that $T_{1,k}$ is superior to the DR method for $k = 2$ and 3. Next we consider whether this superiority is maintained for $k > 3$ in stepwise fashion.

7.5.2 A Simulation Power Comparison

In this section, we conduct a simulation study comparing the performance of the above three methods: the DR method (denoted by DR in Table 7.6), our method (denoted by $T_{q,k}$) and Dunnett's method (denoted by D). Strong control of the familywise error is guaranted by these three methods. The number of dose levels (including a control) is 7 and the nominal error rate α is fixed at 0.05. A typical sigmoidal dose response curve is formulated by equation $f(x) = \frac{A-D}{1+\Big(\frac{x}{C}\Big)^B} + D$, where

x is the dose level and $f(x)$ is the corresponding dose response. A is the dose response of the control. D is the dose response at the kth dose level. C is equal to the ED_{50}, which is the dose producing a 50% response, and B is the slope at C (see Ruberg (1995)). We consider six configurations based on the four-parameter logistic dose response with $A = 0, D = 1$ and $B = 4, 5, 6, C = 1, 2, 3, 4$, they are at the bottom of Table 7.6. For each configuration, the value of q is considered to be the true MED minus 1 (q' in Table 7.6), or the true MED (q in Table 7.6), or the true MED plus 1 (q'' in Table 7.6). The value of δ is equal to 0.4, 0.5, 1.0 and 1.5. There is replicated 10000 times for each configuration. Table 7.6 reports the simulation results.

The probability of detecting the dose response is the cumulative probabilities for identifing effective dose levels from MED to dose level k, which is denoted by D in Table 7.6. The probability of identifying the true MED gives an estimate of the power of the method, which is denoted by I in Table 7.6.

When q is very close to the true MED, Table 7.6 shows that the new method has the highest probabilities (I or D) among these three methods for different values of MED and δ. For the three different q values, the new method has roughly the same probability of detecting the dose response. When q equals the true MED, $T_{q,k}$ always has larger probabilities (I and D) than that of DR method and Dunnett's method. In this case, The maximum gains of $T_{q,k}$ over the DR method and Dunnett method can reach 6.16% (for I) and and 11.67% (for D) and 14.59% (for I) and and 7.8% (for D) respectively. When MED ≥ 2, Dunnett method has the lowest probability of identifing the true MED for all the cases. This is not surprising since Dunnett method does not utilize the dose ordering and the dose response shape.

The probability of identifying the true MED for $q'' = \text{MED} + 1$ are lower than that for $q = \text{MED}$ as expected, but the probability of detecting the dose response for $q'' = \text{MED} + 1$ is higher than that for $q = \text{MED}$. When $q' = \text{MED} - 1$, the $T_{q',k}$ method still outperforms the other two methods for $\text{MED} < k$ in detecting the dose response, and the $T_{q',k}$ method is also better than Dunnett's method in identifying the true MED.

Based on the results in Table 7.6, when the dose-response curve is logistic and the experimenter believes his or her q value is around the true MED, the $T_{q,k}$ method is recommended.

7.6 Conclusions

We have proposed a multiple contrast test and developed it into a stepwise method for the analysis of dose response under the partially monotonically assumption. The proposed method is more powerful than the DR method and Dunnett method.

Table 7.1: Upper Percentage Points for T_m.

		m=treatment groups								
df	α	2	3	4	5	6	7	8	9	10
5	.10	1.712	1.814	1.871	1.908	1.936	1.952	1.972	1.980	1.987
	.05	2.272	2.396	2.454	2.494	2.534	2.548	2.565	2.577	2.588
	.01	3.701	3.878	3.947	3.999	4.032	4.068	4.100	4.106	4.120
10	.10	1.581	1.670	1.720	1.749	1.772	1.792	1.800	1.814	1.819
	.05	2.027	2.121	2.175	2.202	2.227	2.249	2.258	2.269	2.280
	.01	3.005	3.114	3.172	3.205	3.223	3.256	3.265	3.279	3.292
15	.10	1.539	1.622	1.670	1.703	1.723	1.741	1.752	1.761	1.772
	.05	1.957	2.040	2.088	2.119	2.142	2.159	2.173	2.184	2.193
	.01	2.816	2.902	2.958	2.990	3.010	3.035	3.053	3.055	3.071
20	.10	1.523	1.603	1.649	1.676	1.700	1.718	1.729	1.735	1.744
	.05	1.921	2.005	2.052	2.078	2.100	2.119	2.133	2.138	2.145
	.01	2.731	2.818	2.862	2.902	2.919	2.936	2.943	2.953	2.960
25	.10	1.507	1.587	1.634	1.667	1.688	1.703	1.709	1.723	1.729
	.05	1.897	1.977	2.027	2.057	2.078	2.093	2.102	2.114	2.119
	.01	2.676	2.762	2.804	2.839	2.865	2.872	2.884	2.898	2.909
30	.10	1.500	1.581	1.628	1.658	1.676	1.691	1.700	1.712	1.723
	.05	1.884	1.970	2.012	2.045	2.062	2.078	2.086	2.098	2.110
	.01	2.642	2.733	2.777	2.807	2.825	2.839	2.848	2.858	2.872
40	.10	1.493	1.575	1.616	1.643	1.667	1.685	1.691	1.703	1.706
	.05	1.871	1.952	1.995	2.022	2.045	2.059	2.069	2.081	2.086
	.01	2.610	2.694	2.733	2.759	2.782	2.796	2.805	2.812	2.821
50	.10	1.487	1.568	1.609	1.640	1.658	1.676	1.685	1.697	1.700
	.05	1.860	1.942	1.982	2.012	2.030	2.047	2.059	2.071	2.071
	.01	2.579	2.661	2.702	2.735	2.750	2.767	2.782	2.798	2.793
60	.10	1.483	1.565	1.606	1.637	1.655	1.667	1.682	1.691	1.697
	.05	1.857	1.934	1.975	2.005	2.022	2.037	2.052	2.057	2.066
	.01	2.571	2.644	2.687	2.718	2.735	2.750	2.764	2.768	2.775
∞	.10	1.459	1.543	1.584	1.612	1.634	1.646	1.655	1.667	1.673
	.05	1.822	1.897	1.942	1.970	1.985	2.002	2.010	2.025	2.027
	.01	2.492	2.565	2.608	2.636	2.655	2.666	2.680	2.683	2.694

Table 7.2: Level Probabilities for Equal Weights.

l	$m = 2$	$m = 3$	$m = 4$	$m = 5$
1	0.41667	0.38231	0.36355	0.35147
2	0.50000	0.49269	0.48685	0.48208
3	0.08333	0.11769	0.13604	0.14750
4		0.00731	0.01315	0.01532
5			0.00041	0.00091
6				0.00272

Table 7.3: Dose Response Data.

Group	Dosage (mg/kg)	Sample size	Mean response	SD response
0	0	6	25.5	2.6
1	.5	6	23.9	4.0
2	1.0	6	27.7	3.3
3	1.5	6	33.4	2.3
4	2.0	6	40.5	10.5
5	2.5	6	57.9	9.9
6	3.0	6	74.4	14.6
7	3.5	6	73.4	7.6
8	4.0	6	73.5	4.5
9	4.5	6	76.2	7.9

Table 7.4: MED Inferred by Three Methods.

δ	MPGN	DR	$T_{q,9}$								
			1	2	3	4	5	6	7	8	9
[0.00, 37.20]	6	6	6	6	6	6	6	6	6	6	6
[38.50, 39.50]	9	6	6	6	6	6	6	6	6	6	6
(40.4040.46)	NA	8	7	7	7	7	6	6	8	8	8
(40.60, 40.70)	NA	9	8	8	8	7	7	6	9	9	9
(40.92, 41.02)	NA	9	9	9	9	9	8	6	9	9	9
(41.02, 41.17)	NA	9	9	9	9	9	9	6	9	9	9
(41.17, 41.95)	NA	9	9	9	9	9	9	9	9	9	9

Table 7.5: Step-down 95% Confidence Lower Bounds for $\mu_i - \mu_0$.

$\mu_i - \mu_0$	MPGN	DR	$T_{q,9}$								
			1	2	3	4	5	6	7	8	9
$\mu_1 - \mu_0$	-11.06	-9.10	-9.10	-9.10	-9.10	-9.10	-9.10	-9.10	-9.10	-9.10	-9.10
$\mu_2 - \mu_0$	-7.06	-5.30	-6.12	-5.30	-5.30	-5.30	-5.30	-5.30	-5.30	-5.30	-5.30
$\mu_3 - \mu_0$	-1.56	0.40	-0.79	-0.42	0.40	0.40	0.40	0.40	0.40	0.40	0.40
$\mu_4 - \mu_0$	5.08	7.50	6.13	6.31	6.68	7.50	7.50	7.50	7.50	7.50	7.50
$\mu_5 - \mu_0$	22.13	24.90	23.40	23.53	23.71	24.08	24.90	24.90	24.90	24.90	24.90
$\mu_6 - \mu_0$	37.22	40.40	39.82	39.90	40.03	40.21	40.58	41.17	40.40	40.40	40.40
$\mu_7 - \mu_0$	37.22	40.40	40.47	40.53	40.60	40.72	40.87	41.17	40.40	40.40	40.40
$\mu_8 - \mu_0$	37.94	40.50	40.74	40.79	40.85	40.92	41.02	41.17	40.74	40.50	40.50
$\mu_9 - \mu_0$	39.57	43.20	41.95	41.99	42.04	42.10	42.17	42.28	42.26	42.14	43.20

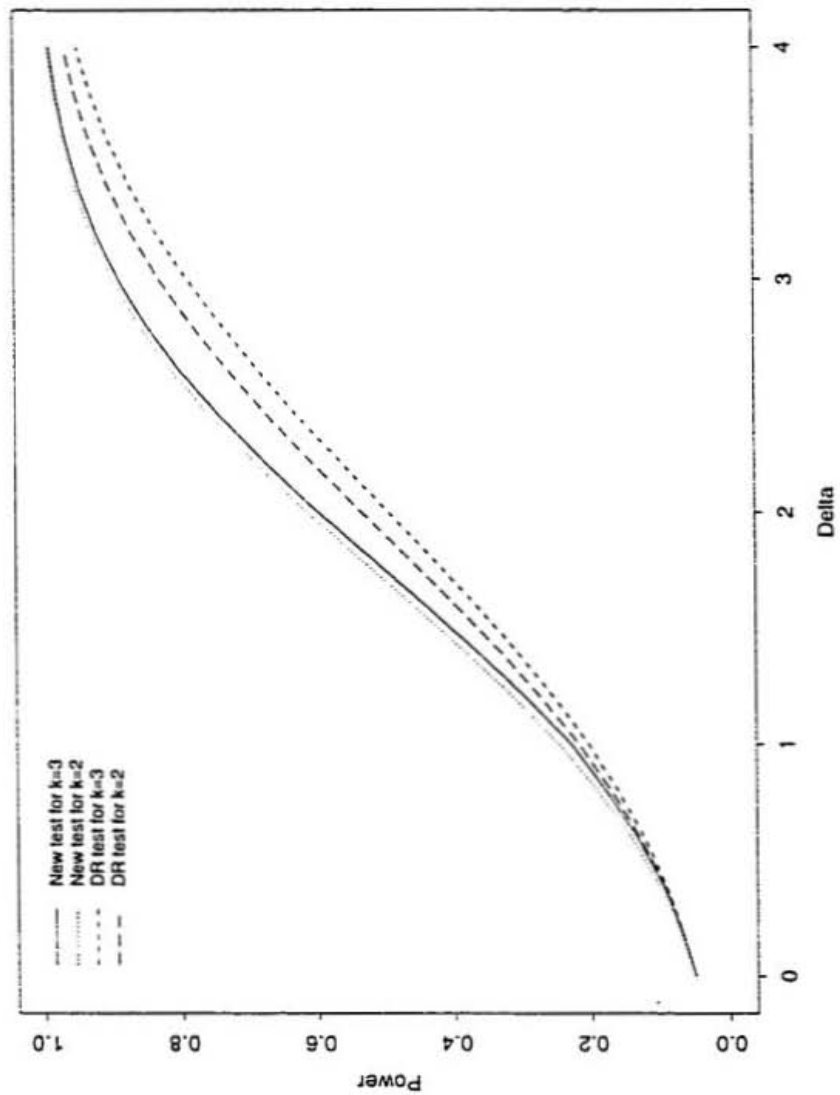

Figure 7.1: Exact Powers of T_k and DR Method for $k = 2$ and 3.

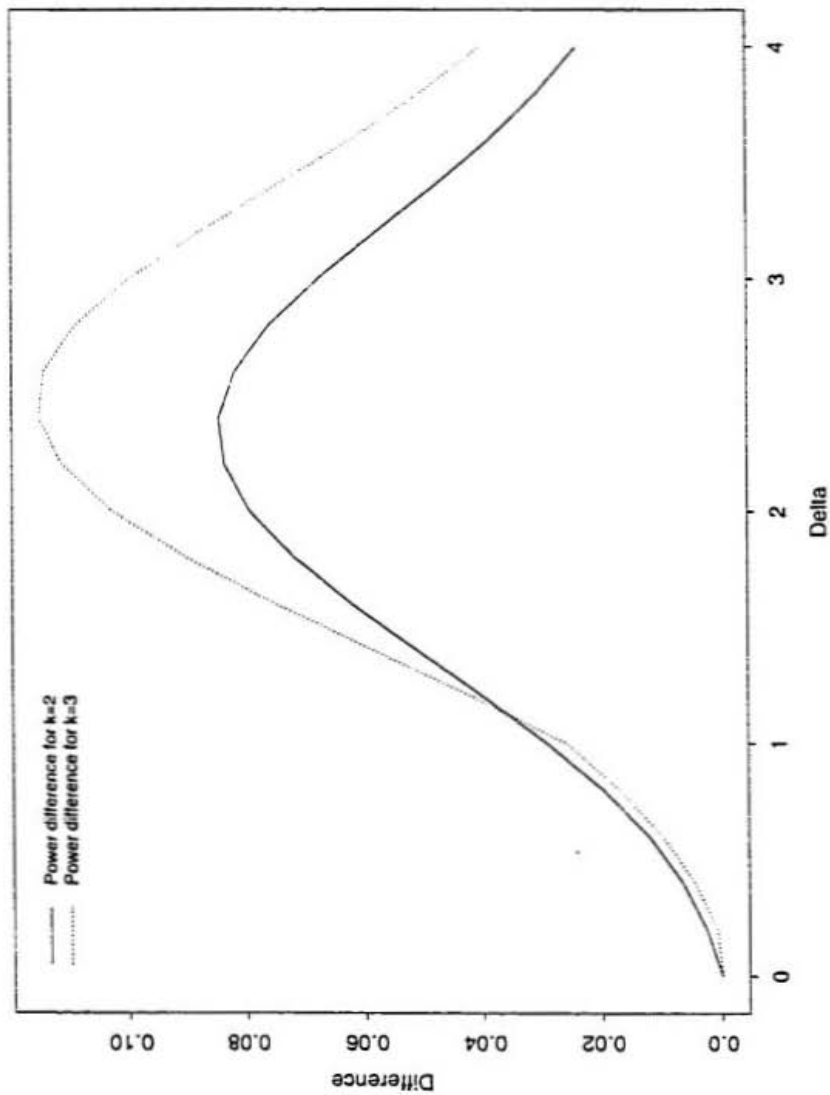

Figure 7.2: Power Differences between T_k and DR Method

Table 7.6: Simulated Probabilities of Identifying the True MED and Simulated Probabilities of Detecting the Dose Responses.

Case	MED	q	δ		Method				
					DR	$T_{q',k}$	$T_{q,k}$	$T_{q'',k}$	D
1	1	1	0.5	I	25.90	NA	31.11	29.41	29.58
				D	83.86	NA	94.02	94.19	90.75
2	2	2	0.5	I	44.77	39.73	46.83	45.58	35.27
				D	94.47	96.03	96.29	96.37	95.32
3	2	2	0.4	I	15.71	16.97	21.87	19.13	13.97
				D	69.92	80.79	81.59	82.12	73.79
4	3	3	1.0	I	33.51	30.18	36.50	34.16	22.99
				D	90.98	93.01	93.63	93.87	90.57
5	3	3	0.5	I	39.67	35.03	41.45	40.28	27.17
				D	92.19	94.26	94.70	94.85	93.08
6	4	4	1.5	I	29.52	26.31	31.67	29.79	17.08
				D	89.02	89.12	90.05	90.86	82.69

Case 1: (0, 2.110, 3.972, 4.169, 4.204, 4.214, 4.217)
Case 2: (0, 0.333, 2.829, 4.725, 5.326, 5.518, 5.590)
Case 3: (0, 0.055, 1.774, 3.262, 3.494, 3.534, 3.543)
Case 4: (0, 0.073, 0.988, 2.996, 4.552, 5.304, 5.639)
Case 5: (0, 0.007, 0.436, 2.698, 4.581, 5.156, 5.314)
Case 6: (0, 0.006, 0.198, 1.255, 3.273, 4.930, 5.784)

Chapter 8

Summary and Further Research

Typically, the goal in comparative clinical trials is to select the treatments that are "better" than the control. Then a one-sided procedure is preferred. On the other hand, when the prior knowledge indicates that treatments are at least effective as the control, sharper statistical procedures can be expected to enhance the inference. It is well known that hypothesis testing does not convey the magnitude of the differences between treatments and the control. However, confidence intervals is more informative than hypothesis testing. In this thesis, we have presented some procedures to yield sharp simultaneous confidence lower bounds for the differences of (combined) treatments and the control. A thorough study has been done on multiple contrast tests. These multiple contrast tests have closed-form null distribution functions. Because all multiple test procedures proposed in this thesis control the type I FWE (see Chapter 1) at level α, only the powers of the procedures are given in the tables.

When the treatments constitute a two-way no presence of interaction model, it is of interest to comparing row (or column) factor means with the control mean.

The one-sided multiple contrast test proposed in Chapter 3 is more efficient than the one-sided Dunntett's procedure.

Once there is a significant difference between the treatments and the control, the magnitude of $\mu_{best} - \mu_0$ is a useful quantity for evaluating the difference between the treatments and the control. We focus on the duality of the maximized confidence lower bound for $\mu_{best} - \mu_0$ and the multiple contrast test statistics T and T^a.

The LRT S_{01} for testing H_0 against H_1 has the strong advantage of good power properties. However, the LRT S_{01} is not cone order monotone. The test statistic T^a for testing H_0 against a more narrow alternative than H_1 in Chapter 5 is cone order monotone. Of course, cone order monotonicity is not necessarily uniformly good, but it is not uniformly bad.

Even if the prior knowledge of treatments are at least as good as the control is available, it is still of interest to consider the equivalence of treatments with respect to the control. The problem is to find the least favorable configuration since the null hypothesis is not a classical homogeneity hypothesis.

In dose-response studies, the typical assumption is monotone response means which is not always the case in real situation. Under more realistic partially monotone assumption, we propose a more efficient test by utilizing the partially monotone assumption to identify the MED. The method in Chapter 7 can also be used to determine the no observed adverse effect level (NOAEL) in safety assessment of toxicological studies.

The constrained optimization problems in Chapter 4 and Chapter 7 are solved through the Kuhn-Tucker equivalence theorem. It is a new insight in comparing treatments with a control. The ideas and approaches presented herein provide a

foundation, and they can be applied to other constrained optimization problems.

In future research relating to this thesis, of particular interest is the situation where the response variable is dichotomous, in which case we are looking at independent binomial populations. This is a very important problem from a practical view, because often an investigator is not measuring a particular response to a drug or medical procedure, but is concerned with the success or failure of the drug or procedure. Future research will also involve the inclusion of variance heterogeneity and the area of nonparametric setting.

References

Ayer, M., Brunk, H. D., Ewing, G. M., Reid, W. T. and Silverman, E. (1955). "An Empirical Distribution Function with Incomplete Information." *Annals of Mathematical Statistics*, **26**, 641-647.

Barlow, R. E., Bartholomew, D. J., Bremner, J. M. and Brunk, H. D. (1972). *Statistical Inference Under Order Restrictions*. John Wiley and Sons, London.

Bartholomew, D. J. (1959a). "A Test of Homogeneity for Ordered Alternatives", *Biometrika*, **46**, 36-48.

Bartholomew, D. J. (1959b). "A Test of Homogeneity for Ordered Alternatives II", *Biometrika*, **46**, 328-335.

Bartholomew, D. J. (1961a). "A Test of Homogeneity of Means Under Restricted Alternatives," *Journal of the Royal Statistical Society, Ser. B*, **23**, 239-281.

Bartholomew, D. J. (1961b). "Ordered Tests in the Analysis of Variance", *Biometrika*, **48**, 325-332.

Bau, J. J., Chen, H. J., and Xiong, M. M. (1993). "Percentage Points of the Studentized Range Test for Dispersion of Normal Means", *Journal of Statistical Computation and Simulation*, **44**, 149-163.

Bauer, P. (1997). "A Note on Multiple Test Procedures for Dose Finding," *Biometrics*, **53**, 1125-1128.

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.

Bechhofer, R. E. and Dunnett, C. W. (1988). "Percentage Points of Multivariate Student t Distributions." *Selected Tables in Mathematical Statistics*, Volume 11, American Mathematical Society, Providence, Rhode Island.

Bechhofer, R. E., Santner, T. J. and Goldsman, D. M. (1995). *Design and Analysis of Experiments for Statistical Selection, Screening and Multiple Comparisons.* John Wiley & Sons, New York.

Berger, J. O. (1985). *Statistical Decision Theory*, 2nd edition, Springer-Verlag, New York.

Berk, R. and Marcus, R. (1996). "Dual Cones, Dual Norms, and Simultaneous Inference for Partially Ordered Means", *Journal of the American Statistical Association*, **91**, 318-328.

Bofinger, E. and Bofinger, M. (1993). "Equivalence of Normal Means Compared with a Control," *Communication in Statistics, Part A - Theory and Methods.* **22**, 3117-3141.

Bofinger, E. and Bofinger, M. (1995). "Equivalence With Respect to a Control: Stepwise Tests," *Journal of the Royal Statistical Society, Ser. B*, **57**, 721-733.

Bofinger, E. and Mengersen, K. (1988). "Simultaneous Comparisons with a Control and with the "Best"", *Communications in Statistics, Part A - Theory and*

Methods, **17**, 1507-1540.

Bohrer, R. (1967). "On Sharpening Scheffé's Bounds." *Journal of the Royal Statistical Society. Ser. B.* **29**, 110-114.

Bohrer, R. and Francis, G. K. (1972). "Sharp One-sided Confidence Bounds Over Positive Regions," *Annals of Mathematical Statistics*, **43**, 1541-1548.

Chakraborti, S. and Desu, M. M. (1991). "Linear Rank Tests for Comparing Several Treatments with a Control when Data Are Subject to Unequal Patterns of Censorship." *Satistica Neerlandica*, **45**, 227-254.

Chakraborti, S. and Hettmansperger, T. P. (1996). "Multi-Sample Inference for the Simple-Tree Alternative Based on One-Sample Confidence Intervals", *Communication in Statistics, Part A - Theory and Methods*, **25**, 2819-2837.

Chen, H. J. and Lam, K. (1991). "Percentage Points of a Studentized Range Statistic From Nonidentical Normal Random Variables," *Communication in Statistics, Part B - Simulation and Computing*, **20**, 995-1047.

Cheung, S. and Holland, B. (1991). "Extension of Dunnett's Multiple Comparison Procedure to the Case of Several Groups", *Biometrics*, **47**, 21-32.

Cheung, S. and Holland, B. (1992). "Extension of Dunnett's Multiple Comparison Procedure with Different Sample Sizes to the Case of Several Groups", *Computational Statistics and Data Analysis*, **14**, 165-182.

Cochran, W. G. and Cox, G. M. (1957). *Experimental Designs*, John Wiley and Sons, New York.

Cohen, A., Kemperman, J. H. B. and Sackrowitz, H. B. (2000). "Properties of Likelihood Inference for Order Restricted Models," *Journal of Multivariate Analysis*, **72**, 50-77.

Cohen, A. and Sackrowitz, H. B. (2000). "Properties of Bayes Testing Procedures in Order Restricted Inference," *Statistics & Probability Letters*, **49**, 205-209.

Cohen, A. and Sackrowitz, H. B. (1998). "Inference for the Model of Several Treatments and a Control." *Technical Report* #98-002. Rutgers University.

Conaway, M., Pillers, C., Robertson, R., and Sconing, J. (1991). "A Circular Cone Test for Testing Homogeneity Against a Simple Tree Order". *The Canadian Journal of Statistics*, **19**, 283-296.

Dunnett, C. W. (1984). "Selection of the Best Treatment in Comparison to a Control with an Application to a Medical Trial". *Design of Experiments*-Ranking and Selection, 47-66.

Dunnett, C. W. (1955). "A Multiple Comparison Procedure for Comparing Several Treatments with a Control," *Journal of the American Statistical Association*, **50**, 1096-1121.

Dunnett, C. W., and Tamhane, A. C. (1997). "Multiple Testing to Establish Superiority/Equivalence of a New Treatment Compared with k Standard Treatments," *Statistics in Medicine*, **16**, 2489-2506.

Dunnett, C. W., and Tamhane, A. C. (1998). "Some New Multiple Test Procedures for Dose Finding," *Journal of Biopharmaceutical Statistics*, **8**, 353-366.

Edwards, D. G. and Hsu, J. C. (1983). "Multiple Comparisons with the Best Treatment," *Journal of the American Statistical Association*, **78**, 965-971.

Fligner, M. A. and Wolfe, D. A. (1982). "Distribution-free Tests for Comparing Treatments With a Control," *Statistica Neerlandica*, **36**, 119-127.

Dykstra, R. L. and Lee, C. I. C. (1991). "Multinomial Estimation Procedures for Isotonic Cones," *Statistics & Probability Letter*, **11**, 155-60.

Genz , A. (1992). "Numerical Computation of Multivariate Normal Probabilities," *Journal of Computational and Graphical Statistics* , **1**, 141-149.

Giani, G. and Straßburger, K. (1994). "Testing and Selecting for Equivalence With Respect to a Control". *Journal of the American Statistical Association*. **89**. 320-329.

Gibbons, J. D., Olkin,I., Sobel. M. (1977). *Selecting and Ordering Populations: A New Statistical Methodology*. John Wiley & Sons, New York.

Gupta, S. S. and Sobel, M. (1958). "On Selecting a Subset Which Contains All Populations Bettre Than a Standard", *Annals of Mathematical Statistics*, **29**, 235-244.

Gupta, S. S. (1965). "On Some Multiple Decision Methods for Ranking Population Means," *Technometrics*, **33**, 248-54.

Hayter, A. J., Liu, W. (1999). "A New Test against an Umbrella Alternative and the Associated Simultaneous Confidence Intervals," *Computational Statistics and Data Analysis*, **30**, 393-401.

Hochberg, Y. and Tamhane, A. (1987). *Multiple Comparison Procedures*, New York: John Wiley.

Hogg, R. (1965). "On Models and Hypotheses with Restricted Alternatives." *Journal of the American Statistical Association*, **60**, 1153-1162.

Hoover, D. R. (1991). "Simultaneous Comparisons of Multiple Treatments to Two (Or More) Controls." *Biometrical Journal*, **33**, 913-921.

Hsu, J. C. and Berger, R. L. (1999). "Stepwise Confidence Intervals Without Multiplicity Adjustment for Dose-Response and Toxicity Studies". *Journal of the American Statistical Association*, **94**, 468-482.

Hsu, J. C. (1996). *Multiple Comparisons: Theory and Methods*, Chapman & Hall. London.

Kwong, K. (2001). "A Modified Dunnett and Tamhane Step-up Approach for Establishing Superiority/Equivalence of a New Treatment Compared with k Standard Treatments." *Journal of Statistical Planning and Inference*, **97**, 359-366.

Korn, E. L. (1982). "Confidence Bands for Isotonic Dose-response Curves," *Applied Statistics*, **31**, 59-63.

Kuhn, H. W. and Tucker, A. W. (1951). "Nonlinear Programming", *2nd Berkeley Symposium., Math. Statist. Probab. (University of California Press, Berkeley)*, 481-492.

Lam, K. (1986). "A New Procedure for Selecting Good Populations", *Biometrika*, **73**, 201-206.

Lee, C. I. C. (1981). "The Quadratic Loss of Isotonic Regression under Normality," *The Annals of Statistics*, **9**, 686-668.

Lee, C. I. C. (1983). "The Min-max Algorithm and Isotonic Regression," *The Annals of Statistics*, **11**, 467-477.

Lee, C. I. C. (1984). "Truncated Bayesian Confidence Region and Its Corresponding Simultaneous Confidnece Intervals in Restricted Normal Models," Department of Mathematics and Statistics, Memorial University of Newfoundland, Technical Report #1984-01.

Lee, C. I. C. (1987). "On the Power Function of $\bar{\chi}^{2}_{3}$", Department of Mathematics and Statistics, Memorial University of Newfoundland, Technical Report #1987-01.

Lee, C. I. C. (1988). "Quadratic Loss of Order Restricted Estimators for Treatment Means with a Control," *The Annals of Statistics*, **16**, 751-758.

Lee, C. I. C. (1996). "On Estimation for Monotone Dose-Response Curves," *Journal of the American Statistical Association*, **91**, 1110-1119.

Lee, C. I. C., Robertson, T. and Wright, F. T. (1993). "Bounds on Distributions Arising in Order Restricted Inferences with Restricted Weights," *Biometrika*, **80**, 405-416.

Liu, W. (1996). "A Group Sequential Procedure for Comparing Several Treatments with a Control", *Journal of Applied Statistics*, **23**, 371-383.

Liu, W. (1997a). "Some Results on Step-up Tests for Comparing Treatments with a Control in Unbalanced One-way Layouts," *Biometrics*, **53**, 1508–1512.

Liu, W. (1997b). "On Sample Size Determination of Dunnett's Procedure for Comparing Several Treatments with a Control," *Journal of Statistical Planning and Inference*, **62** , 255–261.

Marcus, R. (1976). "The Power of Some Tests of the Equality of Normal Means Against an Ordered Alternative", *Biometrika*, **63**, 177-183.

Marcus, R. (1978). "Further Results on Simultaneous Confidence Bounds in Normal Models With Restricted Alternatives", *Communication in Statistics, Part A - Theory and Methods*, **A7**, 573-590.

Marcus, R. and Genizi, A. (1994). "Simultaneous Confidence Intervals for Umbrella Contrasts of Normal Means," *Computational Statistics and Data Analysis*, **17**, 393-407.

Marcus, R. and Peritz, E. (1976). "Some Simultaneous Confidence Bounds in Normal Models with Restricted Alternatives," *Journal of the Royal Statistical Society, Ser. B*, **38**, 157-165.

Marcus, R., Peritz, E. and Gabriel, K. R. (1976). "On Closed Testing Procedures with Special Reference to Ordered Analysis of Variance," *Biometrika*, **63**, 655-660.

Marcus, R. and Talpaz, H. (1992). "Further results on testing homogeneity of normal means against simple tree alternatives," *Communications in Statistics, Part A-Theory and Methods*, **21**, 2135-2149.

Miwa, T., Hayter, A. J. and Liu, W. (2000). "Calculations of Level Probabilities for Normal Random Variables with Unequal Variances with Applications to Bartholomew's Test in Unbalanced One-way Models", *Computational Statistics and Data Analysis*, **34**, 17-32.

Miller, R. G. Jr. (1981). *Simultaneous Statistical Inference*, New York: Spring-Verlag.

McDermott, M. P. (1999). "Generalized Orthogonal Contrast Tests for Homogeneity of Ordered Means", *The Canadian Journal of Statistics*, **3**, 457-470.

Mukerjee, H., Robertson, T., and Wright, F. T. (1987). "Comparison of Several Treatments with a Control Using Multiple Contrasts," *Journal of the American Statistical Association*, **82**, 902-910.

Peng, J., Lee, C. C., and Liu, L. (2000). "A Multiple Comparisons Procedure for Detecting Differences Between Treatments and a Control in Two-factor Experiments", *Communication in Statistics, Part A - Theory and Methods*, **29**, 1197-1218.

Qian, S. (1992). "Minimum Lower Sets Algorithms for Isotonic Regression," *Statistics & Probability Letters*, **15**, 31–35.

Robertson, T., and Wegman, E. J. (1978). "Likelihood Ratio Tests for Order Restrictions in Exponential Families," *The Annals of Statistics*, **6**, 485-505.

Robertson, T., Wright, F. T. (1985). "One-sided Comparisons for Treatments with a Control," *The Canadian Journal of Statistics*, **13**, 109-122.

Robertson, T., Wright, F. T. and Dykstra, R. L. (1988). *Ordered Restricted Statistical Inference*, Chichester, U.K.: John Wiley.

Rockafellar, R. T. (1972). *Convex Analysis*, Princeton, New Jersey: Princeton university Press.

Rom, D. M., Costello, R. J., and Connell, L. T. (1994). "On Closed Test Procedures for Dose-Response Analysis," *Statistics in Medicine*, **13**, 1583-1596.

Ruberg, S. J. (1989). "Constrasts for Identifying the Mininmum Effective Dose," *Journal of the American Statistical Association*, **84**, 816-822.

Ruberg, S. J. (1995). "Dose-Response Studies II. Analysis and Interpretation," *Journal of Biopharmaceutical Statistics*, **5**, 15-42.

Ryan, L. (1992). "Quantitative Risk Assessment for Developmental Toxicity," *Biometrics*, **41**, 295-301.

Scheffé, H. (1953). "A Method for Judging All Contrasts in the Analysis of Variance", *Biometrika*, **40**, 87-104.

Schoenfeld, D. A. (1986). "Confidence Bounds for Normal Means Under Order Restrictions, with Applications to Dose-response Curves, Toxicology Experi-