SIMULTANEOUS STATISTICAL INFERENCE FOR MONOTONE DOSE-RESPONSE MEANS

LIN LIU

Your file Votre référence
ISBN: 0-612-89695-1
Our file Notre référence
ISBN: 0-612-89695-1

Simultaneous Statistical Inference for Monotone Dose-Response Means

by

©Lin Liu

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

June 2001

St. John's Newfoundland Canada

Abstract

Statistical inference under order restrictions is an important field in statistical science and has been studied and practiced widely. The utilization of the assumption of monotonicity increases the efficiency of statistical inference procedures. This can be found in the literature such as Ayer, Brunk, Ewing, Reid and Silverman (1955), Robertson and Wright (1974), Barlow and Ubhaya (1971), Lee (1981), Kelly (1989), Korn (1982), Schoenfeld (1986), Hayter (1990) and Lee (1996). In Chapter 2, we review some fundamental theories about the order restricted statistical inference including isotonic regression and test of a simply ordered hypothesis.

In Chapter 3, we study a max-min interval procedure, a modification of Tukey's studentized range technique, to construct simultaneous confidence intervals for pairwise comparisons of response means by utilizing the prior knowledge of the monotonicity of the means. The improvement of the proposed max-min interval procedure is substantial.

The one-sided simultaneous confidence lower bound is studied in Chapter 4. We investigate the incomplete optimization problem of maximizing simultaneous lower bounds for nonnegative contrasts considered by Marcus (1978). Significant improvements over Marcus' (1978) results, including a necessary and sufficient condition for the optimal solution and an efficient computation algorithm to compute the optimal lower bounds, are made.

In Chapter 5, we introduce a one-sided multiple comparison test (OMCT) for testing the homogeneity of the means against the simple order alternative.

It gives sharper one-sided simultaneous confidence lower bounds. This OMCT approach compares favorably with Hayter's (1990) and Marcus' (1978) approaches and it may be comparable to the least significant difference approach.

The simultaneous statistical inference for response means with a control is considered in Chapter 6. An orthant test statistic is introduced. With the prior knowledge that the response means are monotone, a more efficient simultaneous confidence lower bound can be inverted from this test to detect the difference between response means and the control mean. An algorithm to compute the optimal lower bound is included.

In Chapter 7, we demonstrate that the stepwise test procedure based on likelihood ratio test is a more efficient test procedure for detecting the minimum efficient dose in dose-response studies.

Acknowledgments

I am very grateful to my supervisor, Dr. Chu-In Charles Lee, for leading me to the research field of this thesis, and for his many helpful and thoughtful comments, discussions and suggestions throughout the preparation of this thesis. He has always been generous with his time and ideas. This work would not have been completed without his guidance, encouragement, patience, and financial support. I came as a student, and I hope I leave as a researcher.

I would also like to thank my co-supervisor, Dr. Hong Wang, for his concern, encouragement and help during my graduate study and my colleague, Jianan Peng for his sincere friendship.

I would express my thanks to the School of Graduate Studies and the Department of Mathematics and Statistics, Memorial University of Newfoundland for supporting me with a Graduate Fellowship and a Teaching Assistantship.

My gratitude will also go to Dr. Farroll Tim Wright, Dr. Gary Sneddon and Dr. Paul Peng. They read my thesis very carefully and gave many helpful comments for the better presentation of this thesis.

I would like to express my sincere appreciation to my husband, Yang, for his continued encouragement, care, understanding and love. Finally, I would especially like to thank my mother and brother for their love and support during this program and dedicate this thesis to the memory of my father, X. J. Liu.

Lin Liu

St. John's

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Order restricted statistical inference has been researched and practiced for the last 50 years. Many types of problems are concerned with identifying meaningful structure in real world situations. Structure involving orderings and inequalities has many useful applications. For example, the probability of a particular response may increase with the treatment level; a regression function may be nondecreasing; the failure rate of a component may increase as it ages; or the treatment response may stochastically dominate the control. Hundreds of research papers have been published on this topic and many of them can be traced through the bibliographies of two books: Barlow, Bartholomew, Bremner and Brunk (1972), and Robertson, Wright and Dykstra (1988).

Utilizing the prior knowledge of ordering, including the ordering of parameters, the ordering of distribution functions, and other related constraints can increase the efficiency of statistical inference procedures. The incorporation of this prior knowledge into estimation makes the estimates superior to the ordinary one. For example, the isotonic regression (see definition 1.3.3 in Robertson, Wright and Dykstra 1988) can reduce the total square error

(Ayer, Brunk, Ewing, Reid and Silverman 1955) and the maximum absolute error (Robertson and Wright 1974, Barlow and Ubhaya 1971). The reduction of mean square error for the normal means problem with a simple order was deduced by Lee (1981). Lee (1988) also observed that this property does not hold, in general, for partial order restrictions. Furthermore, Kelly (1989) showed that the isotonic regression estimator of the normal mean is superior to the ordinary one under any nonconstant loss which is a nondecreasing function of absolute error.

It is also a common view that a more powerful test can be obtained by taking the additional knowledge into account. For example, considering a one-sided alternative leads to more powerful tests. But caution should be taken to interpret the result from such a test. In particular, without prior knowledge that strongly supports the assumption of one-sidedness, it may be misleading to interpret the rejection of the null hypothesis as evidence supporting the alternative hypothesis.

The classical likelihood ratio test (LRT), which is denoted by $\bar{\chi}^2_{01}$ or $\bar{E}^2_{01}$, for testing the equality of partially ordered means from several normal populations was first proposed by Bartholomew (1959a, 1959b, 1961a, 1961b). It is known to possess generally superior operating characteristics to those of its competitors (Robertson, Wright and Dykstra 1988). Tests for identifying the structures with order restrictions often require good estimates under inequality constraints. However, difficulties in computing the restricted maximum likelihood estimates and determining the null distribution of the test statistics make the LRT difficult to implement in many instances. Therefore, the approximations to these distributions are of considerable interest. Bartholomew (1959a, 1959b) proposed a two-moment Chi-square approximation for the null distribution of $\bar{\chi}^2_{01}$. Siskind (1976) and Grove (1980) conjectured that the null distributions of LRT would not be sensitive to moderate variations in the weights and this has been investigated by Robertson and Wright (1983), and Wright and Tran (1985) for the simple order and the simple tree order. Another approach has been to obtain sharp upper and lower bounds on the tail probabilities for the LRT. These bounds, which give the most extreme possible error for the equal weights approximation, were studied by Robertson and Wright (1982), Wright and Tran (1985), and Lee, Robertson and Wright (1993).

Several other researchers, including Abelson and Tukey (1963), Hogg (1965), Schaafsma and Smid (1966), and more recently Snidjers (1979), considered the tests based on contrasts. One advantage of these tests is that the contrast statistic is normally distributed with easily computed mean and variance under both the null and alternative hypotheses. Such a contrast test is easily shown to be uniformly most powerful for alternatives in a certain direction. Consequently it is very powerful in some subregion of the alternative hypothesis and less powerful in other directions. While the LRT is not most powerful at any particular point, it maintains a more uniform power over all the alternative regions. The aforementioned contrast tests can not compete with the LRT in general. The multiple contrast test is another approach that may be comparable to the LRT. Dunnett's test (1955) for testing against a simple tree alternative is surely the best known and most widely used. Van Eeden (1958) and Williams (1971, 1972) proposed *ad hoc* tests. The properties of the *ad hoc* tests have also been shown to be generally inferior to those of the LRT (Chase 1974, Robertson and Wegman 1978). Mukerjee, Robertson and Wright (1987) introduced the multiple contrast test based on orthogonal contrasts. And most recently, McDermott (1999) proposed a class of tests based on an orthant approximation which can be viewed as generalizations of the orthogonal contrasts test proposed by Mukerjee, Robertson and Wright (1987).

Significant contributions have been made in the literature for testing homogeneity against ordered alternatives. But confidence interval procedures involving order restrictions have been somewhat slow in developing. The pioneering work in the development of simultaneous confidence intervals for restricted settings was made by Bohrer (1967) and Bohrer and Francis (1972). 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) described simultaneous one-sided confidence bounds in this restricted setting. Marcus and Peritz (1976) also developed methodology for finding simultaneous confidence intervals for linear combinations of normal means with certain restrictions on the coefficients. Marcus (1978) was able to improve Bohrer and Francis bounds when prior information is available on the parameters. The evaluation of the improved simultaneous confidence lower bound is a concave programming problem. Deriving a computation algorithm to search for an optimal solution to this concave programming problem is a new and challenging work and has not received much attention. Kuhn-Tucker equivalence theorem (Kuhn and Tucker 1951) will help us to resolve the difficulties and the application of this theorem will be discussed in detail in this thesis.

Simultaneous statistical inference received interest after the development of research on multiple comparisons and simultaneous confidence intervals. The fundamental contributions by Tukey and Scheffé on this area can be found in the monograph by Miller (1981). Berk and Marcus (1996) studied simultaneous inference for partially ordered means. Other simultaneous inference procedures can be obtained in Hochberg and Tamhane (1987) and Hsu (1996). In this thesis our interest will focus on simultaneous statistical inferences with order restrictions.

It is of considerable interest to study the monotone regression curves with independent normal errors. In the dose-response studies, we usually assume the dose-response mean $\mu_i = f(x_i), i = 1, ..., k$, is a monotone, nondecreasing function of the dose level x_i. The prior knowledge of monotonicity of regression curves can be used to increase the efficiency of the maximum likelihood estimate as shown by Lee (1981). Korn (1982), Schoenfeld (1986) and Lee (1996) all sought confidence intervals for each individual mean μ_i by incorporating this monotonicity. The generalized studentized maximum modulus procedure by Lee (1996) gains much over the Scheffé-type procedure by Schoenfeld (1986) and the studentized maximum modulus by Korn (1982). Hayter (1990) proposed the one-sided studentized range test (OSRT) to construct a one-sided simultaneous confidence lower bound for the pairwise mean comparison $\mu_j - \mu_i$ for the balanced one-way analysis of variance model. Hayter (1992) generalized the OSRT procedure to an unbalanced model with three populations. How to detect the difference between the monotone nondecreasing means efficiently is the main subject which will be pursued in this thesis.

In Chapter 2, we introduce some basic concepts in order restricted inference. Section 2.1 consists of the definition of a simple order, isotonic regression for a simple order restriction and the algorithms to obtain the isotonic regression. In Section 2.2, the likelihood ratio tests for testing the simply ordered alternative and their relationship with the linear contrasts are given. In Section 2.3, some results about interval estimation of the simply ordered parameters are introduced. The Kuhn-Tucker equivalence theorem, which will be used in Chapter 4 and Chapter 6, is given in Section 2.4.

In Chapter 3, the two-sided simultaneous inference will be studied. A simple novel procedure that modifies Tukey's studentized range technique is proposed to construct simultaneous confidence intervals for pairwise comparisons of means by utilizing the prior knowledge of the monotonicity of the response curve. The new procedure is a substantial improvement over its predecessor.

In Chapter 4, we will study the problem considered by Marcus (1978). She introduced the optimization problem of maximizing simultaneous lower bounds for nonnegative contrasts $\sum_{i=1}^{k} n_i c_i \mu_i, \sum_{i=j}^{k} n_i c_i \geq 0, j = 2, ..., k$ and $\sum_{i=1}^{k} n_i c_i = 0$ with the prior knowledge that $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_k$. However, her result is incomplete. We will propose a necessary and sufficient condition for the optimal solution and an efficient computation algorithm to compute the optimal lower bounds for pairwise comparisons and nonnegative contrasts.

In Chapter 5, a new simple one-sided multiple comparison test (OMCT) is introduced to test the null hypothesis $H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$ against the alternative hypothesis $H_1 : \mu_1 \leq \mu_2 \leq \cdots \leq \mu_k$. It can be used to construct the efficient one-sided simultaneous confidence lower bounds for pairwise comparisons and nonnegative contrasts. It is advantageous in categorizing dosage levels. This OMCT approach compares favorably with Hayter's (1990) and Marcus' (1978) approaches and it may be comparable to the least significant difference approach.

In Chapter 6, we will consider the simultaneous statistical inference with a zero-dose control. An orthant test statistic is introduced. Its power compares favorably with other procedures. With the prior knowledge that the dose-response curve is monotone, a more efficient simultaneous confidence lower bound can be inverted from this test to detect the difference between the dose response mean and the zero dose control mean. An algorithm to compute the optimal lower bound is also included.

In Chapter 7, we will study the stepwise procedure for detecting the minimum efficient dose when the control mean and dose-response means satisfy the simple order $\mu_0 \leq \mu_1 \leq \cdots \leq \mu_k$. Likelihood ratio test and multiple comparison tests are considered. It will be shown by a simulation study that LRT is a more efficient test procedure.

In Chapter 8, we will give a brief summary of the studies in this thesis.

Chapter 2

Order Restricted Statistical Inference

2.1 Maximum Likelihood Estimate Under Order Restrictions

2.1.1 Simple Order

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

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

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

3. it is *antisymmetric*: $x, y \in X, x \preceq y, y \preceq x$ imply $x = y$;

4. every two elements are *comparable*: $x, y \in X$ implies either $x \preceq y$ or $y \preceq x$.

A simple order on the finite set X is in the form of $x_1 \preceq x_2 \preceq \cdots \preceq x_k$. A binary relation $\preceq$ is a *partial order* if it is reflexive, transitive, and antisymmetric, but there may be noncomparable elements. The simple tree order: $x_1 \preceq x_i, i = 2, ..., k$, is an example of a partial order.

The simple order is one of the most important orders and has many useful applications. This will be evident throughout this thesis.

2.1.2 Isotonic Regression with a Simple Order Restriction

Let X be a finite set $\{x_1, x_2, ..., x_k\}$ with a simple order $x_1 \preceq x_2 \preceq \cdots \preceq x_k$. Then a real valued function f on X is isotonic if $f(x_1) \leq f(x_2) \leq \cdots \leq f(x_k)$.

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

$$\sum_{i=1}^{k}[g(x_i) - f(x_i)]^2 w(x_i)$$

in the class of all isotonic functions on X.

Suppose g and w are functions defined on X, set

$$Av(s,t) = \frac{\sum_{i=s}^{t} w(x_i)g(x_i)}{\sum_{i=s}^{t} w(x_i)}$$

for $s \leq t$. $Av(s,t)$ depends on g, this will not be made explicit in the notation.

Theorem 2.1.1 *The isotonic regression of g is given by*

$$\begin{aligned} g^*(x_i) &= \max_{1\leq s\leq i}\min_{i\leq t\leq k} Av(s,t) \\ &= \max_{1\leq s\leq i}\min_{s\leq t\leq k} Av(s,t) \\ &= \min_{i\leq t\leq k}\max_{1\leq s\leq i} Av(s,t) \\ &= \min_{i\leq t\leq k}\max_{1\leq s\leq t} Av(s,t) \end{aligned}$$

(Robertson, Wright and Dykstra, 1988)

Theorem 2.1.2 *If* $\{s : g^*(x_s) = c\} = \{i, i+1, ..., j\}$, *then* $c = Av(i,j)$. *(Th1.3.5 Robertson, Wright and Dykstra, 1988)*

Theorem 2.1.2 reduces the problem of computing g^* to finding the sets on which g^* is constant (i.e. its level sets). The calculation of g^*, given g, the weight w, and the simple order on X, can be accomplished via quadratic programming. An extensive literature on methods for computing quadratic programming solutions for such a problem exists. A number of algorithms have been proposed for computing the isotonic regression. We will introduce two of them in the next subsection that have been used extensively, namely the *pool-adjacent-violators algorithm* (PAVA) and the *minimum lower sets algorithm.*

The utilization of the simple order information in estimation makes the estimates superior to the ordinary one. Lee (1981) shows that mean square error is reduced for every individual mean by using the order restricted maximum likelihood estimate (MLE) of the simply ordered normal means. Kelly (1989) obtained an even stronger result that the absolute error of each component of the isotonic regression estimator is stochastically smaller than that of the usual estimator.

2.1.3 Algorithms for Isotonic Regression for a Simple Order

Pool-Adjacent-Violators algorithm (PAVA)

The PAVA starts with g. If g is isotonic then $g^* = g$. Otherwise, there must exist a subscript $i, 2 \leq i \leq k$, such that $g(x_{i-1}) > g(x_i)$. These two values are then replaced by their weighted average, namely, $Av(i-1,i) = [g(x_{i-1})w(x_{i-1}) + g(x_i)w(x_i)]/[w(x_{i-1}) + w(x_i)]$ and their weights by $w(x_{i-1}) + w(x_i)$. If this new set of $k-1$ values is isotonic, then $g^*(x_{i-1}) = g^*(x_i) = Av(i-1,i)$ and $g^*(x_j) = g(x_j)$ otherwise. If this new set of values is not isotonic, then this process is repeated using the new values and new weights until an isotonic set of values is obtained. The value of $g^*(x_i)$ is the weighted average over the block in which x_i is contained.

Minimum lower sets algorithm

A subset L of X is called a lower set with respect to the simple order $\preceq$ if $y \in L$ and $x \preceq y$ imply $x \in L$. A subset U of X is called an upper set if $x \in U$ and $x \preceq y$ imply $y \in U$. 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 $B = L \cap U$. There are exactly k nonempty lower sets and exactly k nonempty upper sets. The set X is both a lower set and an upper set. The other lower sets are sets of the form $\{x_1, x_2, ..., x_i\}, i = 1, 2, ..., k-1$, and the upper sets are sets of the form $\{x_i, x_{i+1}, ..., x_k\}, i = 2, ..., k$.

Set

$$g^*(x_i) = Av(1, i_1) = \min\{Av(1,j) : 1 \leq j \leq k\} \quad \text{for} \quad i = 1, 2, ..., i_1.$$

Now consider the averages of the sets $\{i_1 + 1, ..., i\}$ for all $i_1 < i \leq k$ and set

$$g^*(x_i) = Av(i_1+1, i_2) = \min\{Av(i_1+1, j) : i_1 < j \leq k\} \quad \text{for} \quad i = i_1+1, i_1+2, ..., i_2.$$

This process is continued until $g^*(x_k)$ is determined.

2.2 Test of a Simply Ordered Hypothesis

Many of the methods of statistical inference are derived from the problem of comparing several normal populations. It is often useful to begin the analysis by testing the null hypothesis that the means are equal. However, in applications, a researcher may believe a *priori* that the means are simply ordered. When this is so, it would be expected that more powerful tests could be devised. In this section, the likelihood ratio tests (LRTs) for homogeneity of normal means with a simple order restricted alternative are introduced. If the simple order imposed on the alternative is in question, one may wish to test this order restriction as the null hypothesis with an unrestricted alternative. The LRTs in this setting are also given in this section. In the meantime, we will demonstrate the relationship between the LRT functions and the class of linear functions of the sample means.

Let $X = \{1, 2, ..., k\}$ and assume that the simple order $\preceq$ is defined on X. Let μ_i be the mean of a normal population with variance σ_i^2 for $i = 1, ..., k$. We denote the mean vector by $\mu = (\mu_1, ..., \mu_k)'$. We are interested in the following hypotheses

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k,$$

$$H_1 : \mu_1 \leq \mu_2 \leq \cdots \leq \mu_k,$$

and

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

Suppose that $\bar{Y}_i$ is the mean of a random sample of size n_i from a normal population with unknown mean μ_i and variance of the form $\sigma_i^2 = a_i\sigma^2$ with the a_i being known positive constants and suppose that the samples are independent. Under H_0, the MLE of $\mu_i, i = 1, ..., k$, is given by $\hat{\mu} = \sum_{i=1}^k w_i\bar{Y}_i / \sum_{i=1}^k w_i$ with $w_i = n_i/a_i$. Under H_1, the maximum likelihood estimate (MLE) of μ is $\mu^* = (\mu_1^*, ..., \mu_k^*)$, the isotonic regression of $\bar{\mathbf{Y}} = (\bar{Y}_1, ..., \bar{Y}_k)$ with weight vector $\mathbf{w} = (w_1, w_2, ..., w_k)$ and the simple order $\preceq$ which determines H_1. The unrestricted MLE of μ is $\bar{\mathbf{Y}}$. Let s^2 be an estimator of σ^2 which is independent of $\bar{\mathbf{Y}}$ with $\nu s^2/\sigma^2 \sim \chi_\nu^2$ with $\nu = \sum_{i=1}^k n_i - k > 0$ (χ_ν^2 denotes a Chi-square variable with ν degrees of freedom).

Consider testing H_0 versus $H_1 - H_0$, the LRT rejects H_0 for large values of

$$S_{01} = \frac{\nu\bar{\chi}_{01}^2}{\bar{\chi}_{12}^2 + Q(\nu)}. \tag{2.1}$$

where

$$\bar{\chi}_{01}^2 = \sum_{i=1}^k w_i(\mu_i^* - \hat{\mu})^2/\sigma^2, \tag{2.2}$$

$$\bar{\chi}_{12}^2 = \sum_{i=1}^k w_i(\bar{Y}_i - \mu_i^*)^2/\sigma^2, \tag{2.3}$$

and $Q(\nu) = \nu s^2/\sigma^2$. If σ^2 is known, $\bar{\chi}_{01}^2$ is the LRT for testing H_0 versus $H_1 - H_0$.

The LRT of H_1 versus $H_2 - H_1$ rejects H_1 for large values of

$$S_{12} = \frac{\bar{\chi}_{12}^2}{Q(\nu)/\nu}. \tag{2.4}$$

and $\bar{\chi}_{12}^2$ is the LRT for testing H_1 versus $H_2 - H_1$ when σ^2 is known.

Let $P_S(l, k; \mathbf{w})$ denote the level probability that there are exactly l distinct values (levels) for the MLE μ^* satisfying the simple order $\preceq$ when H_0 is true.

The $P_S(l,k;\mathbf{w})$'s depend on the sample sizes and the population variances through the weights w_i. Let $1 \leq m \leq k$ and let $B_1, B_2, ..., B_m$ be a partition of X where $B_j = \{i_{j-1}+1, i_{j-1}+2, ..., i_j\}, j = 1, ..., m (i_0 = 0)$. Let $\mathcal{L}_{mk}$ be the collection of all the possible decompositions $(B_1, B_2, ..., B_m)$ of X. Set $W_{B_j} = \sum_{i \in B_j} w_i, C_{B_j} = i_j - i_{j-1}$ and $\mathbf{w}(B_j) = (w_{i_{j-1}+1}, w_{i_{j-1}+2}, ..., w_{i_j})$. For a given decomposition, define $\preceq'$ on $\{1, 2, ..., m\}$ by $i \preceq' j$ if $i \leq j$. Let $P_S(m, m; W_{B_1}, W_{B_2}, ..., W_{B_m})$ be the probability of m levels with the simple order $\preceq'$ and the weight vector $(W_{B_1}, W_{B_2}, ..., W_{B_m})$ and let $P_S(1, C_{B_j}; \mathbf{w}(B_j))$ be the probability of one level with the simple order $\preceq$ and the weight vector $\mathbf{w}(B_j)$.

Theorem 2.2.1 *For* $m \in \{2, 3, ..., k-1\}$,

$$P_S(m,k;\mathbf{w}) = \sum_{\{B_1, B_2, ..., B_m\} \in \mathcal{L}_{mk}} P_S(m, m; W_{B_1}, W_{B_2}, ..., W_{B_m}) \prod_{i=1}^{m} P_S(1, C_{B_i}; \mathbf{w}(B_i)).$$

(Robertson, Wright and Dykstra, 1988)

The above theorem provides a recursive formula for calculating $P_S(l,k;\mathbf{w})$; however, it can be tedious to use. When the weights are equal, $P_S(l,k;\mathbf{w})$ is denoted by $P_S(l,k)$ and it can be obtained by the following theorem.

Theorem 2.2.2

$$P_S(1,k) = \frac{1}{k}$$
$$P_S(k,k) = \frac{1}{k!}$$

and

$$P_S(l,k) = \frac{1}{k} P_S(l-1, k-1) + \frac{k-1}{k} P_S(l, k-1) \quad \textit{for} \quad i = 2, 3, ..., k-1.$$

(Robertson, Wright and Dykstra, 1988)

Numerical values of $P_S(l,k)$ are given in Table $A.10$ (Robertson, Wright and Dykstra 1988). Robertson and Wright (1983) have shown that $P_S(l,k;\mathbf{w})$ are robust to small deviations in the weights and give an approximation for these mixing coefficients for unequal weights. The null distributions of $\bar{\chi}^2_{01}, \bar{\chi}^2_{12}, S_{01}$ and S_{12} are given by the following theorem which is equivalent to the corollary of Theorem 2.3.1 by Robertson, Wright and Dykstra (1988).

Theorem 2.2.3 *For $\mu \in H_0$, ν a positive integer and $N = \sum_{i=1}^k n_i$*

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

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

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

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

for any $c > 0$.

For the case in which the weights are equal, i.e., $w_1 = \cdots = w_k$, the critical values for the above tests are given in Table $A.4, A.6$ and $A.7$ of Robertson, Wright and Dykstra (1988).

Hogg (1965) discussed the relationship between the likelihood ratio function and the class of linear functions of the sample mean $\bar{Y}_i$. Without loss of generality, we assume that $\sum_{i=1}^k n_i c_i = 0$ and $\sum_{i=1}^k n_i c_i^2 = 1$ for the linear contrast $\sum_{i=1}^k n_i c_i \bar{Y}_i$, and the k populations have an equal known variance σ^2.

Suppose that $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_k$. The following result was given by Hogg (1965).

Theorem 2.2.4 *(Hogg, 1965)*

$$\sqrt{\bar{\chi}^2_{01}} = \{\max \sum_{i=1}^{k} n_i c_i \bar{Y}_i / (\sigma^2 \sum_{i=1}^{k} n_i c_i^2)^{1/2}\} \tag{2.5}$$

subject to c_i *satisfies the simple order as* μ_i. *The maximum is attained at* $c_i^* = (\mu_i^* - \hat{\mu})/\sqrt{\sum_{i=1}^{k} n_i(\mu_i^* - \hat{\mu})^2}$.

The results discussed in this section can be generalized to the other partial orders (see Robertson, Wright and Dykstra, 1988).

2.3 Interval Estimations

The pioneering work in the development of simultaneous confidence intervals for restricted settings was carried out by Bohrer (1967), with further refinements found in Bohrer and Francis (1972). Scheffé (1959) provided a method of constructing confidence bounds on a linear function

$$f(\mathbf{x}) = \beta \mathbf{x} = \sum_{i=1}^{k} \beta_i x_i.$$

These bounds are based on samples $y(\mathbf{x})$ (observed values of $\sum_{i=1}^{k} \beta_i x_i$) which are normally distributed with mean $f(\mathbf{x})$. Bohrer (1967) gave sharper confidence bounds for a linear function of nonnegative arguments by extending Scheffé's (1959) confidence bounds. In particular, assume $\hat{\beta}_1, \hat{\beta}_2, ..., \hat{\beta}_k$, the least squares estimator of $\beta_1, \beta_2, ..., \beta_k$, are independent normal random variables with respective means $\beta_1, \beta_2, ..., \beta_k$ and known variances $\sigma_1^2, \sigma_2^2, ..., \sigma_k^2$.

Let $X^+ = \{\mathbf{x} : x_i \geq 0 \text{ for } 1 \leq i \leq k\}$, Bohrer (1967) proposed the following $100(1-\alpha)\%$ simultaneous confidence intervals for $\sum_{i=1}^k \beta_i x_i$,

$$\sum_{i=1}^k \hat{\beta}_i x_i - c(\sum_{i=1}^k x_i^2 \sigma_i^2)^{1/2} \leq \sum_{i=1}^k \beta_i x_i \leq \sum_{i=1}^k \hat{\beta}_i x_i + c(\sum_{i=1}^k x_i^2 \sigma_i^2)^{1/2}$$

where c is determined by

$$\sum_{i=0}^k 2^{-k} \binom{k}{i} P(\chi_i^2 \leq c^2) P(\chi_{k-i}^2 \leq c^2) = 1 - \alpha.$$

The table values $c = c(\alpha, k)$ were given by Bohrer (1967). For large k, the simultaneous confidence bounds for $\sum_{i=1}^k \beta_i x_i$ when $\mathbf{x}$ is restricted to the positive orthant are up to 30 percent shorter than Scheffé's (1959) bound. Bohrer and Francis (1972) extended the above development to the case when $\hat{\beta}_1, \hat{\beta}_2, ..., \hat{\beta}_k$ are not independent and gave one-sided simultaneous confidence bounds for $\sum_{i=1}^k \beta_i x_i$.

Marcus and Peritz (1976) employed the critical point d_α of Bartholomew's LRT (1959a, 1959b, 1961a) for the simple order alternative to construct the one-sided simultaneous confidence lower bound for monotone contrasts $\sum_{i=1}^k n_i c_i \mu_i$ where $\sum_{i=1}^k n_i c_i = 0$ and $c_1 \leq c_2 \leq \cdots \leq c_k$. Assume that $\bar{Y}_i, i = 1, ..., k$, are normal random variables with mean μ_i and variance σ^2/n_i where σ^2 is known. The lower bound for $\sum_{i=1}^k n_i c_i \mu_i$ is of the form

$$\sum_{i=1}^k n_i c_i \bar{Y}_i - \sigma d_\alpha (\sum_{i=1}^k n_i c_i^2)^{1/2}.$$

If the means μ_i are simply ordered, i.e., $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_k$, the lower bound for the monotone contrast $\sum_{i=1}^k n_i c_i^* \mu_i$ can be improved to

$$\max\{\sum_{i=1}^k n_i c_i \bar{Y}_i - \sigma d_\alpha (\sum_{i=1}^k n_i c_i^2)^{1/2}\} \quad (2.6)$$

subject to $\sum_{i=1}^{k} n_i c_i \mu_i \leq \sum_{i=1}^{k} n_i c_i^* \mu_i, c_i \leq c_{i+1}, \sum_{i=1}^{k} n_i c_i = 0, \mu_1 \leq \mu_2 \leq \cdots \leq \mu_k$. Further work on simultaneous confidence intervals for the class of monotone contrasts can be found in Williams (1977) and Marcus (1982).

Marcus (1978) studied the one-sided simultaneous confidence lower bound for more general nonnegative contrasts. The nonnegative contrast $\sum_{i=1}^{k} n_i c_i \mu_i$, where $\sum_{i=j}^{k} n_i c_i \geq 0, j = 2, ..., k$, and $\sum_{i=1}^{k} n_i c_i = 0$, includes the monotone contrasts and all types of pairwise mean comparisons: $\mu_j - \mu_i, 1 \leq i < j \leq k$. The confidence lower bounds for nonnegative contrasts by Marcus (1978) were given by

$$\sum_{i=1}^{k} n_i c_i \bar{Y}_i - \sigma \tilde{d}_\alpha (\sum_{i=1}^{k} n_i c_i^2)^{1/2}$$

where $\tilde{d}_\alpha$ is the positive square root of the critical value for $\bar{\chi}_{12}^2$ (see Section 2.2). With the simple order restriction on treatment means, the lower bound for the nonnegative contrast $\sum_{i=1}^{k} n_i c_i^* \mu_i$ can be improved to

$$\max\{\sum_{i=1}^{k} n_i c_i \bar{Y}_i - \sigma \tilde{d}_\alpha (\sum_{i=1}^{k} n_i c_i^2)^{1/2}\} \tag{2.7}$$

subject to $\sum_{i=1}^{k} n_i c_i \mu_i \leq \sum_{i=1}^{k} n_i c_i^* \mu_i, \sum_{i=1}^{k} n_i c_i = 0, \sum_{i=j}^{k} n_i c_i \geq 0, j = 2, .., k$, and $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_k$.

When the treatment means are monotone, the ordered pairwise mean comparison $\mu_j - \mu_i, 1 \leq i < j \leq k$, is of particular interest. It can be used to determine whether μ_j is larger than μ_i. For the equal sample size case, Hayter (1990) proposed an efficient one-sided simultaneous confidence lower bound for $\mu_j - \mu_i$ as

$$\bar{Y}_j - \bar{Y}_i - s h_{k,\alpha,\nu} / \sqrt{n}, \tag{2.8}$$

where $h_{k,\alpha,\nu}$ is defined by

$$P_\mu\{\max_{1\le i<j\le k}(\bar{Y}_j - \bar{Y}_i)/(s/\sqrt{n}) \ge h_{k,\alpha,\nu}\} = \alpha$$

when $\mu_1 = \mu_2 = \cdots = \mu_k$. Some critical values $h_{k,\alpha,\nu}$ were given by Hayter (1990). If σ is known, s is replaced by σ and $h_{k,\alpha,\infty}$ is used. Furthermore, the one-sided lower bound for nonnegative contrasts $\sum_{i=1}^k c_i\mu_i$ can be formulated as

$$\sum_{i=1}^k c_i\bar{Y}_i - \sum_{i=1}^k |c_i| s h_{k,\alpha,\nu}/(2\sqrt{n}).$$

Hayter(1992) generalizes the above lower bounds in (2.8) to the unequal sample size cases for three ordered normal means. By the similar discussion as Marcus and Peritz (1976), and Marcus (1978), incorporating the simple order restriction on μ_i improves the lower bounds for pairwise comparisons and nonnegative contrasts.

Korn (1982) studied confidence bands for monotone dose-response curves. With the assumption that the response means are monotone nondecreasing, the $100(1-\alpha)\%$ simultaneous confidence intervals for μ_i's were given by

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

where $m_{k,\nu}$ is the upper α point of the studentized maximum modulus distribution with parameters k and ν (Miller 1981). Both Schoenfeld (1986) and Lee (1996) sought confidence intervals for each individual mean μ_i by incorporating the monotonicity of the response means. The generalized studentized maximum modulus procedure by Lee (1996) gains much over the Scheffé-type procedure by Schoenfeld (1986) and the studentized maximum modulus by Korn (1982).

2.4 Kuhn-Tucker Conditions

The evaluation of the improved simultaneous confidence lower bounds such as in (2.6) and (2.7) is a maximization problem subject to a mixture of equality and inequality constraints. Particularly, let $\mathbf{x}$ be an $n \times 1$ vector and $H(\mathbf{x})$ be an $m \times 1$ vector whose components $h_1(\mathbf{x}), .., h_m(\mathbf{x})$ are differentiable concave functions for $\mathbf{x} \geq 0$. Let $g(\mathbf{x})$ be a differentiable concave function for $\mathbf{x} \geq 0$ as well. The Kuhn-Tucker equivalence theorem will enable us to find an $\mathbf{x}^o$ that maximizes $g(\mathbf{x})$ constrained by $H(\mathbf{x}) \geq 0$ and $\mathbf{x} \geq 0$. A vector $\mathbf{x}$ is said to be feasible if $\mathbf{x}$ satisfies all the constraints. 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 $\phi(\mathbf{x}, \mathbf{u}) = g(\mathbf{x}) + \mathbf{u}'H(\mathbf{x})$. Let $[\frac{\partial \phi}{\partial x_i}]^o$ and $[\frac{\partial \phi}{\partial u_j}]^o$ denote the partial derivatives evaluated at a particular point $\mathbf{x}^o$ and $\mathbf{u}^o$, respectively.

Theorem 2.4.1 ***(Equivalence theorem)*** *Let $h_1(\mathbf{x}), ..., h_m(\mathbf{x}), g(\mathbf{x})$ be concave as well as differentiable for $\mathbf{x} \geq 0$. Let $\phi(\mathbf{x}, \mathbf{u}) = g(\mathbf{x}) + \mathbf{u}'H(\mathbf{x})$. Then $\mathbf{x}^o$ is a solution that maximizes $g(\mathbf{x})$ constrained by $H(\mathbf{x}) \geq 0$ and $\mathbf{x} \geq 0$ if and only if $\mathbf{x}^o$ and some $\mathbf{u}^o$ satisfy the following conditions:*

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

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

(Theorem 3 Kuhn-Tucker 1951)

Simple modifications are admitted when the constraints $H(\mathbf{x}) \geq 0, \mathbf{x} \geq 0$ are changed to the following three cases:

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

Here, using $\phi(\mathbf{x}, \mathbf{u}) = g(\mathbf{x}) + \mathbf{u}'H(\mathbf{x})$ defined for all $\mathbf{x}$ and constrained only by $\mathbf{u} \geq 0$, one must replace condition (1) by

(1)* $$[\frac{\partial \phi}{\partial x_i}]^o = 0$$

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

Here, using $\phi(\mathbf{x}, \mathbf{u}) = g(\mathbf{x}) + \mathbf{u}'H(\mathbf{x})$ defined for all $\mathbf{u}$ and constrained only by $\mathbf{x} \geq 0$, one must replace condition (2) by

(2)* $$[\frac{\partial \phi}{\partial u_j}]^o = 0$$

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

Here, using $\phi(\mathbf{x}, \mathbf{u}) = g(\mathbf{x}) + \mathbf{u}'H(\mathbf{x})$ defined for all $\mathbf{x}$ and $\mathbf{u}$ without constraints, one must replace conditions (1) and (2) by (1*) and (2*). This corresponds to the customary use of the method of Lagrange multipliers.

Chapter 3

Max-Min Multiple Comparison Procedure

The effects of a drug or a toxin are estimated by an experiment in which increasing doses $x_1, x_2, \ldots, x_k$ are given to k groups of animals and the response Y_{ij} of the jth animal in the ith group is observed. It is frequently of interest to use simultaneous confidence intervals for pairwise differences of dose-response means to assess the significance of dose levels. If a parametric family of dose-response curves is hypothesized, then the parameters and the curve can be estimated from the data using a nonlinear regression. A confidence region calculated for these parameters yields confidence bands for pairwise comparisons of the dose-response curves in a straightforward manner. But in most environmental toxicology applications, the response at lower doses is of interest and no parametric dose-response model is assumed to hold in general. In these applications, the response mean μ_i can be estimated by the sample mean $\bar{Y}_i$ at various doses. Assuming normality of the response data, simultaneous confidence intervals for pairwise mean differences can be constructed using the studentized range technique. The simultaneous confidence interval estimation procedures for successive comparisons of ordered treatment effects were studied by Lee and Spurrier (1995) and Liu, Miwa and Hayter (2000).

In this chapter we propose a max-min technique to compare pairwise mean differences. The procedure given in Section 3.1 is a modification of the studentized range technique and it can be used when the dose-response curve is isotonic. Our max-min multiple comparison procedure is an improvement over Tukey's technique since our technique utilizes the prior knowledge of monotonicity. This improvement can be found in an example given in Section 3.2 and its expected gains are given in Section 3.3. A discussion is presented in Section 3.4.

3.1 Model-Free Confidence Intervals

3.1.1 Max-Min Simultaneous Confidence Intervals

The dose-response curve $y = f(x)$ is to be estimated from k independent samples $Y_{i1}, Y_{i2}, \ldots, Y_{in}$ taken at increasing dose level $x_i, i = 1, 2, \ldots, k$. The Y_{ij} are independent normal random variables with mean $\mu_i = f(x_i)$ and with an equal unknown variance σ^2. If a parametric model for $f(x)$ is not hypothesized, then $f(x_i)$ can be estimated by the means $\bar{Y}_i = \sum_{j=1}^{n} Y_{ij}/n$ of the responses at the dose levels x_i. The usual model-free approach to form the $100(1-\alpha)\%$ simultaneous confidence interval for the pairwise mean differences $\mu_j - \mu_i$ is given by

$$\bar{Y}_j - \bar{Y}_i - q^{\alpha}_{k,\nu}\frac{s}{\sqrt{n}} \leq \mu_j - \mu_i \leq \bar{Y}_j - \bar{Y}_i + q^{\alpha}_{k,\nu}\frac{s}{\sqrt{n}} \tag{3.1}$$

where $\nu = k(n-1)$, $s^2 = \sum_{i,j}(Y_{ij} - \bar{Y}_i)^2/\nu$ and $q^{\alpha}_{k,\nu}$ is the upper 100α percentage point of the studentized range test with parameters k and ν (see Miller 1981).

If the dose-response curve $f(x)$ is known to be monotone nondecreasing, then the isotonic regression offers natural estimates of the $\mu_i = f(x_i)$ and it can be computed from the sample mean $\bar{Y}_i$ by the pool-adjacent-violators algorithm (see Section 2.1). Under the assumption that the regression function is monotone nondecreasing, for any $l \leq j, l' \geq i, m \geq j, m' \leq i$, we have that

$$\mu_l - \mu_{l'} \leq \mu_j - \mu_i \leq \mu_m - \mu_{m'}.$$

Note that it is possible that $l' \leq l$. Therefore, $\mu_j - \mu_i$ will be bounded from below by the lower confidence bound for $\mu_l - \mu_{l'}$ and from above by the upper confidence bound for $\mu_m - \mu_{m'}$. One may have another set of confidence interval

$$\bar{Y}_l - \bar{Y}_{l'} - q^{\alpha}_{k,\nu}\frac{s}{\sqrt{n}} \leq \mu_j - \mu_i \leq \bar{Y}_m - \bar{Y}_{m'} + q^{\alpha}_{k,\nu}\frac{s}{\sqrt{n}}.$$

When $f(x)$ is known to be nondecreasing, the following $100(1-\alpha)\%$ simultaneous improved confidence intervals are proposed:

$$\max_{l \leq j, l' \geq i}(\bar{Y}_l - \bar{Y}_{l'} - q^{\alpha}_{k,\nu}\frac{s}{\sqrt{n}}) \leq \mu_j - \mu_i \leq \min_{m \geq j, m' \leq i}(\bar{Y}_m - \bar{Y}_{m'} + q^{\alpha}_{k,\nu}\frac{s}{\sqrt{n}}). \quad (3.2)$$

These simultaneous confidence intervals are not derived from the estimated isotonic regression. They are derived from the sample means by utilizing the monotone assumption on $f(x)$. We have just shown that any nondecreasing sequence μ_i satisfying (3.1) will satisfy (3.2). On the other hand, it is obvious that the nondecreasing sequence μ_i satisfying (3.2) will automatically satisfy (3.1). Thus, the simultaneous confidence intervals (3.2) for pairwise differences of the true dose-response means have an exact $1 - \alpha$ coverage probability. The above modified procedure applies when $f(x)$ is known to be monotone nondecreasing. A computation procedure to find the lower bounds and the upper bounds of (3.2) will be illustrated in the next section.

Utilizing the one-sided studentized range test, Hayter (1990) constructed a one-sided $100(1-\alpha)\%$ simultaneous confidence lower bound for $\mu_j - \mu_i, j > i$. By the similar discussion as Marcus (1982), a conservative $100(1-\alpha)\%$ two-sided confidence interval can be obtained as follows:

$$\bar{Y}_j - \bar{Y}_i - h_{k,\nu}^{\alpha/2}\frac{s}{\sqrt{n}} \leq \mu_j - \mu_i \leq \bar{Y}_j - \bar{Y}_i + h_{k,\nu}^{\alpha/2}\frac{s}{\sqrt{n}},$$

where the critical value $h_{k,\nu}^{\alpha/2}$ for one-sided studentized range test statistic was tabled by Hayter(1990) for $\alpha = .10, .05, .01$. The improved confidence interval under the assumption of monotonicity is

$$\max_{l \leq j, l' \geq i}(\bar{Y}_l - \bar{Y}_{l'} - h_{k,\nu}^{\alpha/2}\frac{s}{\sqrt{n}}) \leq \mu_j - \mu_i \leq \min_{m \geq j, m' \leq i}(\bar{Y}_m - \bar{Y}_{m'} + h_{k,\nu}^{\alpha/2}\frac{s}{\sqrt{n}}). \quad (3.3)$$

For a fixed α, we can see that $q_{k,\nu}^{\alpha} < h_{k,\nu}^{\alpha/2}$ for any positive integers k and ν. Therefore, the max-min confidence interval (3.2) by studentized range test is always shorter than the confidence interval (3.3) by one-sided studentized range test.

3.1.2 Unequal Sample Size Case

Let $\bar{Y}_i$ be the sample mean of n_i observations on the ith dose level. A modification of Tukey's simultaneous confidence intervals can be obtained from the conservative property of the Tukey-Kramer multiple comparison procedure. Hayter (1984) showed that if n_i are unequal, simultaneous confidence intervals (3.1) can be modified by replacing $\frac{1}{\sqrt{n}}$ by $\{\frac{1}{2}(\frac{1}{n_j}+\frac{1}{n_i})\}^{\frac{1}{2}}$ in the confidence interval for $\mu_j - \mu_i$, and the coverage probability is at least $1-\alpha$, that is

$$P\{|(\bar{Y}_j - \bar{Y}_i) - (\mu_j - \mu_i)| \le q^{\alpha}_{k,\nu} s\sqrt{\frac{1}{2}(\frac{1}{n_j}+\frac{1}{n_i})}, \quad \text{for all} \quad 1 \le i,j \le k\} \ge 1-\alpha,$$

with the degrees of freedom $\nu = \sum_{i=1}^{k} n_i - k > 0$. The $100(1-\alpha)\%$ simultaneous confidence intervals for $\mu_j - \mu_i$ are given by

$$\bar{Y}_j - \bar{Y}_i - q^{\alpha}_{k,\nu} s\sqrt{\frac{1}{2}(\frac{1}{n_j}+\frac{1}{n_i})} \le \mu_j - \mu_i \le \bar{Y}_j - \bar{Y}_i + q^{\alpha}_{k,\nu} s\sqrt{\frac{1}{2}(\frac{1}{n_j}+\frac{1}{n_i})}. \quad (3.4)$$

If $f(x)$ is monotone nondecreasing , the $100(1-\alpha)\%$ max-min simultaneous confidence intervals for $\mu_j - \mu_i, 1 \le i < j \le k$ are

$$\max_{l \le j, l' \ge i}\{\bar{Y}_l - \bar{Y}_{l'} - q^{\alpha}_{k,\nu} s\sqrt{\frac{1}{2}(\frac{1}{n_l}+\frac{1}{n_{l'}})}\} \le \mu_j - \mu_i \le \min_{m \ge j, m' \le i}\{\bar{Y}_m - \bar{Y}_{m'} + q^{\alpha}_{k,\nu} s\sqrt{\frac{1}{2}(\frac{1}{n_m}+\frac{1}{n_{m'}})}\}. \quad (3.5)$$

The simultaneous confidence intervals (3.4) and (3.5) are analogues of (3.1) and (3.2), respectively, when sample sizes are unequal. They are conservative because their coverage probability is at least $1-\alpha$. A FORTRAN program for computing the max-min simultaneous confidence interval (3.5) is given in the Appendices.

Hayter (1984) also noted that if interest is restricted to pairwise comparisons of the means, the Tukey-Kramer procedure (3.4) will provide shorter intervals than Scheffé's procedure and the classical Bonferroni's procedure. Therefore, the max-min simultaneous confidence interval procedure is good in comparing pairwise means under the monotone assumption.

3.2 A Numerical Example

For illustration, we consider the data, given in Table 3.1, from a binding inhibition assay which was described fully by Kanowith-Klein, Vitetta, Korn and Ashman (1979). For each dilution of antiserum, the number of rosettes formed was counted and compared to the number of rosettes formed with no antiserum present. The analysis here proceeds conditionally on the numbers of rosettes formed with no antiserum present. The percentage inhibitions can be taken to be statistically independent (see Korn 1982). In this set of data, there are $k = 9$ different dilutions of one antiserum.

For the 24 observations in Table 3.1, the pooled estimate of the variance, s^2, is 86.48 with 15 degrees of freedom. The 90% Tukey's simultaneous confidence intervals of $\mu_j - \mu_i, 1 \leq i < j \leq 9$, calculated according to (3.4), are provided in Table 3.2 with the upper percentage point $q_{9,15}^{0.10} = 4.52$. If $l' < l$, the Tukey's lower bound can be found in the bottom half of the table, whereas if $l' > l$ the values are the negatives of the top half of the table with the indices l' and l interchanged. The 90% max-min simultaneous confidence intervals can be computed using the values in Table 3.2. To compute the max-min lower bound, we change the sign on each value of the upper bounds in Table 3.2. For example, in row 5 and column 6, the value 17.94 is Tukey's upper bound for $\mu_6 - \mu_5$; hence -17.94 is Tukey's lower bound for $\mu_5 - \mu_6$. The max-min lower bound for $\mu_5 - \mu_4$ is the maximum of the values of the first 5 rows in columns 4 to 9. That value is -17.94 which is Tukey's lower bound for $\mu_5 - \mu_6$. The 90% max-min simultaneous confidence intervals for $\mu_j - \mu_i, 1 \leq i < j \leq 9$, calculated according to (3.5), are provided in Table 3.3. The notation -17.94^* is used in Table 3.3 to indicate that the max-min lower bound for $\mu_5 - \mu_4$ is zero from our prior knowledge and the value -17.94 indicates the lower bound computed by (3.5). The max-min upper bound for $\mu_j - \mu_i$ is the minimum of the values of the first i rows in columns j to 9. The max-min upper bound for $\mu_5 - \mu_4$ is 20.80, which is Tukey's upper bound of $\mu_6 - \mu_4$.

In general, by utilizing the prior knowledge of order relationship on μ_i, Tukey's simultaneous lower bound and upper bound can be improved by the max-min technique. For example, the 90% max-min simultaneous confidence interval of $\mu_9 - \mu_2$ is (0.48, 55.22); however, the corresponding Tukey's simultaneous confidence interval is (-4.22, 55.22). One may not conclude a significant difference between level 2 and level 9 using Tukey's procedure but this difference can be detected by the proposed max-min procedure. Comparing the max-min confidence intervals (Table 3.3) with Tukey's confidence intervals (Table 3.2), 25 of the 36 lower bounds had considerable improvements, as did 14 of the 36 upper bounds.

The confidence interval (3.3) obtained by Hayter's one-sided studentized range test can also be generalized to the unequal sample size case. The critical value is replaced by $h^{0.05}_{9,15,\mathbf{n}} = 4.68$, with the sample size $\mathbf{n} = (2,2,4,2,3,3,2,4,2)$. Comparing this to the critical value $q^{0.10}_{9,15} = 4.52$, we realize that the two-sided confidence bound constructed by one-sided test is less efficient than the one obtained by two-sided test.

For this example, the critical values (coefficients of the pooled variance $s\sqrt{\frac{1}{n_i} + \frac{1}{n_j}}$) of Tukey's, Scheffé's and Bonferroni's procedure are $q^{0.10}_{9,15}/\sqrt{2} =$

3.20, $\sqrt{8F^{0.10}_{8,15}} = 4.12$ and $t^{0.0014}_{15} = 3.57$, respectively. Tukey's procedure yields the shortest confidence intervals for $\mu_j - \mu_i, j > i$, therefore, since the max-min procedure is an improvement over Tukey's procedure, the max-min simultaneous confidence intervals are effective for pairwise comparisons of the means.

3.3 Expected Gains of Max-Min Confidence Bounds

We shall consider the equal sample size case. The results for unequal sample size case follow similarly. The $100(1-\alpha)\%$ Tukey's simultaneous confidence intervals for $\mu_j - \mu_i, i \leq j$, are

$$(\bar{Y}_j - \bar{Y}_i - q^{\alpha}_{k,\nu}\frac{s}{\sqrt{n}}, \quad \bar{Y}_j - \bar{Y}_i + q^{\alpha}_{k,\nu}\frac{s}{\sqrt{n}}).$$

The expected lower and upper bounds are

$$\mu_j - \mu_i - q^{\alpha}_{k,\nu}\frac{E(s)}{\sqrt{n}},$$

and

$$\mu_j - \mu_i + q^{\alpha}_{k,\nu}\frac{E(s)}{\sqrt{n}},$$

respectively.

Let

$$L_{ij} = \max_{\beta \leq j, \alpha \geq i}(\bar{Y}_\beta - \bar{Y}_\alpha - q^{\alpha}_{k,\nu}\frac{s}{\sqrt{n}})$$

and

$$U_{ij} = \min_{\alpha \leq i, \beta \geq j}(\bar{Y}_\beta - \bar{Y}_\alpha + q^{\alpha}_{k,\nu}\frac{s}{\sqrt{n}})$$

be the max-min lower bound and the max-min upper bound for (3.2). The expected max-min lower bound is

$$\begin{aligned} E(L_{ij}) &= E\{\max_{\beta \leq j, \alpha \geq i}(\bar{Y}_\beta - \bar{Y}_\alpha)\} - q^\alpha_{k,\nu}\frac{E(s)}{\sqrt{n}} \\ &= E(\max_{\beta \leq j}\bar{Y}_\beta) - E(\min_{\alpha \geq i}\bar{Y}_\alpha) - q^\alpha_{k,\nu}\frac{E(s)}{\sqrt{n}}. \end{aligned}$$

The expected gain, denoted by $g_{ij}(L)$, of L_{ij} over Tukey's lower bound is

$$g_{ij}(L) = E\{\max_{\beta \leq j}(\bar{Y}_\beta - \mu_j)\} - E\{\min_{\alpha \geq i}(\bar{Y}_\alpha - \mu_i)\}. \quad (3.6)$$

The distributions of $\max_{\beta \leq j}(\bar{Y}_\beta - \mu_j)$ and $\min_{\alpha \geq i}(\bar{Y}_\alpha - \mu_i)$ can be obtained in a straightforward manner, but the computations of their expected values are very complicated. Since the gain is nonnegative, the expected gain is always nonnegative. Similarly, we obtain the expected gain, denoted by $g_{ij}(U)$, of U_{ij} over Tukey's upper bound as

$$g_{ij}(U) = E\{\max_{\alpha \leq i}(\bar{Y}_\alpha - \mu_i)\} - E\{\min_{\beta \geq j}(\bar{Y}_\beta - \mu_j)\}. \quad (3.7)$$

The gains $g_{ij}(L)$ and $g_{ij}(U)$ are illustrated by the regression curve $\mu_i = f(x_i)$ with $\mu_1 = \cdots = \mu_t = \mu$ and $\mu_{t+1} = \cdots = \mu_k = \mu + \delta$. We restrict our study to pairwise comparisons of $\mu_j - \mu_i$ with $i \leq t$ and $j > t$. Without loss of generality, we may assume that $\sigma/\sqrt{n} = 1$. The expected gain $g_{ij}(U)$ in (3.7) becomes

$$g_{ij}(U) = E(Z_{i:i}) - E(Z_{1:k-j+1}),$$

where $Z_{i:n}$ is the ith smallest order statistic in a random sample of size n from $N(0, 1)$. The exact expected gain $g_{ij}(L)$ is difficult to compute. However, its bound can be obtained as follows. We have that

$$E\{\max_{\beta\leq j}(\bar{Y}_\beta - \mu_j)\} \geq E\{\max_{t<\beta\leq j}(\bar{Y}_\beta - \mu_j)\} = E(Z_{j-t:j-t})$$

and

$$E\{\min_{\alpha\geq i}(\bar{Y}_\alpha - \mu_i)\} \leq E\{\min_{t\geq\alpha\geq i}(\bar{Y}_\alpha - \mu_i)\} = E(Z_{1:t-i+1}).$$

Therefore,

$$g_{ij}(L) \geq E(Z_{j-t:j-t}) - E(Z_{1:t-i+1}). \tag{3.8}$$

The lower bounds of (3.8) are given in Table 3.4 for the case of $k = 9, t = 3$. They can be computed using the mean of normal order statistics (see Arnold, Balakrishnan and Nagaraja 1992). The largest lower bound for $g_{ij}(L)$ is $\mu_9 - \mu_1$ and the smallest lower bound for $g_{ij}(L)$ is zero, located at $g_{34}(L)$. The further apart the indices i and j are from the mean change point t, the larger the gain. The expected gain, $g_{ij}(U)$, of U_{ij} over Tukey's upper bound can also be found in Table 3.4 by replacing i and j by $4-i$ and $13-j$, respectively. On the other hand, the expected max-min confidence lower bound can be rewritten as

$$E(L_{ij}) = \delta - \{q^\alpha_{k,\nu}\frac{E(s)}{\sqrt{n}} - g_{ij}(L)\}.$$

As $\nu s^2/\sigma^2$ has a Chi-square distribution with ν degrees of freedom, we have

$$\begin{aligned} E(s) &= \frac{\sigma}{\sqrt{\nu}}\int_0^\infty \sqrt{w}\frac{1}{2^{\nu/2}\Gamma(\frac{\nu}{2})}w^{\frac{\nu}{2}-1}e^{-\frac{w}{2}}dw \\ &= \sigma\sqrt{2}\Gamma(\frac{\nu+1}{2})/\sqrt{\nu}\Gamma(\frac{\nu}{2})\int_0^\infty \frac{1}{2^{\frac{\nu+1}{2}}\Gamma(\frac{\nu+1}{2})}w^{\frac{\nu+1}{2}-1}e^{-\frac{w}{2}}dw \\ &= \sigma\sqrt{2}\Gamma(\frac{\nu+1}{2})/\sqrt{\nu}\Gamma(\frac{\nu}{2}). \end{aligned}$$

Whenever the size of the shift δ exceeds $q^{\alpha}_{k,\nu}\frac{E(s)}{\sqrt{n}} - g_{ij}(L)$, one would expect to detect a change in mean around t by the max-min simultaneous confidence lower bounds.

3.4 Discussion

The proposed modification of Tukey's studentized range technique is a simple and effective method to construct simultaneous confidence intervals for pairwise differences in monotone dose-response curves. As suggested by Scheffé (1953), if we are interested exclusively in the difference $\mu_j - \mu_i, j \neq i$, when all $\bar{Y}_i$ have the same variance and all pairs $\bar{Y}_j - \bar{Y}_i$ have the same covariance, Tukey's method will yield shorter simultaneous confidence intervals. Hence, the max-min simultaneous confidence intervals can be applied specifically for pairwise mean differences under the monotonic assumption. For the equal sample size case, the max-min simultaneous confidence intervals have an exact $1 - \alpha$ coverage probability.

The max-min simultaneous confidence bounds can also be used to detect the range of the change point for normal variables. This approach is effective for detecting $\mu_j - \mu_i$ when i and j are not adjacent indices.

Table 3.1: Inhibition of Rosette Formation

Level	Log_{10} dilution	Percentage inhibition
1	3.519	-12 , 5
2	3.114	12, 27
3	2.778	14, 18, 25, 36
4	2.399	44, 46
5	2.000	44, 45, 46
6	1.399	27, 33, 56
7	1.000	38, 40
8	0.699	32, 43, 50, 54
9	0.301	43, 47

Table 3.2: 90% Tukey's Simultaneous Confidence Intervals for $\mu_j - \mu_i, j > i$

	Upper bound									
	1	2	3	4	5	6	7	8	9	j/i
1		52.72	52.49	78.22	75.63	69.30	72.22	73.99	78.22	1
2	-6.72		29.49	55.22	52.63	46.30	49.22	50.99	55.22	2
3	1.01	-21.99		47.49	44.45	38.12	41.49	42.52	47.49	3
4	18.78	-4.22	-3.99		27.13	20.80	23.72	25.49	29.72	4
5	21.37	-1.63	-0.95	-27.13		17.94	21.13	22.45	27.13	5
6	15.03	-7.97	-7.28	-33.47	-30.60		27.47	28.78	33.47	6
7	12.78	-10.22	-9.99	-35.72	-33.13	-26.80		31.49	35.72	7
8	22.51	-0.49	0.48	-25.99	-22.95	-16.62	-19.99		25.99	8
9	18.78	-4.22	-3.99	-29.72	-27.13	-20.80	-23.72	-25.49		
j/i	1	2	3	4	5	6	7	8		
	Lower bound									

Table 3.3: 90% Max-Min Simultaneous Confidence Intervals for $\mu_j - \mu_i, j > i$

	Upper bound									
	1	2	3	4	5	6	7	8	9	j/i
1		52.49	52.49	69.30	69.30	69.30	72.22	73.99	78.22	1
2	-6.72*		29.49	46.30	46.30	46.30	49.22	50.99	55.22	2
3	1.01	-21.99*		38.12	38.12	38.12	41.49	42.52	47.49	3
4	18.78	-3.99*	-3.99*		20.80	20.80	23.72	25.49	29.72	4
5	21.37	-0.95*	-0.95*	-17.94*		17.94	21.13	22.45	27.13	5
6	21.37	-0.95*	-0.95*	-17.94*	-17.94*		21.13	22.45	27.13	6
7	21.37	-0.95*	-0.95*	-17.94*	-17.94*	-17.94*		22.45	27.13	7
8	22.51	0.48	0.48	-16.62*	-16.62*	-16.62*	-19.99*		25.99	8
9	22.51	0.48	0.48	-16.62*	-16.62*	-16.62*	-19.99*	-22.45*		
j/i	1	2	3	4	5	6	7	8		
	Lower bound									

Table 3.4: Lower Bounds $E(Z_{j-t:j-t}) - E(Z_{1:t-i+1})$ When $t = 3, k = 9$

$i\backslash j$	4	5	6	7	8	9
1	0.85	1.41	1.69	1.88	2.01	2.11
2	0.56	1.13	1.41	1.59	1.73	1.83
3	0.00	0.56	0.85	1.03	1.16	1.27

Chapter 4

Simultaneous Confidence Lower Bounds

The regression curve $y = f(x)$ is to be estimated from the observations $Y_{i1}, Y_{i2}, ..., Y_{in_i}$ collected at the quantitative level $x_i, i = 1, 2, ..., k$. Let Y_{ij} be independent normal variates with means $\mu_i = f(x_i)$ and a common variance σ^2, where μ_i are monotone nondecreasing. We are interested in the one-sided confidence lower bounds for the pairwise comparisons $\mu_j - \mu_i, 1 \leq i < j \leq k$, and nonnegative linear combinations of pairwise comparisons (nonnegative contrasts). The development of simultaneous confidence bounds for restricted settings was first carried out by Bohrer (1967) and Bohrer and Francis (1972). By use of the likelihood ratio statistic, Marcus and Peritz (1976) obtained one-sided simultaneous confidence intervals for monotone contrasts $\sum_{i=1}^{k} n_i c_i \mu_i$, for which $\sum_{i=1}^{k} n_i c_i = 0$ and $c_1 \leq c_2 \leq \cdots \leq c_k$. Their results subsume those of Bohrer and Francis (1972). However, apart from $\mu_k - \mu_1$, none of the ordered pairwise comparisons $\mu_j - \mu_i$ are monotone contrasts. Marcus (1978) studied the confidence lower bounds for the nonnegative contrasts, which include monotone contrasts and pairwise comparisons, when the common variance σ^2 is known.

If several treatment means are to be compared with one another and the experimenter has a reason to believe that the treatment means are simply ordered, then this order assumption can improve confidence bounds. The use of prior knowledge that the regression curve is monotone, $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_k$, to sharpen confidence bounds first appeared in Marcus and Peritz (1976). The technique can also be found in Marcus (1978), Korn (1982), Schoenfeld (1986), Hayter (1990) and Lee (1996). Marcus (1978) studied the improved simultaneous confidence lower bounds for nonnegative contrasts while utilizing prior knowledge of the monotonicity of the means μ_i. This improved lower bound is the solution to an optimization problem of maximizing the simultaneous confidence lower bounds. However, Marcus' results are incomplete.

In this chapter, we improve significantly over the results of Marcus (1978). In Section 4.1, we provide a necessary and sufficient condition for the solution to the optimization problem of maximizing simultaneous confidence lower bounds. An efficient computation algorithm for the improved one-sided confidence lower bounds of pairwise comparisons and nonnegative contrasts is given in Section 4.2. A numerical example illustrating the algorithm is given in Section 4.3. Section 4.4 contains all technical results and a conclusion is included in Section 4.5.

4.1 The Optimization Problem

4.1.1 Simultaneous Confidence Lower Bounds

For the monotone nondecreasing regression means μ_i, the class of monotone contrasts is defined as $\sum_{i=1}^k n_i c_i \mu_i$ where $c_{i-1} \leq c_i, i = 2, ..., k$. The class of nonnegative contrasts is defined by $\sum_{i=1}^k n_i c_i \mu_i = \sum_{i=1}^{k-1} \sum_{j=i+1}^k \lambda_{ij}(\mu_j - \mu_i)$ with $\lambda_{ij} \geq 0$, which is a nonnegative linear combination of $\mu_j - \mu_i, i < j$. The coefficient $c_1, ..., c_k$ can be rewritten as $\mathbf{c} \succeq 0$, where the partial order $\mathbf{c} \preceq \mathbf{c}^*$ is defined by $\sum_{i=j+1}^k n_i c_i \leq \sum_{i=j+1}^k n_i c_i^*, j = 1, ..., k-1$, and $\sum_{i=1}^k n_i c_i = \sum_{i=1}^k n_i c_i^* = 0$. Monotone contrasts are special cases of nonnegative contrasts.

Example 4.1.1 *Let $k = 5, n_i = n$ for $i = 1, ..., 5$. $\mu_5 - (\mu_1 + \mu_2)/2$ is a monotone contrast. However, $\mu_4 - \mu_3$ is a nonnegative contrast but not a monotone contrast.*

As not all pairwise mean differences are monotone contrasts while they are nonnegative contrasts, it will be of considerable interest to construct one-sided simultaneous confidence lower bounds for pairwise comparisons $\mu_j - \mu_i, 1 \leq i < j \leq k$, and nonnegative linear combinations of pairwise comparisons.

A $100(1-\alpha)\%$ one-sided simultaneous confidence bound for the nonnegative contrast $\sum_{i=1}^k n_i c_i \mu_i$ is denoted by

$$l(\sum_{i=1}^k n_i c_i \mu_i) = \sum_{i=1}^k n_i c_i \bar{Y}_i - \tilde{t}_\alpha s(\sum_{i=1}^k n_i c_i^2)^{1/2}, \tag{4.1}$$

where $\bar{Y}_i = \sum_{j=1}^{n_i} Y_{ij}/n_i, s^2 = \sum_{i,j}(Y_{ij} - \bar{Y}_i)^2/\nu$ with $\nu = \sum_{i=1}^k n_i - k > 0$, and $\tilde{t}_\alpha$ will be given below. Marcus (1978) studied the case when σ is known, and some of the critical values $\tilde{t}_\alpha$ can be found there. As a special case, the $100(1-\alpha)\%$ one-sided simultaneous confidence lower bound for $\mu_j - \mu_i$ is

$$l(\mu_j - \mu_i) = \bar{Y}_j - \bar{Y}_i - \tilde{t}_\alpha s(n_j^{-1} + n_i^{-1})^{1/2}. \tag{4.2}$$

4.1.2 The Critical Value $\tilde{t}_\alpha$

The critical value $\tilde{t}_\alpha$ is the solution to the equation:

$$P_\mu\{\sum_{i=1}^k n_i c_i \mu_i \geq \sum_{i=1}^k n_i c_i \bar{Y}_i - \tilde{t}_\alpha s(\sum_{i=1}^k n_i c_i^2)^{1/2}, \forall \mathbf{c} \succeq 0\} = 1 - \alpha.$$

The left-hand side can be rewritten as

$$\begin{aligned} & P_\mu\{\max_{\mathbf{c}\in N} \sum_{i=1}^k n_i c_i(\bar{Y}_i - \mu_i)/s(\sum_{i=1}^k n_i c_i^2)^{1/2} \leq \tilde{t}_\alpha\} \\ = \; & P_{\mu=0}\{\max_{\mathbf{c}\in N} \sum_{i=1}^k n_i c_i \bar{Y}_i / s(\sum_{i=1}^k n_i c_i^2)^{1/2} \leq \tilde{t}_\alpha\} \\ = \; & P_{\mu=0}\{\sum_{i=1}^k n_i \mu_i^{o2}/s^2 \leq \tilde{t}_\alpha^2\} \end{aligned}$$

and the last identity follows a similar argument as in Hogg (1965) where $\mu^o = (\mu_1^o, ..., \mu_k^o)$ is the weighted least square projection of $(\bar{Y}_1, ..., \bar{Y}_k)$ onto $N = \{\mathbf{c} : \mathbf{c} \succeq 0, \sum_{i=1}^k n_i c_i = 0\}$ with weights $n_1, n_2, ..., n_k$. The statistic $S_{01}^o = \sum_{i=1}^k n_i \mu_i^{o2}/s^2$ has the same distribution as the statistic S_{12} in (2.4) in Section 2.2 when $\mu = 0$ and its critical value $\tilde{t}_\alpha^2$ can be found in Table A.7 of Robertson, Wright and Dykstra (1988).

4.1.3 The Optimization Problem

The monotone nondecreasing property of regression curves can be used to improve the confidence lower bound for $\sum_{i=1}^k n_i c_i^* \mu_i$. If $\mu_6 - \mu_1 \geq \mu_5 - \mu_1$, then the simultaneous confidence lower bound for $\mu_6 - \mu_1$ is bounded from below by that for $\mu_5 - \mu_1$. By Abel's method of summation, $\sum_{i=1}^k n_i c_i \mu_i \leq \sum_{i=1}^k n_i c_i^* \mu_i$ for all $\mu = (\mu_1, ..., \mu_k), \mu_1 \leq \mu_2 \leq \cdots \leq \mu_k$, if and only if $\mathbf{c} \preceq \mathbf{c}^*$. The improved confidence lower bound for $\sum_{i=1}^k n_i c_i \mu_i^*$ is denoted by

$$L(\sum_{i=1}^k n_i c_i^* \mu_i) = \max_{0 \preceq \mathbf{c} \preceq \mathbf{c}^*} l(\sum_{i=1}^k n_i c_i \mu_i). \qquad (4.3)$$

It can be shown that $S_{01}^o > \tilde{t}_\alpha^2$ if and only if there exists a $\mathbf{c}, \mathbf{c} \succeq 0$, such that $l(\sum_{i=1}^k n_i c_i \mu_i) > 0$. The latter indicates that there are differences between the dose levels, in particular, $\mu_k - \mu_1 > 0$. In this chapter, we shall restrict our attention to the case $l(\sum_{i=1}^k n_i c_i \mu_i) > 0$ for some $\mathbf{c} \succeq 0$. The following theorem establishes a necessary and sufficient condition for an optimal solution to (4.3) and its proof is found in Section 4.4.

Theorem 4.1.1 *Given a contrast* $\sum_{i=1}^k n_i c_i^* \mu_i$ *where* $\mu_1 \leq \cdots \leq \mu_k$, *let* $a_j^* = \sum_{i=j+1}^k n_i c_i^*, j = 1, ..., k-1$, *and let* $Z = \{j < k : a_j^* = 0\}$. *Let* $\mathbf{c}^o$ *be such that* $0 \preceq \mathbf{c}^o \preceq \mathbf{c}^*$, *let* $a_j^o = \sum_{i=j+1}^k n_i c_i^o$, *let* $R = \{j < k : a_j^o = a_j^* > 0\}$ *and* $S = \{j < k : a_j^o = 0, a_j^* > 0\}$. *Let* p, q *and* r *be three consecutive indices in* $R \cup S \cup Z \cup \{0, k, k+1\} (q \neq 0, k+1)$. *Let* $\tilde{n}_{pq} = \sum_{j=p+1}^q n_j, \tilde{c}_{pq} = (a_p^o - a_q^o)/\tilde{n}_{pq}$ *and* $\tilde{Y}_{pq} = \sum_{j=p+1}^q n_j \bar{Y}_j / \tilde{n}_{pq}$. *The* $\mathbf{c}^o$ *maximizes* $l(\sum_{i=1}^k n_i c_i \mu_i) = \sum_{i=1}^k n_i c_i \bar{Y}_i - \tilde{t}_\alpha s (\sum_{i=1}^k n_i c_i^2)^{1/2}$ *subject to* $0 \preceq \mathbf{c} \preceq \mathbf{c}^*$ *if and only if*

$$c_j^o = \tilde{c}_{pq} + b(\bar{Y}_j - \tilde{Y}_{pq}), \quad \text{if } p < j \leq q, \qquad (4.4)$$

$$\tilde{c}_{qr} - \tilde{c}_{pq} \leq b(\tilde{Y}_{qr} - \tilde{Y}_{pq}), \quad \text{if } q \in R, \qquad (4.5)$$

$$\tilde{c}_{qr} - \tilde{c}_{pq} \geq b(\tilde{Y}_{qr} - \tilde{Y}_{pq}), \quad \text{if } q \in S, \qquad (4.6)$$

$$SSW = \sum_q \sum_{j=p+1}^{q} n_j(\bar{Y}_j - \tilde{Y}_{pq})^2 < \tilde{t}_\alpha^2 s^2, \quad (4.7)$$

where

$$b = \{\sum_q \tilde{n}_{pq}\tilde{c}_{pq}^2/(\tilde{t}_\alpha^2 s^2 - SSW)\}^{1/2}. \quad (4.8)$$

Remark: For the case $SSW = 0$, the optimization problem (4.3) is reduced to minimizing $\sum_{i=1}^{k} n_i c_i^2$ subject to $0 \preceq \mathbf{c} \preceq \mathbf{c}^*$. The optimal solution $\mathbf{c}^o$ is determined by $R \supseteq \{j < k : \bar{Y}_j < \bar{Y}_{j+1}, a_j^* > 0\}$ and $S \supseteq \{j < k : \bar{Y}_j > \bar{Y}_{j+1}$ and $a_j^* > 0\}$. An index j such that $\bar{Y}_j = \bar{Y}_{j+1}$ and $a_j^* > 0$ shall have the value $c_j^o = c_{j+1}^o$ if $0 \preceq \mathbf{c}^o \preceq \mathbf{c}^*$.

From the discussion following (4.9) in Section 4.4, it suffices to consider the case that $Z = \emptyset$ in the remainder of this chapter. Marcus (1978) proposed a method to compute the solution for a particular partition R, S and $T = \{1, ..., k-1\} - (R \cup S)$. Part of the results of Lemma 4.4.2 in Section 4.4, including (4.10), (4.13) and (4.14), were given by Marcus (1978). The formulas (4.4), (4.7) and (4.8) in Theorem 4.1.1 are respectively their simplifications. However, which partition yields the optimal solution is unresolved by Marcus (1978). Theorem 4.1.1 provides a necessary and sufficient condition for the optimal solution. These are the two significant improvements over those of Marcus (1978). Furthermore, we make another improvement by providing an efficient computation algorithm as below.

4.2 Stepwise Optimal Partition Algorithm

When $L(\sum_{i=1}^{k} n_i c_i \mu_i) > 0$, the feasible partition is the one with nonempty R. It has as many as $3^{k-1} - 2^{k-1}$ choices. This is a very large number even for a moderate k. For example, when $k = 6$, there are 211 feasible partitions. Hence, it is important to have an efficient algorithm. This section provides an efficient algorithm to identify the optimal partitions (R_i, S_i, T_i). For a given $Y_{ij}, i = 1, ..., k, j = 1, ..., n_i$, each partition (R_i, S_i, T_i) is optimal for a different range of confidence level $1 - \alpha$, starting from the lowest level and continuing until a desired level is reached.

Algorithm

In each step, let p, q and r be three consecutive indices in $R_i \cup S_i \cup Z \cup \{0, k, k+1\}$.

(0) Let $a_h^* = \sum_{j=h+1}^{k} n_j c_j^*, h = 1, ..., k-1$. Set $i = 0, R_0 = \{j < k : \bar{Y}_j < \bar{Y}_{j+1}, a_j^* > 0\}, S_0 = \{j < k : \bar{Y}_j > \bar{Y}_{j+1}, a_j^* > 0\}$ and $T_0 = \emptyset$ (for the case that $\bar{Y}_j = \bar{Y}_{j+1}$ for some j, see Remark after Theorem 4.1.1 for the initial partition R_0, S_0 and T_0).

Let $a_q = a_q^*$ if $q \in R_0, a_q = 0$ if $q \in S_0, a_0 = 0$ and $a_k = 0$. Let $\tilde{c}_{pq} = (a_{q-1} - a_q)/n_q, \tilde{Y}_{pq} = \bar{Y}_q, \tilde{n}_{pq} = n_q, q = 1, ..., k$. Let $A_0 = \sum_{q=1}^{k} \tilde{n}_{pq} \tilde{c}_{pq}^2$ and $B_0 = 0$.

(i) Let $\delta_{i1} = \sup\{(\tilde{c}_{qr} - \tilde{c}_{pq})/(\tilde{Y}_{qr} - \tilde{Y}_{pq}) < \delta_{i-1} : \tilde{Y}_{qr} \neq \tilde{Y}_{pq}, q \in R_i \cup S_i\}$ and the restriction of δ_{i-1} applies to $i \geq 1$ only.

(ii) For $i \geq 1$, let

$$\delta_{i2} = \sup\{(a_h^* - a_q - \sum_{j=h+1}^{q} n_j \tilde{c}_{pq})/[\sum_{j=h+1}^{q} n_j (\bar{Y}_j - \tilde{Y}_{pq})] < \delta_{i-1} : h \in T_i, p < h < q\}.$$

(iii) Let $\delta_i = \max\{\delta_{i1}, \delta_{i2}\}$ and $t_i = (B_i + A_i/\delta_i^2)^{1/2}/s$. If $\tilde{t}_\alpha \leq t_i$, the optimal partition is R_i, S_i and T_i. Otherwise, go to Step (iv) if $\delta_i = \delta_{i1}$ and Step (v) if $\delta_i = \delta_{i2}$.

(iv) If the supremum of δ_i is obtained at $q \in R_i$, define $R_{i+1} = R_i - \{q\}, S_{i+1} = S_i$ and $T_{i+1} = T_i \cup \{q\}$. If the supremum of δ_i is obtained at $q \in S_i$, define $R_{i+1} = R_i, S_{i+1} = S_i - \{q\}, T_{i+1} = T_i \cup \{q\}$. Let $\Delta = (\tilde{Y}_{pq} - \tilde{Y}_{qr})^2/(\tilde{n}_{pq}^{-1} + \tilde{n}_{qr}^{-1})$. Let $A_{i+1} = A_i - \delta_i^2\Delta$ and $B_{i+1} = B_i + \Delta$. Replace $i = i + 1$. Go to Step (i).

(v) If the supremum of δ_i is obtained at $h = T_i, p < h < q$, define $R_{i+1} = R_i \cup \{h\}, S_{i+1} = S_i, T_{i+1} = T_i - \{h\}$. Let $a_h = a_h^*, \Delta = (\tilde{Y}_{ph} - \tilde{Y}_{hq})^2/(\tilde{n}_{ph}^{-1} + \tilde{n}_{hq}^{-1}), A_{i+1} = A_i + \delta_i^2\Delta$ and $B_{i+1} = B_i - \Delta$. Replace $i = i + 1$. Go to Step (i).

Remark: For pairwise comparison $\mu_j - \mu_i$, skip Step (ii) and Step (v).

4.3 A Numerical Example

Let $\bar{Y}_1 = -10, \bar{Y}_2 = 2, \bar{Y}_3 = -2, \bar{Y}_4 = 6, \bar{Y}_5 = 10, \bar{Y}_6 = 4$, let $n_1 = n_2 = \cdots = n_6 = n = 10$ and $s/\sqrt{n} = 3.6$. The $100(1-\alpha)\%$ simultaneous confidence lower bound $L(-\mu_1 + 0.35\mu_2 - 0.35\mu_3 + \mu_6)$ can be computed as follows. Here $\mathbf{a}^* = (1, 0.65, 1, 1, 1)'$ and $\mathbf{a} = (a_1, a_2, ..., a_5)'$.

(0) The initial partition is $R_0 = \{1, 3, 4\}, S_0 = \{2, 5\}$ and $T_0 = \emptyset$. We have $\mathbf{a} = (1, 0, 1, 1, 0)', n\tilde{\mathbf{c}} = (-1, 1, -1, 0, 1, 0)', \tilde{Y}_{p-1,p} = \bar{Y}_p, \tilde{n}_{p-1,p} = n_p, p =$

$1, ..., k$. Therefore,

$$\delta_0 = \delta_{01} = \sup\{\frac{1}{60}, \frac{1}{20}, \frac{1}{80}, \frac{1}{40}, \frac{1}{60}\} = 1/20.$$

Since $A_0 = 2/5$ and $B_0 = SSW = 0$, we have $t_0 = 1.11$. The R_0, S_0 and T_0 form the optimal partition for confidence level up to 20.9%.

(1) Since $\delta_0 = 1/20$ is obtained at $q = 2 \in S_0$, define the partition $R_1 = \{1, 3, 4\}, S_1 = \{5\}$ and $T_1 = \{2\}$. We have $\mathbf{a} = (1, 1 - 20b, 1, 1, 0)'$. In this step we have that

$$\delta_1 = \max\{\sup(\frac{1}{100}, 0, \frac{1}{40}, \frac{1}{60}), \sup(\frac{7}{400})\} = 1/40.$$

Since $\Delta = 80, A_1 = 1/5$ and $B_1 = 80$, we have $t_1 = 1.76$. The partition is optimal for confidence level ranging from 20.9% to 52.9%.

(2) Since $\delta_1 = 1/40$ is obtained at $q = 4 \in R_1$, define the partition $R_2 = \{1, 3\}, S_2 = \{5\}$ and $T_2 = \{2, 4\}$. We have $\mathbf{a} = (1, 1 - 20b, 1, \frac{1}{2} + 20b, 0)'$ and

$$\delta_2 = \max\{\sup(\frac{1}{100}, \frac{1}{160}, \frac{1}{80}), \sup(\frac{7}{400})\} = 7/400.$$

Since $\Delta = 80, A_2 = 3/20$ and $B_2 = 160$, we have $t_2 = 2.24$. The partition is optimal for confidence level ranging from 52.9% to 74.7%.

(3) Since $\delta_2 = 7/400$ is obtained at $q = 2 \in T_2$, define $R_3 = \{1, 2, 3\}, S_3 = \{5\}$ and $T_3 = \{4\}$. We have $\mathbf{a} = (1, 0.65, 1, \frac{1}{2} + 20b, 0)'$ and

$$\delta_3 = \max\{\sup(\frac{9}{800}, \frac{17}{2000}), \sup(\frac{1}{80})\} = 1/80.$$

Since $\Delta = 80, A_3 = 349/2000$ and $B_3 = 80$, we have $t_3 = 3.04$. The partition is optimal for confidence level ranging from 74.7% to 94.1%.

(4) Since $\delta_3 = 1/80$ is obtained at $q = 5 \in S_3$, define $R_4 = \{1,2,3\}, S_4 = \emptyset$ and $T_4 = \{4,5\}$. We have $\mathbf{a} = (1, 0.65, 1, \frac{2}{3} + \frac{20}{3}b, \frac{1}{3} - \frac{80}{3}b)'$ and

$$\delta_4 = \max\{\sup(\frac{9}{800}, \frac{41}{5200}), \sup(-\frac{1}{40})\} = 9/800.$$

Since $\Delta = 320/3, A_4 = 947/6000$ and $B_4 = 560/3$, we have $t_4 = 3.33$. The partition is optimal for confidence level ranging from 94.1% to 96.9%.

(5) Since $\delta_4 = 9/800$ is obtained at $q = 1 \in R_4$, define $R_5 = \{2,3\}, S_5 = \emptyset$ and $T_5 = \{1,4,5\}$. We have $t_5 = 3.91$. The partition is optimal for confidence level ranging from 96.9% to 99.3%.

(6) Since $\delta_5 = 41/5200$ is obtained at $q = 3 \in R_5$, define $R_6 = \{2\}, S_6 = \emptyset$ and $T_6 = \{1,3,4,5\}$. We have $t_6 = 4.33$. The partition is optimal for confidence level ranging from 99.3% to 99.8%. Note that the p-value for the test statistic S^o_{01} is 0.002.

When $\alpha = 0.05$, the critical value $\tilde{t}_{0.05,6,54}$ with $k = 6$ and $\nu = 54$ is 3.116. The 95% simultaneous confidence lower bound $L(-\mu_1 + 0.35\mu_2 - 0.35\mu_3 + \mu_6) = 5.06$ can be obtained at Step (4) with

$$n\mathbf{c}^o = (-1, 0.35, -0.35, 0.252, 0.738, 0.010)'.$$

If we are interested 95% simultaneous confidence lower bound for the pairwise comparison $\mu_6 - \mu_1$, Step (0) remains the same as above. In Step (1), we have $\delta_1 = \max\{\frac{1}{100}, 0, \frac{1}{40}, \frac{1}{60}\} = 1/40$. However, $A_1 = 1/5, B_1 = 80, t_1 = 1.76$ remain also the same as in Step (1). But in Step (2), we have $\delta_2 = \max\{\frac{1}{100}, \frac{1}{160}, \frac{1}{80}\} = 1/80, A_2 = 3/20, B_2 = 160$ and $t_2 = 2.94$. Therefore, the R_2, T_2 and S_2 form the optimal partition for confidence level between 20.9% and 92.7%. In Step (3) we have that $R_3 = \{1,3\}, S_3 = \emptyset$ and $T_3 = \{2,4,5\}$. Since $\delta_3 = \max\{\frac{1}{100}, \frac{1}{200}\} = 1/100, \Delta = 320/3, A_3 = 2/15, B_3 = 800/3$ and $t_3 = 3.51$. The R_3, T_3 and S_3 form the optimal partition for confidence level between 92.7% and 98.0%. The 95% simultaneous confidence lower bound for $\mu_6 - \mu_1$ is 5.17 with $n\mathbf{c}^o = (-1, 0.232, -0.232, 0.256, 0.720, 0.024)'$. Since $\mu_6 - \mu_1 > -\mu_1 + 0.35\mu_2 - 0.35\mu_3 + \mu_6$, it follows that $L(\mu_6 - \mu_1) = 5.17$ is bounded from below by $L(-\mu_1 + 0.35\mu_2 - 0.35\mu_3 + \mu_6) = 5.06$.

4.4 Technical Results

4.4.1 Derivation of the Optimal Solution

Consider the transformations $X_i = \bar{Y}_{i+1} - \bar{Y}_i, \delta_i = \mu_{i+1} - \mu_i, a_i = \sum_{j=i+1}^{k} n_j c_j$. Then $X_1, X_2, ..., X_{k-1}$ are normally distributed with means δ_i and covariance matrix $\sigma^2\Sigma = \sigma^2[\sigma_{ij}]$, where $\sigma_{ii} = (n_i^{-1} + n_{i+1}^{-1}), \sigma_{i,i+1} = \sigma_{i+1,i} = -n_{i+1}^{-1}$ and $\sigma_{ij} = 0$ if $|j-i| > 1$. Note that $\sum_{i=1}^{k} n_i c_i \mu_i = \sum_{i=1}^{k-1} a_i \delta_i$. Let $\mathbf{X} = (X_1, X_2, ..., X_{k-1})'$, the optimization problem (4.3) becomes the problem

$$\max_{\mathbf{0} \leq \mathbf{a} \leq \mathbf{a}^*} l(\sum_{i=1}^{k-1} a_i \delta_i) = \max_{\mathbf{0} \leq \mathbf{a} \leq \mathbf{a}^*} \{\mathbf{a}'\mathbf{X} - \bar{t}_\alpha s(\mathbf{a}'\Sigma\mathbf{a})^{1/2}\}. \tag{4.9}$$

If $a_i^* = 0$, so is a_i and the corresponding terms on the right-hand side of (4.9) vanish. Without loss of generality we may assume $a_i^* > 0$ for each $i = 1, ..., k-1$.

Let $\mathbf{a}^o$ be a vector such that $\mathbf{0} \leq \mathbf{a}^o \leq \mathbf{a}^*$ and let $R = \{i : 0 < a_i^o = a_i^*\}, S = \{i : a_i^o = 0\}$ and $T = \{i : 0 < a_i^o < a_i^*\}$. Then $\mathbf{a}^o$ and $\mathbf{a}^*$ can be partitioned as $\mathbf{a}^o = [\mathbf{a}_R^{o'}, \mathbf{a}_S^{o'}, \mathbf{a}_T^{o'}]'$ and $\mathbf{a}^* = [\mathbf{a}_R^{*'}, \mathbf{a}_S^{*'}, \mathbf{a}_T^{*'}]'$. The matrix Σ and the vector $\mathbf{X}$

are partitioned likewise. A necessary and sufficient condition for the optimal solution to be attained at $\mathbf{a}^o$ is given by Lemma 4.4.2 which is another version of Theorem 4.1.1. We will introduce Lemma 4.4.1 first, which will be used in the proof of Lemma 4.4.2.

Lemma 4.4.1 *The function* $f(\mathbf{x}) = (\mathbf{x}'\Sigma\mathbf{x})^{1/2}$ *is convex.*

Proof. It suffices to prove that the Hessian matrix $(\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j})$ is positive semi-definite (Rockafellar 1972). By taking the derivatives of the function

$$f(\mathbf{x}) = (\mathbf{x}'\Sigma\mathbf{x})^{1/2} = (\sum_i \sum_j \sigma_{ij} x_i x_j)^{1/2},$$

we have that

$$\frac{\partial f(\mathbf{x})}{\partial x_i} = \frac{1}{f(\mathbf{x})} \sum_j \sigma_{ij} x_j$$

and

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} = \sigma_{ij}/f(\mathbf{x}) - (\sum_s \sigma_{is} x_s)(\sum_t \sigma_{tj} x_t)/f^3(\mathbf{x}).$$

For any $k \times 1$ vector $\mathbf{y}$,

$$\begin{aligned} \mathbf{y}'(\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j})\mathbf{y} &= \sum_i \sum_j y_i \sigma_{ij} y_j / f(\mathbf{x}) - \sum_i \sum_j (\sum_s y_i \sigma_{is} x_s)(\sum_t y_j \sigma_{jt} x_t)/f^3(\mathbf{x}) \\ &= \sum_i \sum_j y_i \sigma_{ij} y_j / f(\mathbf{x}) - (\sum_i \sum_s y_i \sigma_{is} x_s)(\sum_t \sum_j x_t \sigma_{tj} y_j)/f^3(\mathbf{x}) \\ &= [(\mathbf{x}'\Sigma\mathbf{x})(\mathbf{y}'\Sigma\mathbf{y}) - (\mathbf{x}'\Sigma\mathbf{y})^2]/f^3(\mathbf{x}). \end{aligned}$$

Let $\tilde{\mathbf{x}} = \Sigma^{1/2}\mathbf{x}$ and $\tilde{\mathbf{y}} = \Sigma^{1/2}\mathbf{y}$, then by the Cauchy-Schwarz inequality

$$(\mathbf{x}'\Sigma\mathbf{x})(\mathbf{y}'\Sigma\mathbf{y}) - (\mathbf{x}'\Sigma\mathbf{y})^2 = (\tilde{\mathbf{x}}'\tilde{\mathbf{x}})(\tilde{\mathbf{y}}'\tilde{\mathbf{y}}) - (\tilde{\mathbf{x}}'\tilde{\mathbf{y}})^2 \geq 0.$$

□

Lemma 4.4.2 *The maximum of $l(\sum_{i=1}^{k-1} a_i\delta_i)$ subject to $\mathbf{0} \le \mathbf{a} \le \mathbf{a}^*$ is attained at $\mathbf{a}^o$ if and only if $\mathbf{a}^o$ satisfies*

$$\mathbf{a}_T^o = -\Sigma_{TT}^{-1}\Sigma_{TR}\mathbf{a}_R^* + b\Sigma_{TT}^{-1}\mathbf{X}_T \tag{4.10}$$

$$\mathbf{X}_R - b^{-1}(\Sigma_{RR}\mathbf{a}_R^* + \Sigma_{RT}\mathbf{a}_T^o) \ge 0; \tag{4.11}$$

$$\mathbf{X}_S - b^{-1}(\Sigma_{SR}\mathbf{a}_R^* + \Sigma_{ST}\mathbf{a}_T^o) \le 0. \tag{4.12}$$

$$\mathbf{X}_T'\Sigma_{TT}^{-1}\mathbf{X}_T < \tilde{t}_\alpha^2 s^2. \tag{4.13}$$

where

$$b^2 = \mathbf{a}_R^{*\prime}\Sigma_{RR.T}\mathbf{a}_R^*/[\tilde{t}_\alpha^2 s^2 - \mathbf{X}_T'\Sigma_{TT}^{-1}\mathbf{X}_T] \tag{4.14}$$

with the convention $\Sigma_{RR.T} = \Sigma_{RR} - \Sigma_{RT}\Sigma_{TT}^{-1}\Sigma_{TR}$.

Proof. Consider the optimization problem

$$\text{maximize} \quad l(\sum_{i=1}^{k-1} a_i\delta_i) \quad \text{subject to} \quad 0 \le \mathbf{a} \le \mathbf{a}^*. \tag{4.15}$$

By Lemma 4.4.1, $l(\sum_{i=1}^{k-1} a_i\delta_i)$ is concave. Let $\phi(\mathbf{a}, \mathbf{u}) = \mathbf{a}'\mathbf{X} - \tilde{t}_\alpha s(\mathbf{a}'\Sigma\mathbf{a})^{1/2} + \mathbf{u}'(\mathbf{a}^* - \mathbf{a})$ and let $\frac{\partial l}{\partial \mathbf{a}^o}$ denote the partial derivatives evaluated at the point $(\mathbf{a}^o, \mathbf{u}^o)$. By the Kuhn-Tucker equivalence theorem (Kuhn and Tucker 1951), $\mathbf{a}^o$ is the solution to the problem (4.15) if and only if

(i) $\frac{\partial l}{\partial \mathbf{a}^o} - \mathbf{u}^o \le 0, (\frac{\partial l}{\partial \mathbf{a}^o} - \mathbf{u}^o)'\mathbf{a}^o = 0$ and $\mathbf{a}^o \ge 0$,

(ii) $\mathbf{a}^* - \mathbf{a}^o \ge 0, (\mathbf{a}^* - \mathbf{a}^o)'\mathbf{u}^o = 0$ and $\mathbf{u}^o \ge 0$.

It is trivial that (i) and (ii) are equivalent to

$$\frac{\partial l}{\partial \mathbf{a}_R^o} = \mathbf{u}_R^o \ge 0, \tag{4.16}$$

$$\frac{\partial l}{\partial \mathbf{a}_S^o} \leq \mathbf{u}_S^o = 0, \tag{4.17}$$

and

$$\frac{\partial l}{\partial \mathbf{a}_T^o} = \mathbf{u}_T^o = 0, \tag{4.18}$$

where $\mathbf{u}$ has the same partition, $\mathbf{u} = [\mathbf{u}'_R, \mathbf{u}'_S, \mathbf{u}'_T]'$. The objective function $l(\sum_{i=1}^{k-1} a_i\delta_i)$ can be written as

$$l(\sum_{i=1}^{k-1} a_i\delta_i) = \mathbf{a}_R'\mathbf{X}_R + \mathbf{a}_S'\mathbf{X}_S + \mathbf{a}_T'\mathbf{X}_T - (\tilde{t}_\alpha s)c(\mathbf{a})^{1/2}$$

where $c(\mathbf{a}) = \mathbf{a}'\Sigma\mathbf{a}$. The identity (4.18) is

$$\frac{\partial l}{\partial \mathbf{a}_T^o} = \mathbf{X}_T - \tilde{t}_\alpha s(\Sigma_{TT}\mathbf{a}_T^o + \Sigma_{TR}\mathbf{a}_R^*)/c(\mathbf{a}^o)^{1/2} = 0. \tag{4.19}$$

It follows that $\mathbf{a}_T^o = -\Sigma_{TT}^{-1}\Sigma_{TR}\mathbf{a}_R^* + c(\mathbf{a}^o)^{1/2}(\tilde{t}_\alpha s)^{-1}\Sigma_{TT}^{-1}\mathbf{X}_T$. But

$$\begin{aligned}
c(\mathbf{a}^o) &= \mathbf{a}_R^{*\prime}\Sigma_{RR}\mathbf{a}_R^* + 2\mathbf{a}_R^{*\prime}\Sigma_{RT}\mathbf{a}_T^o + \mathbf{a}_T^{o\prime}\Sigma_{TT}\mathbf{a}_T^o \\
&= \mathbf{a}_R^{*\prime}\Sigma_{RR}\mathbf{a}_R^* + 2\mathbf{a}_R^{*\prime}\Sigma_{RT}[-\Sigma_{TT}^{-1}\Sigma_{TR}\mathbf{a}_R^* + c(\mathbf{a}^o)^{1/2}(\tilde{t}_\alpha s)^{-1}\Sigma_{TT}^{-1}\mathbf{X}_T] \\
&\quad +[-\Sigma_{TT}^{-1}\Sigma_{TR}\mathbf{a}_R^* + c(\mathbf{a}^o)^{1/2}(\tilde{t}_\alpha s)^{-1}\Sigma_{TT}^{-1}\mathbf{X}_T]'\Sigma_{TT}[-\Sigma_{TT}^{-1}\Sigma_{TR}\mathbf{a}_R^* \\
&\quad +c(\mathbf{a}^o)^{1/2}(\tilde{t}_\alpha s)^{-1}\Sigma_{TT}^{-1}\mathbf{X}_T] \\
&= \mathbf{a}_R^{*\prime}\Sigma_{RR}\mathbf{a}_R^* - \mathbf{a}_R^{*\prime}\Sigma_{RT}\Sigma_{TT}^{-1}\Sigma_{TR}\mathbf{a}_R^* + \mathbf{X}_T'\Sigma_{TT}^{-1}\mathbf{X}_T c(\mathbf{a}^o)(\tilde{t}_\alpha^2 s^2)^{-1}. \\
&= \mathbf{a}_R^{*\prime}\Sigma_{RR.T}\mathbf{a}_R^* + \mathbf{X}_T'\Sigma_{TT}^{-1}\mathbf{X}_T c(\mathbf{a}^o)(\tilde{t}_\alpha^2 s^2)^{-1}.
\end{aligned}$$

Hence

$$c(\mathbf{a}^o) = \tilde{t}_\alpha^2 s^2 \mathbf{a}_R^{*\prime}\Sigma_{RR.T}\mathbf{a}_R^*/(\tilde{t}_\alpha^2 s^2 - \mathbf{X}_T'\Sigma_{TT}^{-1}\mathbf{X}_T).$$

Let $b = c(\mathbf{a}^o)^{1/2}/(\tilde{t}_\alpha s)$. Then it has the same expression as (4.14) and expression (4.19) becomes (4.10). The inequalities (4.16) and (4.17) are, respectively,

$$\frac{\partial l}{\partial \mathbf{a}_R^o} = \mathbf{X}_R - b^{-1}(\Sigma_{RR}\mathbf{a}_R^* + \Sigma_{RT}\mathbf{a}_T^o) \geq 0$$

and

$$\frac{\partial l}{\partial \mathbf{a}_S^o} = \mathbf{X}_S - b^{-1}(\Sigma_{SR}\mathbf{a}_R^* + \Sigma_{ST}\mathbf{a}_T^o) \leq 0.$$

and they are (4.11) and (4.12), respectively.

For the case when T is empty, (4.10) does not apply. (4.11) and (4.12) are reduced to

$$\mathbf{X}_R \geq b^{-1}\Sigma_{RR}\mathbf{a}_R^*$$

and

$$\mathbf{X}_S \leq b^{-1}\Sigma_{SR}\mathbf{a}_R^*.$$

This completes the proof. □

4.4.2 Computation and Proof of Theorem 4.1.1

The following lemmas will be used to simplify the computation. The transformations in Section 4.4.1 will be used here. Let $\bar{\mathbf{Y}} = [\bar{Y}_1, ..., \bar{Y}_k]'$. Then $\mathbf{X} = A\bar{\mathbf{Y}}$ where $A = [a_{ij}]_{(k-1)\times k}$ is such that $a_{ii} = -1, a_{i,i+1} = 1, a_{ij} = 0$ otherwise.

Lemma 4.4.3 *The inverse matrix of Σ has the following expression:* $\Sigma^{-1} = [\sigma^{ij}], \sigma^{ij} = \sigma^{ji} = \frac{\tilde{n}_{0i}\tilde{n}_{jk}}{\tilde{n}_{0k}}$, *if* $i \leq j < k$.

Proof. It is trivial that $\Sigma_{11}^{-1} = (\frac{n_1 n_2}{\tilde{n}_{02}})$. Assume $\Sigma^{-1} = [\sigma^{ij}]$ holds for k. For the case $k+1$, we have that

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

with $\Sigma_{11} = [\sigma_{ij}]_{(k-1)\times(k-1)}, \Sigma_{21} = [0, \cdots, 0, -\frac{1}{n_k}]_{1\times(k-1)}$, and $\Sigma_{22} = \frac{1}{n_k} + \frac{1}{n_{k+1}}$. It is trivial that

$$\Sigma^{-1} = \begin{pmatrix} \Sigma_{11}^{-1} + \Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22.1}^{-1}\Sigma_{21}\Sigma_{11}^{-1} & -\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22.1}^{-1} \\ -\Sigma_{22.1}^{-1}\Sigma_{21}\Sigma_{11}^{-1} & \Sigma_{22.1}^{-1} \end{pmatrix}$$

where $\Sigma_{22.1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$. By the assumption we have that $\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} = \sigma^{k-1,k-1}/n_k^2 = \tilde{n}_{0,k-1}n_k/(n_k^2\tilde{n}_{0k}) = \tilde{n}_{0,k-1}/(n_k\tilde{n}_{0k})$, and hence $\Sigma_{22.1}^{-1} = \tilde{n}_{0k}n_{k+1}/\tilde{n}_{0,k+1} = \sigma^{kk}$. It follows that

$$\begin{aligned} -\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22.1}^{-1} &= (-1)(-\frac{1}{n_k})\frac{\tilde{n}_{0k}n_{k+1}}{\tilde{n}_{0,k+1}}\Sigma_{11}^{-1}\mathbf{e}_{k-1} \\ &= \frac{\tilde{n}_{0k}n_{k+1}}{n_k\tilde{n}_{0,k+1}}[\sigma^{i,k-1}] = \frac{1}{\tilde{n}_{0,k+1}}[\tilde{n}_{0i}n_{k+1}] = [\sigma^{ik}] \end{aligned}$$

where $\mathbf{e}_{k-1} = [0, ..., 0, 1]'_{1\times(k-1)}$ and

$$\begin{aligned} [\Sigma_{11}^{-1} + \Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22.1}^{-1}\Sigma_{21}\Sigma_{11}^{-1}]_{ij} &= \frac{\tilde{n}_{0i}\tilde{n}_{jk}}{\tilde{n}_{0k}} + \frac{\tilde{n}_{0i}n_{k+1}}{\tilde{n}_{0,k+1}}\frac{1}{n_k}\frac{\tilde{n}_{0j}n_k}{\tilde{n}_{0k}} \\ &= \frac{\tilde{n}_{0i}}{\tilde{n}_{0k}\tilde{n}_{0,k+1}}(\tilde{n}_{0,k+1}\tilde{n}_{jk} + \tilde{n}_{0j}n_{k+1}) \\ &= \frac{\tilde{n}_{0i}}{\tilde{n}_{0k}\tilde{n}_{0,k+1}}(\tilde{n}_{0k}\tilde{n}_{jk} + n_{k+1}\tilde{n}_{jk} + \tilde{n}_{0j}n_{k+1}) \\ &= \frac{\tilde{n}_{0i}}{\tilde{n}_{0k}\tilde{n}_{0,k+1}}\tilde{n}_{0k}\tilde{n}_{j,k+1} = \frac{\tilde{n}_{0i}\tilde{n}_{j,k+1}}{\tilde{n}_{0,k+1}} = \sigma^{ij}. \end{aligned}$$

□

Lemma 4.4.4 *The vector* $\Sigma^{-1}\mathbf{X}$ *and the quadratic form* $\mathbf{X}'\Sigma^{-1}\mathbf{X}$ *have the following expressions:* $[\Sigma^{-1}\mathbf{X}]_i = \tilde{n}_{0i}(\tilde{Y}_{0k} - \tilde{Y}_{0i})$ *and* $\mathbf{X}'\Sigma^{-1}\mathbf{X} = \sum_{i=1}^{k} n_i(\bar{Y}_i - \tilde{Y}_{0k})^2$.

Proof. By Lemma 4.4.3

$$
\begin{aligned}
[\Sigma^{-1}\mathbf{X}]_i &= \sum_{j=1}^{k-1} \sigma^{ij} X_j \\
&= \sum_{j=1}^{i} \frac{\tilde{n}_{0j}\tilde{n}_{ik}}{\tilde{n}_{0k}}(\bar{Y}_{j+1} - \bar{Y}_j) + \sum_{j=i+1}^{k-1} \frac{\tilde{n}_{0i}\tilde{n}_{jk}}{\tilde{n}_{0k}}(\bar{Y}_{j+1} - \bar{Y}_j).
\end{aligned}
$$

By Abel's method of summation, $\sum_{r=\alpha}^{\beta} a_r b_r = \sum_{r=\alpha}^{\beta-1}(a_r - a_{r+1}) \sum_{t=\alpha}^{r} b_t + a_\beta \sum_{t=\alpha}^{\beta} b_t$, we have that

$$
\sum_{j=1}^{i+1} n_j \bar{Y}_j = \sum_{j=1}^{i} (\bar{Y}_j - \bar{Y}_{j+1})\tilde{n}_{0j} + \bar{Y}_{i+1}\tilde{n}_{0,i+1},
$$

and

$$
\sum_{j=i+1}^{k} n_j \bar{Y}_j = \sum_{j=i+1}^{k-1} (\bar{Y}_j - \bar{Y}_{j+1})\tilde{n}_{ij} + \bar{Y}_k \tilde{n}_{ik}.
$$

It follows that

$$
\sum_{j=1}^{i} \tilde{n}_{0j}(\bar{Y}_{j+1} - \bar{Y}_j) = \tilde{n}_{0i}\bar{Y}_{i+1} - \sum_{j=1}^{i} n_j \bar{Y}_j,
$$

and

$$
\begin{aligned}
\sum_{j=i+1}^{k-1} \tilde{n}_{jk}(\bar{Y}_{j+1} - \bar{Y}_j) &= \sum_{j=i+1}^{k-1} (\bar{Y}_{j+1} - \bar{Y}_j)(\tilde{n}_{ik} - \tilde{n}_{ij}) \\
&= \tilde{n}_{ik}(\bar{Y}_k - \bar{Y}_{i+1}) + \sum_{j=i+1}^{k-1} (\bar{Y}_j - \bar{Y}_{j+1})\tilde{n}_{ij} \\
&= \sum_{j=i+1}^{k} n_j \bar{Y}_j - \tilde{n}_{ik}\bar{Y}_{i+1}.
\end{aligned}
$$

Hence

$$
\begin{aligned}
[\Sigma^{-1}\mathbf{X}]_i &= \frac{\tilde{n}_{ik}}{\tilde{n}_{0k}}(\tilde{n}_{0i}\bar{Y}_{i+1} - \sum_{j=1}^{i} n_j \bar{Y}_j) + \frac{\tilde{n}_{0i}}{\tilde{n}_{0k}}(\sum_{j=i+1}^{k} n_j \bar{Y}_j - \tilde{n}_{ik}\bar{Y}_{i+1}) \\
&= \frac{\tilde{n}_{0i}\tilde{n}_{ik}}{\tilde{n}_{0k}}(\tilde{Y}_{ik} - \tilde{Y}_{0i}) = \tilde{n}_{0i}(\tilde{Y}_{0k} - \tilde{Y}_{0i}).
\end{aligned}
$$

Let $B = A'\Sigma^{-1}\mathbf{X}$. Then we have that $B_1 = n_1(\bar{Y}_1 - \tilde{Y}_{0k}), B_i = \tilde{n}_{0i-1}(\tilde{Y}_{0k} - \tilde{Y}_{0i-1}) - \tilde{n}_{0i}(\tilde{Y}_{0k} - \tilde{Y}_{0i}) = n_i(\bar{Y}_i - \tilde{Y}_{0k}), i = 2, ..., k-1$, and

$$\begin{aligned} B_k &= \tilde{n}_{0,k-1}(\tilde{Y}_{0k} - \tilde{Y}_{0,k-1}) = \tilde{n}_{0,k-1}\tilde{Y}_{0k} - \sum_{j=1}^{k-1} n_j \bar{Y}_j \\ &= \tilde{n}_{0,k-1}\tilde{Y}_{0k} + n_k\bar{Y}_k - \tilde{n}_{0k}\tilde{Y}_{0k} = n_k(\bar{Y}_k - \tilde{Y}_{0k}). \end{aligned}$$

It follows that $\mathbf{X}'\Sigma^{-1}\mathbf{X} = \bar{\mathbf{Y}}'A'\Sigma^{-1}\mathbf{X} = \sum \bar{Y}_i n_i(\bar{Y}_i - \tilde{Y}_{0k}) = \sum n_i(\bar{Y}_i - \tilde{Y}_{0k})^2$. This completes the proof. □

Let $Q = R \cup S = \{r_1, r_2, ..., r_{l-1}\}$, let $T_i = \{r_{i-1}+1, ..., r_i - 1\}$ if $r_{i-1}+1 \leq r_i - 1$ and let $T = \{t_1, t_2, ..., t_{k-l}\} = T_1 \cup T_2 \cup \cdots \cup T_l$ with the conventions $r_0 = 0$ and $r_l = k$. Note that T_i is an empty set $\emptyset$ if r_{i-1} and r_i are consecutive integers and $R \cup S \cup T = \{1, \cdots, k-1\}$. Let p, q and r be three consecutive indices in $Q \cup \{0, k\}$ $(q \neq 0, k)$. We shall denote $q = r_i \in Q$. If $i = 1$ then $p = r_{i-1} = 0$; if $i = l-1$ then $r = r_{i+1} = k$. Let π be the permutation

$$\pi = \begin{pmatrix} 1 & \cdots & l-1 & l & \cdots & k-1 \\ r_1 & \cdots & r_{l-1} & t_1 & \cdots & t_{k-l} \end{pmatrix}$$

and let Γ be the corresponding elementary operation matrix which permutes rows according to π, i.e.,

$$\Gamma\Sigma\Gamma' = \begin{bmatrix} \Sigma_{QQ} & \Sigma_{QT} \\ \Sigma_{TQ} & \Sigma_{TT} \end{bmatrix}.$$

Note that

$$\Sigma_{TT} = \begin{bmatrix} \Sigma_{T_1T_1} & 0 & \cdots & 0 \\ 0 & \Sigma_{T_2T_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Sigma_{T_lT_l} \end{bmatrix},$$

$\Sigma_{TQ} = [\Sigma_{Tr_1}, \cdots, \Sigma_{Tr_{l-1}}]$, $\Sigma_{Tr_i} = [\Sigma'_{T_1 r_i}, \cdots, \Sigma'_{T_l r_i}]'$ and

$$\begin{cases} \Sigma'_{T_i q} = [0, \cdots, 0, -1/n_q] & \text{if } T_i \neq \emptyset, \\ \Sigma'_{T_{i+1} q} = [-1/n_{q+1}, 0, \cdots, 0] & \text{if } T_{i+1} \neq \emptyset, \\ \Sigma'_{T_j q} = 0 & \text{otherwise.} \end{cases} \tag{4.20}$$

From Lemma 4.4.3, we have $[\Sigma^{-1}_{T_i T_i}]_{\alpha\beta} = \frac{\tilde{n}_{p\alpha}\tilde{n}_{\beta q}}{\tilde{n}_{pq}}$ if $p < \alpha \leq \beta < q$ and $[\Sigma^{-1}_{T_{i+1} T_{i+1}}]_{\alpha\beta} = \frac{\tilde{n}_{q\alpha}\tilde{n}_{\beta r}}{\tilde{n}_{qr}}$ if $q < \alpha \leq \beta < r$. Therefore,

$$\begin{cases} -\Sigma^{-1}_{T_i T_i}\Sigma_{T_i q} = [\frac{n_{p+1}}{\tilde{n}_{pq}}, \cdots, \frac{\tilde{n}_{p,q-1}}{\tilde{n}_{pq}}]' & \text{if } T_i \neq \emptyset, \\ -\Sigma^{-1}_{T_{i+1} T_{i+1}}\Sigma_{T_{i+1} q} = [\frac{\tilde{n}_{q+1,r}}{\tilde{n}_{qr}}, \cdots, \frac{n_r}{\tilde{n}_{qr}}]' & \text{if } T_{i+1} \neq \emptyset, \\ -\Sigma^{-1}_{T_j T_j}\Sigma_{T_j q} = 0 & \text{if } j \neq i, i+1. \end{cases} \tag{4.21}$$

Lemma 4.4.5 *Let* $\Sigma_{QQ.T} = \Sigma_{QQ} - \Sigma_{QT}\Sigma^{-1}_{TT}\Sigma_{TQ} = [\tau_{ij}]$, *then* $\tau_{qq} = \frac{1}{\tilde{n}_{pq}} + \frac{1}{\tilde{n}_{qr}}$, $\tau_{qr} = \tau_{rq} = -\frac{1}{\tilde{n}_{qr}}$ *and* $\tau_{ij} = 0$ *otherwise.*

Proof. By (4.20) and (4.21), we have that

$$\begin{aligned} \tau_{qq} &= [\Sigma_{QQ} - \Sigma_{QT}\Sigma^{-1}_{TT}\Sigma_{TQ}]_{qq} \\ &= \frac{1}{n_q} + \frac{1}{n_{q+1}} - \frac{1}{n_q}\frac{\tilde{n}_{p,q-1}}{\tilde{n}_{pq}} - \frac{1}{n_{q+1}}\frac{\tilde{n}_{q+1,r}}{\tilde{n}_{qr}} \\ &= \frac{1}{n_q}(1 - \frac{\tilde{n}_{p,q-1}}{\tilde{n}_{pq}}) + \frac{1}{n_{q+1}}(1 - \frac{\tilde{n}_{q+1,r}}{\tilde{n}_{qr}}) \\ &= \frac{1}{\tilde{n}_{pq}} + \frac{1}{\tilde{n}_{qr}}; \end{aligned}$$

$$\tau_{qr} = [\Sigma_{QQ} - \Sigma_{QT}\Sigma^{-1}_{TT}\Sigma_{TQ}]_{qr} = -\frac{1}{n_{q+1}} = -\frac{1}{\tilde{n}_{qr}} \quad \text{if } r = q+1;$$

$$\tau_{qr} = (-\frac{1}{n_{q+1}})\frac{n_{q+1}}{\tilde{n}_{qr}} = -\frac{1}{\tilde{n}_{qr}} \quad \text{if } r > q+1.$$

It is trivial that $[\Sigma_{QQ} - \Sigma_{QT}\Sigma^{-1}_{TT}\Sigma_{TQ}]_{\alpha\beta} = 0$ if α, β are not consecutive indices in Q. This completes the proof. □

Let $\tilde{a}_i = a_i^*$ if $i \in R$ and $\tilde{a}_i = 0$ if $i \in S$. The expression (4.10) can be rewritten as

$$\mathbf{a}_T^o = -\Sigma_{TT}^{-1}\Sigma_{TQ}\tilde{\mathbf{a}}_Q + b\Sigma_{TT}^{-1}\mathbf{X}_T. \tag{4.22}$$

By the fact that $\Sigma_{TT}^{-1}\mathbf{X}_T = [(\Sigma_{T_1T_1}^{-1}\mathbf{X}_{T_1})', \cdots, (\Sigma_{T_lT_l}^{-1}\mathbf{X}_{T_l})']'$ and by Lemma 4.4.4, $[\Sigma_{T_iT_i}^{-1}\mathbf{X}_{T_i}]_j = \tilde{n}_{pj}(\tilde{Y}_{pq} - \tilde{Y}_{pj}),\ p < j < q$. Therefore,

$$a_j^o = (\tilde{n}_{jq}\tilde{a}_p + \tilde{n}_{pj}\tilde{a}_q)/\tilde{n}_{pq} + b\tilde{n}_{pj}(\tilde{Y}_{pq} - \tilde{Y}_{pj}),\ p < j < q. \tag{4.23}$$

By (4.22), the left-hand sides of the inequalities (4.11) and (4.12) can be combined as

$$\mathbf{X}_Q - b^{-1}(\Sigma_{QQ}\tilde{\mathbf{a}}_Q + \Sigma_{QT}\mathbf{a}_T^o) = (\mathbf{X}_Q - \Sigma_{QT}\Sigma_{TT}^{-1}\mathbf{X}_T) - b^{-1}(\Sigma_{QQ.T}\tilde{\mathbf{a}}_Q)$$

where

$$\begin{aligned}[\mathbf{X}_Q - \Sigma_{QT}\Sigma_{TT}^{-1}\mathbf{X}_T]_q &= \bar{Y}_{q+1} - \bar{Y}_q \\ &+ (0, ..., 0, \frac{n_{p+1}}{\tilde{n}_{pq}}, ..., \frac{\tilde{n}_{p,q-1}}{\tilde{n}_{pq}}, \frac{\tilde{n}_{q+1,r}}{\tilde{n}_{qr}}, ..., \frac{n_r}{\tilde{n}_{qr}}, 0, ..., 0)\mathbf{X}_T \\ &= \tilde{Y}_{qr} - \tilde{Y}_{pq}\end{aligned}$$

and $[\Sigma_{QQ.T}\tilde{\mathbf{a}}_Q]_q = (\tilde{a}_q - \tilde{a}_r)/\tilde{n}_{qr} - (\tilde{a}_p - \tilde{a}_q)/\tilde{n}_{pq}$.

Therefore, (4.11) and (4.12) become, respectively,

$$b(\tilde{Y}_{qr} - \tilde{Y}_{pq}) \geq (\tilde{a}_q - \tilde{a}_r)/\tilde{n}_{qr} - (\tilde{a}_p - \tilde{a}_q)/\tilde{n}_{pq} \ \text{ if } q \in R; \tag{4.24}$$

and

$$b(\tilde{Y}_{qr} - \tilde{Y}_{pq}) \leq (\tilde{a}_q - \tilde{a}_r)/\tilde{n}_{qr} - (\tilde{a}_p - \tilde{a}_q)/\tilde{n}_{pq} \ \text{ if } q \in S. \tag{4.25}$$

Proof of Theorem 4.1.1

From (4.22), we have that

$$\mathbf{a}^o_\pi = \begin{pmatrix} \mathbf{a}^o_Q \\ \mathbf{a}^o_T \end{pmatrix} = \begin{pmatrix} I_{l-1} \\ -\Sigma^{-1}_{TT}\Sigma_{TQ} \end{pmatrix} \tilde{\mathbf{a}}_Q + b \begin{pmatrix} 0 \\ \Sigma^{-1}_{TT}\mathbf{X}_T \end{pmatrix}.$$

By the inverse permutation, we have that $\mathbf{a}^o = \Gamma'\mathbf{a}^o_\pi$. The optimal coefficient $\mathbf{nc}^o = A'\mathbf{a}^o$ is expressed as

$$\mathbf{nc}^o = A'\Gamma' \begin{pmatrix} I_{l-1} \\ -\Sigma^{-1}_{TT}\Sigma_{TQ} \end{pmatrix} \tilde{\mathbf{a}}_Q + bA'\Gamma' \begin{pmatrix} 0 \\ \Sigma^{-1}_{TT}\mathbf{X}_T \end{pmatrix}. \qquad (4.26)$$

The $k \times (l-1)$ matrix $[A'\Gamma' \begin{pmatrix} I_{l-1} \\ -\Sigma^{-1}_{TT}\Sigma_{TQ} \end{pmatrix}]$ can be evaluated as:

$$[A'\Gamma' \begin{pmatrix} I_{l-1} \\ -\Sigma^{-1}_{TT}\Sigma_{TQ} \end{pmatrix}]_{ji} = \begin{cases} -\frac{n_j}{\tilde{n}_{pq}}; & \text{if } p < j \leq q \\ \frac{n_j}{\tilde{n}_{qr}} & \text{if } q < j \leq r \\ 0 & \text{otherwise.} \end{cases}$$

The first term of (4.26) is

$$[A'\Gamma' \begin{pmatrix} I_{l-1} \\ -\Sigma^{-1}_{TT}\Sigma_{TQ} \end{pmatrix} \tilde{\mathbf{a}}_Q]_j = \frac{n_j}{\tilde{n}_{pq}}(\tilde{a}_p - \tilde{a}_q) = n_j \tilde{c}_{pq} \text{ if } p < j \leq q. \qquad (4.27)$$

The second term of (4.26) can be evaluated as

$$[A'\Gamma' \begin{pmatrix} 0 \\ \Sigma^{-1}_{TT}\mathbf{X}_T \end{pmatrix}]_j = n_j(\bar{Y}_j - \tilde{Y}_{pq}), \text{ if } p < j \leq q. \qquad (4.28)$$

By (4.26), (4.27) and (4.28), the identity (4.10) is equivalent to (4.4). By (4.24) and (4.25),

$$\tilde{c}_{qr} - \tilde{c}_{pq} \leq b(\tilde{Y}_{qr} - \tilde{Y}_{pq}) \text{ if } q \in R;$$

$$\tilde{c}_{qr} - \tilde{c}_{pq} \geq b(\tilde{Y}_{qr} - \tilde{Y}_{pq}) \text{ if } q \in S.$$

By Lemma 4.4.4, $\mathbf{X}'_T\Sigma^{-1}_{TT}\mathbf{X}_T = \sum_q \sum_{j=p+1}^q n_j(\bar{Y}_j - \tilde{Y}_{pq})^2$. Note that $\mathbf{a}^{*\prime}_R\Sigma_{RR.T}\mathbf{a}^*_R = \tilde{\mathbf{a}}'_Q\Sigma_{QQ.T}\tilde{\mathbf{a}}_Q = \sum_q \tilde{n}_{pq}\tilde{c}^2_{pq}$. This completes the proof. □

4.4.3 Justification of the Algorithm

The following proofs are derived from Theorem 4.1.1 as well as (4.23), (4.24) and (4.25).

Let $R_0 \supseteq \{j < k : \bar{Y}_j < \bar{Y}_{j+1}, a_j^* > 0\}, S_0 \supseteq \{j < k : \bar{Y}_j > \bar{Y}_{j+1}, a_j^* > 0\}$ and let p, q and r be three consecutive indices in $R_0 \cup S_0 \cup \{0, k, k+1\}(q \neq 0, k+1)$. Then $\bar{Y}_j = \tilde{Y}_{pq}$ if $p < j \leq q$ and hence $SSW = 0$. By Theorem 4.1.1, the optimal solution $\mathbf{c}^o$ is the one such that $c_j^o = \tilde{c}_{pq}$, $p < j \leq q$, if

$$b \geq \delta_0 = \delta_{01} = \sup\{(\tilde{c}_{qr} - \tilde{c}_{pq})/(\tilde{Y}_{qr} - \tilde{Y}_{pq}) : q \in R_0 \cup S_0, \tilde{Y}_{qr} \neq \tilde{Y}_{pq}\}$$

where $b = (\Sigma_q \tilde{n}_{pq} \tilde{c}_{pq}^2)^{1/2}/(\tilde{t}_\alpha s)$. The above inequality is equivalent to $\tilde{t}_\alpha \leq t_{\alpha_0} = A_0^{1/2} s/\delta_0$. Confidence lower bound (4.3) is solved for confidence level up to $1 - \alpha_0$.

Let R_i, S_i and T_i be the optimal partition satisfying (4.4), (4.5), (4.6) and (4.7) of Theorem 4.1.1 for a given $\tilde{t}_\alpha < t_{\alpha_i}(\alpha > \alpha_i)$. We shall show that $S_i \supseteq S_{i+1}$. Let $q \in S_{i+1}$ have an immediate predecessor p and an immediate successor r respectively in $R_{i+1} \cup S_{i+1} \cup \{0, k\}$. Then by (4.25), we have that for any $b(\alpha), 0 < b(\alpha) \leq \delta_i$,

$$\tilde{Y}_{qr} - \tilde{Y}_{pq} \leq (-\tilde{a}_p/\tilde{n}_{pq} - \tilde{a}_r/\tilde{n}_{qr})/b(\alpha) \leq 0.$$

Suppose that $q \in T_i$. From (4.23), we have that for any $b(\alpha) > \delta_i$

$$\begin{aligned} a_q^o[b(\alpha)] &= (\tilde{n}_{qr}\tilde{a}_p + \tilde{n}_{pq}\tilde{a}_r)/\tilde{n}_{pr} + b(\alpha)\tilde{n}_{pq}\tilde{n}_{qr}(\tilde{Y}_{qr} - \tilde{Y}_{pq})/\tilde{n}_{pr} \\ &= a_q^o(\delta_i) + [b(\alpha) - \delta_i]\tilde{n}_{pq}\tilde{n}_{qr}(\tilde{Y}_{qr} - \tilde{Y}_{pq})/\tilde{n}_{pr}. \end{aligned} \quad (4.29)$$

Since $a_q^o(\delta_i) = 0$, $a_q^o[b(\alpha)] \leq 0$ for any $b(\alpha) > \delta_i$. This contradicts that $q \in T_i$. It follows that $S_i \supseteq S_{i+1}$.

As the confidence level $1-\alpha$ (and hence $\tilde{t}_\alpha$) increases, the optimal partition holds until either

(I) $\tilde{t}_\alpha \leq t_{\alpha_i}$ and there exists a $q \in R_i$ so that $R_{i+1} = R_i - \{q\}, S_{i+1} = S_i$ and $T_{i+1} = T_i \cup \{q\}$ is the optimal partition for $\tilde{t}_\alpha > t_{\alpha_i}$, or

(II) $\tilde{t}_\alpha \leq t_{\alpha_i}$ and there exists a $q \in S_i$ so that $R_{i+1} = R_i, S_{i+1} = S_i - \{q\}$ and $T_{i+1} = T_i \cup \{q\}$ is the optimal partition for $\tilde{t}_\alpha > t_{\alpha_i}$, or

(III) $\tilde{t}_\alpha < t_{\alpha_i}$ and there exists a $j \in T_i$ so that $R_{i+1} = R_i \cup \{j\}, S_{i+1} = S_i$ and $T_{i+1} = T_i - \{j\}$ is the optimal partition for $\tilde{t}_\alpha \geq t_{\alpha_i}$.

We shall prove the Case (I) only and the proofs for the Case (II) and (III) follow similarly.

Let $q \in R_i$ have an immediate predecessor p and an immediate successor r respectively in $R_i \cup S_i \cup \{0, k\}$. For $\tilde{t}_\alpha \leq t_{\alpha_i}$, we have that $b^2(\alpha) = A_i/(\tilde{t}_\alpha^2 s^2 - B_i)$ where

$$A_i = \sum_q \tilde{n}_{pq}(\tilde{c}_{pq})^2$$

and

$$B_i = \sum_q \sum_{j=p+1}^{q} n_j(\bar{Y}_j - \tilde{Y}_{pq})^2.$$

Then

$$\tilde{c}_{qr} - \tilde{c}_{pq} = \delta_i(\tilde{Y}_{qr} - \tilde{Y}_{pq}). \tag{4.30}$$

For $\tilde{t}_\alpha > t_{\alpha_i}$, we have that $b^2(\alpha) = A_{i+1}/(\tilde{t}_\alpha^2 s^2 - B_{i+1})$ where $A_{i+1} = A_i - \delta_i^2 \Delta$ and $B_{i+1} = B_i + \Delta$ with

$$\Delta = (\tilde{Y}_{pq} - \tilde{Y}_{qr})^2 / (\tilde{n}_{pq}^{-1} + \tilde{n}_{qr}^{-1}).$$

Therefore, $\lim_{\alpha \to \alpha_i -} b^2(\alpha) = A_{i+1}/(t_i^2 s^2 - B_{i+1})$. However,

$$\frac{A_{i+1}}{t_i^2 s^2 - B_{i+1}} = \frac{A_i - \delta_i^2 \Delta}{t_i^2 s^2 - B_i - \Delta} = \frac{A_i - \delta_i^2 \Delta}{A_i/\delta_i^2 - \Delta} = \delta_i^2.$$

It follows that $\lim_{\alpha \to \alpha_i -} b(\alpha) = \delta_i$ and the coefficient $b(\alpha)$ is a continuous, increasing function of α.

By (4.4), when $b(\alpha) = \delta_i$ we have that

$$c_j^{(i)} = \tilde{c}_{pq} + \delta_i(\bar{Y}_j - \tilde{Y}_{pq}), \quad \text{if } \; p < j \le q.$$

$$c_j^{(i)} = \tilde{c}_{qr} + \delta_i(\bar{Y}_j - \tilde{Y}_{qr}), \quad \text{if } \; q < j \le r.$$

where $c_j^{(i)}$ denotes the optimal solution for the partition R_i, S_i and T_i. By (4.30),

$$\tilde{c}_{pq} - \delta_i \tilde{Y}_{pq} = \tilde{c}_{qr} - \delta_i \tilde{Y}_{qr} = \tilde{c}_{pr} - \delta_i \tilde{Y}_{pr} \tag{4.31}$$

where $\tilde{c}_{pr} = (\tilde{n}_{pq}\tilde{c}_{pq} + \tilde{n}_{qr}\tilde{c}_{qr})/(\tilde{n}_{pq} + \tilde{n}_{qr})$ and $\tilde{Y}_{pr} = (\tilde{n}_{pq}\tilde{Y}_{pq} + \tilde{n}_{qr}\tilde{Y}_{qr})/(\tilde{n}_{pq} + \tilde{n}_{qr})$. It follows that

$$c_j^{(i+1)} = \tilde{c}_{pr} + \delta_i(\bar{Y}_j - \tilde{Y}_{pr}) \;\text{ if } \; p < j \le r.$$

Let (4.4) hold for the partition R_{i+1}, S_{i+1} and T_{i+1} when $b(\alpha) < \delta_i$. Then

$$0 < \sum_{j=h+1}^{k} n_j c_j^{(i+1)}(\alpha) < \sum_{j=h+1}^{k} n_j c_j^*$$

for each $h \in T_{i+1}$, except $h = q$, and hence the inequality holds for b in the neighborhood of δ_i, $\delta_{i+1} < b(\alpha) < \delta_i$. Since $q \in R_i$, the last inequality becomes an identity when $h = q$. By the fact that $\tilde{Y}_{qr} > \tilde{Y}_{pr}$, $0 < \sum_{j=q+1}^{k} n_j c_j^{(i+1)}(\alpha) < \sum_{j=q+1}^{k} n_j c_j^*$ for $\delta_{i+1} < b(\alpha) < \delta_i$.

By the assumption that δ_i is determined by $q \in R_i$, (4.5) holds for each $h \in R_{i+1}$ for $\delta_{i+1} < b(\alpha) < \delta_i$ with the exception of $h = p$ or $h = r$. Suppose that $p \in R_i$ with an immediate predecessor m in $R_i \cup S_i \cup \{0, k\}$. Then

$$\tilde{c}_{pq} - \tilde{c}_{mp} < \delta_i(\tilde{Y}_{pq} - \tilde{Y}_{mp}).$$

By (4.31),

$$\tilde{c}_{pr} - \tilde{c}_{mp} < \delta_i(\tilde{Y}_{pr} - \bar{Y}_{mp}).$$

Therefore, (4.5) holds for the partition R_{i+1}, S_{i+1} and T_{i+1} when $\delta_{i+1} < b(\alpha) < \delta_i$. The proof for the case $h = r$ and the case (4.6) follows similarly. For $\tilde{t}_\alpha > t_{\alpha_i}$ we have that

$$\tilde{t}_\alpha^2 > t_{\alpha_i}^2 = B_i + (A_i/\delta_i^2) = B_{i+1} - \Delta + (A_{i+1} + \delta_i^2\Delta)/\delta_i^2 > B_{i+1}.$$

Therefore, (4.7) holds for the new partition. Since each optimal partition R_i, S_i and T_i holds for a specific range of $1 - \alpha$, $\delta_i < b(\alpha) < \delta_{i-1}$, the algorithm will terminate after a finite number of steps.

For pairwise comparisons $\mu_k - \mu_1$, we have that $a_i^* = 1, i = 1, 2, ..., k-1$. Let $q \in R_{i+1}$ have an immediate predecessor p and an immediate successor r respectively in $R_{i+1} \cup S_{i+1} \cup \{0, k\}$. Then by (4.24), we have that for any $b(\alpha) \leq \delta_i$,

$$\tilde{Y}_{qr} - \tilde{Y}_{pq} \geq [(1 - \tilde{a}_p)/\tilde{n}_{pq} + (1 - \tilde{a}_r)/\tilde{n}_{qr}]/b(\alpha) \geq 0.$$

Suppose that $q \in T_i$. We also have (4.29) holds for any $b(\alpha) > \delta_i$. Since

$a_q^o(\delta_i) = 1$, $a_q^o[b(\alpha)] \geq 1$ for any $b(\alpha) > \delta_i$. This contradicts that $q \in T_i$. It follows that $R_i \supseteq R_{i+1}$. □

4.5 Conclusion

The use of prior knowledge that the regression curve is monotone, $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_k$, can sharpen confidence bounds. The $100(1-\alpha)\%$ simultaneous confidence lower bound in the numerical example in Section 4.3 for $\mu_6 - \mu_1$ is -1.86 without the prior knowledge and it is 5.17 with the prior knowledge.

Incorporating the prior knowledge of monotonicity, Marcus (1978) studied the optimal lower bound for the nonnegative contrasts when the common variance σ is known and her method requires computation of a large number of feasible partitions when R is nonempty. Our algorithm in Section 4.2 simplifies the computations. At each step of the algorithm, an optimal partition is found for an optimal solution with a different confidence coefficient until a desired level of $1-\alpha$ is reached. The algorithm terminates after a finite number of steps.

Theorem 4.1.1, which employs the Kuhn-Tucker equivalence theorem, is the key to the optimization problem and the proposed algorithm. This approach can also be applied to other optimization problems involving ordered restrictions.

Chapter 5

A One-Sided Multiple Comparison Procedure

Marcus (1978) obtained explicit one-sided simultaneous confidence intervals for monotone contrasts and nonnegative contrasts. The most recent improvements were made by Hayter (1990) which were shown to compare well with its predecessors. The purpose of this chapter is to develop a more efficient interval estimation procedure for ordered pairwise mean differences and nonnegative contrasts. In Section 5.1 a one-sided multiple comparison test (OMCT) statistic is introduced. The upper percentage points of its distribution are tabled for tail probabilities $\alpha = .10, .05$, and $.01$. The power comparisons are made with the other test procedures. In Section 5.2, a procedure is proposed to construct one-sided simultaneous confidence lower bounds. This approach makes use of the distribution of the one-sided multiple comparison test statistic. Simulation studies to compare the probabilities of detecting the differences of dosage levels by Hayter's (1990) one-sided studentized range test (OSRT) to those by the OMCT are included in Section 5.3. Our method is more efficient when the number of dosage levels is four or more. The efficiency of the OMCT procedure in some occasions may exceed that of the least significant difference (LSD) procedure - a one-sided t-test with the critical value t_ν^α. The extension of the OMCT procedure to two-sided simultaneous confidence intervals is discussed in Section 5.4. Illustrated is an application to the data of a binding inhibition assay given in Section 5.5. The proofs of the theorems are given in Section 5.6 and a conclusion is included in Section 5.7.

5.1 A One-Sided Multiple Comparison Test

5.1.1 A One-Sided Multiple Comparison Test

The dose-response curve $y = f(x)$ is to be estimated from the observations $Y_{i1}, Y_{i2}, ..., Y_{in_i}$ collected at dose level $x_i, i = 1, 2, ..., k$. Let Y_{ij} be independent normal variates with means $\mu_i = f(x_i)$ and a common unknown variance σ^2.

We are considering the problem of testing the null hypothesis $H_0 : \mu_1 = \cdots = \mu_k$ against the alternative hypothesis $H_1 : \mu_1 \leq \cdots \leq \mu_k$ with at least one strict inequality. The following one-sided multiple comparison test statistic is proposed. We reject the null hypothesis H_0 if

$$\max_{1\leq p\leq q<r\leq s\leq k} \frac{\bar{Y}_{rs} - \bar{Y}_{pq}}{S\sqrt{(\sum_{j=r}^{s} n_j)^{-1} + (\sum_{i=p}^{q} n_i)^{-1}}} \tag{5.1}$$

is large, where $\bar{Y}_i = \sum_{j=1}^{n_i} Y_{ij}/n_i, \bar{Y}_{rs} = \sum_{i=r}^{s} n_i\bar{Y}_i / \sum_{i=r}^{s} n_i, S^2 = \sum_{i,j}(Y_{ij} - \bar{Y}_i)^2/(\sum_{i=1}^{k} n_i - k)$. Its critical value $l_{\mathbf{n},k,\nu}^\alpha$ is defined by

$$P_0(\max_{1\leq p\leq q<r\leq s\leq k} \frac{\bar{Y}_{rs} - \bar{Y}_{pq}}{S\sqrt{(\sum_{j=r}^{s} n_j)^{-1} + (\sum_{i=p}^{q} n_i)^{-1}}} \leq l_{\mathbf{n},k,\nu}^\alpha) = 1 - \alpha, \tag{5.2}$$

when the means are equal, i.e., $\mu_1 = \cdots = \mu_k$, where $\nu = \sum_{i=1}^{k} n_i - k > 0$ is the degrees of freedom for S^2. For the equal sample size case, we shall use the notation $l_{k,\nu}^{\alpha}$.

There are many special cases of the OMCT statistic described by (5.2) found in the literature. They include Hayter's (1990) OSRT, Hayter's (1992) modified OSRT when $r = s$ and $p = q$, and Hirotsu, Kuriki and Hayter's (1992) maximum t method when $s = k, r = q+1$ and $p = 1$. The type of contrast used here is a comparison of μ_{rs} to μ_{pq} which includes Helmert contrasts, reverse Helmert contrasts and step contrasts (see Tamhane, Hochberg and Dunnett 1996). It is of particular interest when neighboring dosage levels have similar responses. The calculation of the critical point $l_{\mathbf{n},k,\nu}^{\alpha}$ is discussed in Section 5.1.2.

A simulation study is conducted to compare the powers of LRT, OSRT and OMCT. The powers are simulated at the 5% level of significance for $k = 4, 6$ and 9, $n_1 = n_2 = \cdots = n_k = n$, $\Delta = 1, 2, 3, 4$ and $\sigma^2/n = 1$ where the non-centrality parameter is $\Delta^2 = \sum_{i=1}^{k} n_i(\mu_i - \mu_{1k})^2$ with $\mu_{1k} = \sum_{i=1}^{k} n_i\mu_i / \sum_{i=1}^{k} n_i$. Two kinds of configurations are considered: Case I, a linear regression function; and Case II, a step regression function with a jump at a midpoint. The results are provided in Table 5.1 with 1,000,000 replications. The powers of the OMCT are much higher than those of the OSRT, particularly at large k and for Case II. They are lower than those of the LRT. These powers are the probabilities of detecting the difference between μ_k and μ_1. Both LRT and OSRT have larger powers along the linear regression curve than the step regression function. However, the OMCT has an identical power over the two regression curves. The advantage of the OMCT over the LRT is that it detects the difference between μ_j and μ_i and is used to construct simultaneous confidence lower bounds for multiple comparisons. The proof for the following theorem is given in Section 5.6.

Theorem 5.1.1 *The OMCT statistic given in* (5.1) *is consistent and unbiased. Its power function*

$$P_\mu\left(\max_{1\le p\le q<r\le s\le k} \frac{\bar{Y}_{rs}-\bar{Y}_{pq}}{S\sqrt{(\sum_{j=r}^{s} n_j)^{-1}+(\sum_{i=p}^{q} n_i)^{-1}}} \ge l^{\alpha}_{\mathbf{n},k,\nu}\right) \qquad (5.3)$$

is monotone increasing in $\mu_2-\mu_1, ..., \mu_k-\mu_{k-1}$ *with an infimum* α *attainable when* $\mu \in H_0$.

5.1.2 Calculation of the Critical Points

The acceptance region of the OMCT statistic for a fixed S is a one-sided polyhedron in $k-1$ dimensional Euclidean space bounded by $\binom{k+2}{4}$ hyperplanes. When $k=3$ and $n_1=n_2=n_3$, the probability (5.2) can be evaluated by

$$\int_0^\infty \{8P_\rho(X\le l^{\alpha}_{3,\nu}s/\sigma, Y\le 0)-3\Phi(l^{\alpha}_{3,\nu}s/\sigma)\}f(s)\,ds,$$

where $f(s)$ is the density of a random variable $(\sigma/\sqrt{\nu})(\chi^2_\nu)^{\frac{1}{2}}$ and X, Y are two standardized bivariate normal random variables having a correlation coefficient $\rho=-\tan(\pi/12)/[1+\tan^2(\pi/12)]^{1/2}$. One may evaluate the percentiles of the OMCT statistic by numerical integrations of k dimensions such as Genz (1992). For higher dimensional cases, the polyhedrons are very complicated and the accuracy of the numerical quadrature of the acceptance region is questionable. A Monte Carlo method is used to simulate the percentiles of the OMCT statistic. A FORTRAN program to calculate $l^{\alpha}_{\mathbf{n},k,\nu}$ is given in the Appendices. The result is provided in Table 5.2 for the equal sample size case with $\alpha = .1, .05$, and $.01$, $k = 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 20$, and degrees of freedom $\nu = 5, 10, 15, 20, 25, 30, 40, 60, \infty$. The accuracy employed is that the simulated tail probabilities have errors no more than 0.01.

For the unequal sample size case, the critical value will depend on the sample size ratios $n_i/n_1 (2 \leq i \leq k)$, as well as k, ν, and α. If sample sizes do not vary much, the critical value $l^{\alpha}_{k,\nu}$ for the equal sample size case can be used to approximate the value of $l^{\alpha}_{\mathbf{n},k,\nu}$. When $k = 9$ and $\mathbf{n} = (2, 2, 4, 2, 3, 3, 2, 4, 2)$, we have that $l^{.05}_{\mathbf{n},9,15} = 3.52$ which is very close to $l^{.05}_{9,15} = 3.53$; when $k = 7$ and $\mathbf{n} = (8, 4, 4, 4, 4, 4, 4)$, we have that $l^{.05}_{\mathbf{n},7,25} = 3.12$ which is also very close to $l^{.05}_{7,25} = 3.11$. Even for the case of a large variation in sample sizes, say $\mathbf{n} = (2, 2, 10, 2, 6, 13, 6, 1, 2)$, the critical value $l^{.05}_{\mathbf{n},9,15} = 3.44$ does not differ greatly from the equal weight case $l^{.05}_{9,15} = 3.53$. This illustrates the robustness of the OMCT to sample size variation, by using Table 5.2 in testing the hypothesis H_0 against H_1 and in interval estimation.

The OMCT statistic in (5.1) is bounded from below by $OSRT/\sqrt{2} = \max_{i \leq j}(\bar{Y}_j - \bar{Y}_i)/S(2/n)^{1/2}$, with critical value $h^{\alpha}_{k,\nu}/\sqrt{2}$, and is bounded from above by a statistic which has the same distribution as $\sqrt{S_{12}}$ (see Section 4.1.2). It follows that their corresponding critical values have the relationship

$$h^{\alpha}_{k,\nu}/\sqrt{2} < l^{\alpha}_{k,\nu} < \sqrt{S^{\alpha}_{12,k,\nu}}$$

for the equal sample size case. When k is small, the differences are relatively small. The difference $l^{\alpha}_{k,\nu} - h^{\alpha}_{k,\nu}/\sqrt{2}$ is a monotone-increasing function of k and a monotone-decreasing function of ν and α and these differences are provided in Table 5.3. For $\alpha = .05$, the difference lies between .04 at $k = 3, \nu = \infty$ and .34 at $k = 12, \nu = 5$ with values .05 at $k = 3, \nu = 5$ and .23 at $k = 12, \nu = \infty$. The pattern of the difference $\sqrt{S^{\alpha}_{12,k,\nu}} - l^{\alpha}_{k,\nu}$ is similar to that of $l^{\alpha}_{k,\nu} - h^{\alpha}_{k,\nu}/\sqrt{2}$, and these differences, provided in Table 5.4, are much larger. For $k = 9, \nu = 15$ and $\alpha = .05$, $l^{.05}_{9,15} - h^{.05}_{9,15}/\sqrt{2} = .21$ and $\sqrt{S^{.05}_{12,9,15}} - l^{.05}_{9,15} = .65$.

The ratios $h^{\alpha}_{k,\nu}(\sqrt{2})^{-1}/l^{\alpha}_{k,\nu}$ are provided in Table 5.5. These ratios are almost identical for each fixed k and they are monotone decreasing in k from .98 at $k = 3$ to .93 at $k = 12$. The ratios $l^{\alpha}_{k,\nu}/\sqrt{S^{\alpha}_{12,k,\nu}}$ are provided in Table 5.6. These ratios are monotone decreasing in k from .99 at $k = 3$ to .79 at $k = 12$. They are also monotone decreasing in α and monotone increasing in ν.

5.2 One-Sided Simultaneous Confidence Lower Bounds

5.2.1 One-Sided Simultaneous Confidence Lower Bounds

Let $\mu_{rs} = \sum_{i=r}^{s} n_i\mu_i/\sum_{i=r}^{s} n_i$ and $\mu_{pq} = \sum_{i=p}^{q} n_i\mu_i/\sum_{i=p}^{q} n_i$ be the mean responses at the dosage levels from r to s and from p to q respectively, where $1 \le p \le q < r \le s \le k$. We are interested in one-sided simultaneous confidence lower bounds for $\mu_{rs} - \mu_{pq}$ without assuming that $\mu_1 \le \mu_2 \le \cdots \le \mu_k$. The numerator of the OMCT statistic in (5.2) can be replaced by $(\bar{Y}_{rs} - \bar{Y}_{pq}) - (\mu_{rs} - \mu_{pq})$. The exact $100(1-\alpha)\%$ simultaneous one-sided confidence lower bounds for $\mu_{rs} - \mu_{pq}$ are as follows:

$$\mu_{rs}-\mu_{pq} \geq \bar{Y}_{rs}-\bar{Y}_{pq}-l^{\alpha}_{\mathbf{n},k,\nu}S\sqrt{(\sum_{j=r}^{s} n_j)^{-1} + (\sum_{i=p}^{q} n_i)^{-1}}, \quad \text{for all} \quad 1 \leq p \leq q < r \leq s \leq k. \tag{5.4}$$

Let $l_{pqrs}(\bar{Y})$ be the simultaneous confidence lower bound in (5.4). The positive value of $l_{pqrs}(\bar{Y})$ indicates that the mean response at the dosage level from r to s is significantly higher than the one from p to q. The simultaneous confidence lower bounds for $\mu_{rs} - \mu_{pq}$ include special cases of pairwise mean differences $\mu_j - \mu_i, i < j$, when $p = q = i, r = s = j$. That is

$$\mu_j - \mu_i \geq \bar{Y}_j - \bar{Y}_i - l^{\alpha}_{\mathbf{n},k,\nu}S\sqrt{n_j^{-1} + n_i^{-1}}. \tag{5.5}$$

Remark: One may be interested in a contrast which is a nonnegative linear combination of the ones in (5.4). For example, when $k = 4$ the linear contrast has an expression

$$3\mu_4 + \mu_3 - \mu_2 - 3\mu_1 = 3(\mu_4 - \mu_3) + 4(\mu_3 - \mu_2) + 3(\mu_2 - \mu_1).$$

The OMCT in (5.2) may be generalized by including such a linear contrast. The corresponding critical value is larger than $l^{\alpha}_{\mathbf{n},k,\nu}$. However, the increment due to the linear contrast is almost negligible. For the equal sample size case, it is no more than 0.002 when $\alpha = 0.1$ and no more than 0.004 when $\alpha = 0.05$ or 0.01.

5.2.2 Efficiency of Confidence Lower Bounds

We consider the case that σ^2 is known (i.e., $\nu = \infty$) and the sample sizes are equal. The height of the confidence lower bound (i.e., the distance the confidence lower bound extends below the difference $\mu_{rs} - \mu_{pq}$) given in (5.4) is $l^{\alpha}_{k,\infty}\sigma\sqrt{(s-r+1)^{-1}+(q-p+1)^{-1}}/\sqrt{n}$. The confidence lower bounds obtained by Marcus [1978, eq.(16)] and Hayter [1990, eq.(1.4)] are similar to those given in (5.4) except their heights are $\sigma\sqrt{\bar{\chi}^2_{12,\alpha}}\sqrt{(s-r+1)^{-1}+(q-p+1)^{-1}}/\sqrt{n}$ and $\sigma h^{\alpha}_{k,\infty}/\sqrt{n}$, respectively. The ratio of the height of the OMCT confidence lower bound to the height of the confidence lower bound given by Marcus (1978) is $l^{\alpha}_{k,\infty}/\sqrt{\bar{\chi}^2_{12,\alpha}}$. These ratios can be found in the last three rows of Table 5.6 and they lie between .99 when $k = 3$ and .80 when $k = 12$. Therefore, the OMCT procedure is more efficient than Marcus' (1978) procedure for comparing μ_{rs} to μ_{pq} .

The ratio of the height of the OSRT confidence lower bound to the height of the OMCT confidence lower bound is

$$R^{\alpha}_{k} = h^{\alpha}_{k,\infty}/\{l^{\alpha}_{k,\infty}\sqrt{(s-r+1)^{-1}+(q-p+1)^{-1}}\}.$$

Since $l^{\alpha}_{k,\nu} > h^{\alpha}_{k,\nu}/\sqrt{2}$, for ordered pairwise differences, i.e., $s = r, q = p$, the height of the OSRT confidence lower bound is shorter than that of the OMCT. But for more complicated contrasts, the converse is true. Some numerical evaluations of R^{α}_{k} are provided in Table 5.7 for $\alpha = .05$ and $k = 3,4,5,6,7,8,9,10,12$. Four types of contrasts are considered, pairwise differences $\mu_j - \mu_i$ and three more complicated comparisons $\mu_j - \mu_{i,i+1}$, $\mu_j - \mu_{i,i+2}$ and $\mu_{j-1,j} - \mu_{i,i+1}$. For complicated comparisons, the heights of the OMCT confidence lower bounds are shorter than those of the OSRT as one would expect. The reduction of the height of the OMCT confidence lower bound relative to that of the OSRT confidence lower bound can reach 27%. Hayter(1990) tabulated the ratios of its height to that of Marcus (1978), $h^{\alpha}_{k,\infty}/[\sqrt{\bar{\chi}^2_{12,\alpha}}\{(s-r+1)^{-1}+(q-p+1)^{-1}\}^{1/2}]$. These ratios are considerably smaller than the ones in Table 5.7. For instance, when $k = 8$ these ratios are .822, .949, 1.006, 1.162 (see Hayter, 1990) as compared to .943, 1.089, 1.155, 1.333 listed in Table 5.7. The OMCT procedure has the highest relative efficiency over the OSRT in detecting the difference between μ_{34} and μ_{12}. The OSRT, a Tukey-type, is efficient for pairwise comparisons, Marcus' (1978) procedure, a Scheffé-type, attains shorter bounds for more complicated contrasts and the OMCT, the one in between, has both advantages. Simultaneous confidence lower bounds can be sharpened substantially when we utilize the prior knowledge of the monotone regression curve as in the next section.

5.3 One-Sided Simultaneous Confidence Lower Bounds for Monotone Dose-Response Means

5.3.1 One-Sided Simultaneous Confidence Lower Bounds for Monotone Dose-Response Means

Simultaneous confidence lower bounds for pairwise mean differences $\mu_j - \mu_i, i < j$, are of great interest to experimenters. For a monotone nondecreasing regression curve $\mu_i = f(x_i)$,

$$\mu_j - \mu_i \geq \mu_{rs} - \mu_{pq}$$

if $i \leq p \leq q < r \leq s \leq j$. It follows from (5.4) that

$$\mu_j - \mu_i \geq l_{pqrs}(\bar{Y}).$$

The $100(1-\alpha)\%$ OMCT simultaneous confidence lower bound for $\mu_j - \mu_i$ is

$$\mu_j - \mu_i \geq \max_{i \leq p \leq q < r \leq s \leq j} l_{pqrs}(\bar{Y}). \tag{5.6}$$

It is noted that the sample means $\bar{Y}_i, \bar{Y}_{i+1}, ..., \bar{Y}_j$ have been used to construct the lower bound (5.6). The lower bound $l_{pqrs}(\bar{Y})$ which maximizes (5.6) occurs on $p \leq q < r \leq s$ with large combined sample sizes $\sum_{i=r}^{s} n_i$ and $\sum_{i=p}^{q} n_i$, and a large difference $\bar{Y}_{rs} - \bar{Y}_{pq}$. It is trivial that for $1 \leq p \leq q < r \leq s \leq k$,

$$\mu_{rs} - \mu_{pq} \geq \mu_{r's'} - \mu_{p'q'},$$

if $p' \leq q' < r' \leq s', p \leq p', q \leq q', r' \leq r, s' \leq s$. From (5.4), the $100(1-\alpha)\%$ OMCT simultaneous confidence lower bound for $\mu_{rs} - \mu_{pq}$ is

$$\mu_{rs} - \mu_{pq} \geq \max_{p \leq p' \leq q' < r' \leq s' \leq s, q \leq q', r' \leq r} l_{p'q'r's'}(\bar{Y}). \tag{5.7}$$

Let $L_{pqrs}(\bar{Y})$ be the simultaneous confidence lower bound in (5.7). By the assumption of the monotone regression curve, $\mu_j - \mu_i$ is bounded from below by zero, so is $\mu_{rs} - \mu_{pq}$. A FORTRAN program for computing the OMCT simultaneous lower bounds is given in the Appendices. For monotone dose-response curves, our primary interest lies in whether one can detect the difference between μ_j and μ_i or the difference between μ_{rs} and μ_{pq}. If the answer is affirmative, then our interest will focus on the value of the lower bound. We can apply (5.6) and (5.7) to construct OMCT simultaneous confidence lower bounds for any nonnegative contrasts as discussed in Section 5.2.1. The improvement of the simultaneous confidence lower bounds for $\mu_j - \mu_i$ and $\mu_{rs} - \mu_{pq}$ while utilizing the assumption of the monotone regression curve can also be found in Marcus (1978) and Hayter (1990). Marcus' (1978) simultaneous confidence lower bound is not as efficient as that of the OMCT.

These simultaneous confidence lower bounds are not derived from the estimated isotonic regression, but result from the sample means by utilizing the isotonic assumption on $f(x)$. Any monotone nondecreasing regression curve which satisfies (5.4) will satisfy (5.7). The coverage probability of these simultaneous confidence lower bounds (5.7) is at least $1-\alpha$ as demonstrated by the following theorem. Its proof is provided in Section 5.6.

Theorem 5.3.1 *Let the simultaneous confidence level be defined by*

$$C(\mu) = P_\mu(\mu_{rs} - \mu_{pq} \geq L_{pqrs}(\bar{Y}), \quad \text{for all} \;\; p \leq q < r \leq s).$$

Then $C(\mu)$ *is partially ordered by* μ *in the sense that* $C(\mu) \leq C(\nu)$ *if* $\mu_{i+1} - \mu_i \leq \nu_{i+1} - \nu_i$. *Therefore,*

$$\inf_{\mu_1 \leq \cdots \leq \mu_k} C(\mu) = P_0(L_{pqrs}(\bar{Y}) \leq 0, \quad \text{for all} \;\; p \leq q < r \leq s) = 1 - \alpha$$

and the infimum is attainable when $\mu \in H_0$.

In the next two subsections, we investigate the behavior of the OMCT and the OSRT procedures under monotone regression curves using simulation studies. For simplicity, the studies are restricted to the equal sample size case with $\alpha = .05$ and $\sigma^2/n = 1$.

5.3.2 Pairwise Comparisons

In this subsection, we will study whether the procedures will be able to detect the difference between μ_j and μ_i at a confidence level $1-\alpha$.

The OMCT critical value $l^{\alpha}_{k,\nu}$ is larger than $h^{\alpha}_{k,\nu}/\sqrt{2}$ for every k, ν and α. Hence the lower bound of OSRT for $\mu_{i+1} - \mu_i$ is larger than that of OMCT. However, the situation for $\mu_j - \mu_i$ with $j - i \geq 3$ could be quite different. For example, if we are interested in the confidence lower bound for $\mu_4 - \mu_1$, the OMCT procedure will compare the confidence lower bound of $\mu_4 - \mu_1$ not only with those of $\mu_2 - \mu_1, \mu_3 - \mu_1, \mu_3 - \mu_2, \mu_4 - \mu_2, \mu_4 - \mu_3$ as does the OSRT, but also with the confidence lower bounds of $\mu_{23} - \mu_1, \mu_{34} - \mu_1, \mu_{24} - \mu_1, \mu_3 - \mu_{12}, \mu_4 - \mu_{12}, \mu_{34} - \mu_{12}, \mu_4 - \mu_{13}$, $\mu_{34} - \mu_2$ and $\mu_4 - \mu_{23}$. Furthermore, the height of the OMCT confidence lower bound for $\mu_{rs} - \mu_{pq}$, $l^{\alpha}_{k,\nu}\sigma\sqrt{(s-r+1)^{-1} + (q-p+1)^{-1}}/\sqrt{n}$, is shorter than the corresponding height of the OSRT if $r < s$ or $p < q$ as shown in Table 5.7.

The OMCT confidence lower bound on $f(x_j) - f(x_i), j > i$, will substantially improve the OSRT confidence lower bound when $j - i$ is large. The situation in which the OMCT bounds are most advantageous is when there exist p, q, r, s with $i \leq p < q < r < s \leq j$ such that $f(x_p) = \cdots = f(x_q)$ and $f(x_r) = \cdots = f(x_s)$. The situation in which the OMCT bounds are less advantageous is when $f(x_{i+1}) - f(x_i) \geq \delta, i = 1, ..., k-1$ for a large positive δ.

A simulation study is conducted to compare the efficiency of the new procedure to that of the OSRT procedure. The 95% simultaneous confidence lower bounds are computed by generating 1,000,000 sets of normal variates. The percentages of detecting the difference between level j and level i are computed for the two procedures. Two cases are considered, the linear regression function, $\mu_i = \delta i$ for Case I and the step regression function $\mu_1 = \cdots = \mu_{[k/2]} = 0, \mu_{[k/2]+1} = \cdots = \mu_k = \delta$ for the Case II. The results for comparing μ_j to μ_i are provided in Table 5.8.

By (5.5) and equation (1.2) in Hayter (1990), the probability of detecting the difference between μ_j and μ_1 without the assumption of the monotone regression curve is $\Phi[(\mu_j - \mu_1)/\sqrt{2} - c_{k,\infty}^{.05}]$ where Φ is the distribution function of a standard normal random variable and $c_{k,\infty}^{.05} = h_{k,\infty}^{.05}/\sqrt{2}$ for the OSRT and $c_{k,\infty}^{.05} = l_{k,\infty}^{.05}$ for the OMCT. The probabilities in Table 5.8 are considerably larger than those obtained without the assumption. Consider the comparison of μ_k and μ_1. For Case I and $\Delta = 4$ we have .907 versus $\Phi(20/\sqrt{35} - h_{6,\infty}^{.05}/\sqrt{2}) = .772$ and .947 versus $\Phi(20/\sqrt{35} - l_{6,\infty}^{.05}) = .729$ when $k = 6$; .822 versus $\Phi(16/\sqrt{30} - h_{9,\infty}^{.05}/\sqrt{2}) = .507$ and .924 versus $\Phi(16/\sqrt{30} - l_{9,\infty}^{.05}) = .429$ when $k = 9$. For Case II and $\Delta = 4$ we have .881 versus $\Phi(4/\sqrt{3} - h_{6,\infty}^{.05}/\sqrt{2}) = .373$ and .947 versus $\Phi(4/\sqrt{3} - l_{9,\infty}^{.05}) = .333$ when $k = 6$, and .784 versus $\Phi(6/\sqrt{10} - h_{9,\infty}^{.05}/\sqrt{2}) = .157$ and .922 versus $\Phi(6/\sqrt{10} - l_{9,\infty}^{.05}) = .115$ when $k = 9$. The increase in probability by the OSRT is due to the extra $\binom{k}{2} - 1$ comparisons. The gain by the OMCT is much larger. It is due to the extra $\binom{k+2}{4} - 1$ comparisons, and to the inequality

$$P[L_{11kk}(\bar{Y}) > 0] \geq max_{1 \leq p \leq q < r \leq s \leq k} P[l_{pqrs}(\bar{Y}) > 0]. \tag{5.8}$$

For Case I and $\Delta = 4$, the right hand side of (5.8) is $P[l_{1256}(\bar{Y}) > 0] = \Phi(32/\sqrt{70} - l_{6,\infty}^{.05}) = .854$ when $k = 6$ and it is $P[l_{1379}(\bar{Y}) > 0] = \Phi(12/\sqrt{10} - l_{9,\infty}^{.05}) = .756$ when $k = 9$. For Case II and $\Delta = 4$, it is $P[l_{1346}(\bar{Y}) > 0] = \Phi(4 - l_{6,\infty}^{.05}) = .891$ when $k = 6$ and it is $P[l_{1459}(\bar{Y}) > 0] = \Phi(4 - l_{9,\infty}^{.05}) = .816$ when $k = 9$. These probabilities $P[l_{pqrs}(\bar{Y}) > 0]$ calculated without the monotone regression curve assumption are the lower bounds for the probability of detecting the difference between μ_k and μ_1 by the OMCT procedure. It is noted that for Case II, $\mu_j - \mu_i = 0$ if $i < j \leq [k/2]$ and the difference $\mu_j - \mu_i$ is a constant if $i \leq [k/2] < j$. The increase in the probabilities $P[L_{iijj}(\bar{Y}) > 0]$ in j is due to the assumption of a monotone regression curve.

These probabilities for comparing μ_k to μ_1 are the same as the powers of the two tests in Table 5.1 when $\Delta = 4$. Therefore, the probabilities of detecting the difference between μ_k and μ_1 can be found in Table 5.1 for $k = 4, 6$, and 9, $\Delta = 1, 2, 3, 4$ with the linear regression function or the step regression function. The OMCT procedure has higher probabilities of detecting the difference between μ_k and μ_1 than the OSRT procedure. The improvement increases for large k.

The OSRT procedure is more efficient than the OMCT procedure in detecting the difference between μ_2 and μ_1 but less efficient for comparing μ_k and μ_1. Table 5.8 indicates that for a fixed i, when the probability is small or j is small, the OSRT is more efficient and when the probability is large or j is large, the OMCT is more efficient. When the difference $\mu_j - \mu_1$ is detectable, the OMCT should normally be used. For Case I, the linear regression, the probability of detecting the difference between μ_j and μ_i, $i < j$, is the same as the probability between μ_{j-i+1} and μ_1. For Case II, the step regression function, the probability of detecting the difference between μ_{k-i} and μ_j is the same as the probability between μ_{k-j} and μ_i for $i + j < k$ when $k = 9$ and the probability of detecting the difference between μ_{k+1-i} and μ_j is the same as the probability between μ_{k+1-j} and μ_i for $i + j \leq k$ when $k = 6$.

The OMCT procedure may perform favorably against the least significant difference (LSD) procedure. The probability of detecting the difference between μ_j and μ_i by the latter procedure at 95% confidence level is $\Phi[(\mu_j - \mu_i)/\sqrt{2} - 1.645]$. We observed that the probabilities for the OMCT procedure may exceed the corresponding LSD procedure. They include the comparisons of $\mu_5 - \mu_1, \mu_6 - \mu_2, \mu_6 - \mu_1$ when $\Delta = 3.46$, $\mu_5 - \mu_1, \mu_6 - \mu_1, \mu_5 - \mu_2$ and $\mu_6 - \mu_2$ when $\Delta = 4$ in Case II at $k = 6$ and $\mu_9 - \mu_1$ in Case I at $k = 9$, and $\mu_6 - \mu_1, \mu_7 - \mu_1, \mu_8 - \mu_1, \mu_9 - \mu_1, \mu_7 - \mu_2, \mu_8 - \mu_2, \mu_9 - \mu_2, \mu_8 - \mu_3$ and $\mu_9 - \mu_3$ in Case II at $k = 9$. This superiority will also be seen in the numerical example in Section 5.5.

Also included in Table 5.8 are regression functions with $\Delta = 2.16$ and $\Delta = 3.46$ respectively when $k = 6$. They are part of the regression functions of Case I and Case II respectively when $k = 9$ and $\Delta = 4$. For Case I, the probabilities for the case $k = 6$ and $\Delta = 2.16$ are larger than the corresponding ones for the case $k = 9$ and $\Delta = 4$. This is because the former use the critical values $h^{.05}_{6,\infty} = 3.725$ and $l^{.05}_{6,\infty} = 2.77$, while the latter use the values $h^{.05}_{9,\infty} = 4.107$ and $l^{.05}_{9,\infty} = 3.09$. Similar results hold true for Case II, but comparisons are made between $\mu_j - \mu_i$ when $k = 6$ and $\mu_{j+1} - \mu_{i+1}$ when $k = 9$. One may also compare the results of the same type of the regression curve with two different Δ's when $k = 6$.

It is of interest to compare the mean heights of simultaneous confidence lower bounds when the probability of detecting the difference between μ_j and μ_i is high by both procedures. The probability that both the OMCT and the OSRT can detect the difference in the means indicates that both procedures succeed in detecting the difference in μ_j and μ_i simultaneously. The mean height is the distance between the lower bound and $\mu_j - \mu_i$. Our prior knowledge of the monotone regression curve indicates that $\mu_j \geq \mu_i$ if $j > i$. Therefore, the simultaneous confidence lower bound for $\mu_j - \mu_i$ is always non-negative and it is positive if there is a significant difference between μ_j and μ_i at a confidence level $1 - \alpha$. The mean heights of 95% simultaneous confidence lower bounds for $\mu_j - \mu_i$ by the OMCT and the OSRT procedures are provided in Table 5.9 for the case that the probability of detecting the difference between μ_j and μ_i by both procedures is at least 60%.

Comparing these probabilities with the corresponding ones in Table 5.8, it can be seen that these probabilities are less than the ones obtained by the OSRT by no more than .015, but they are less than the ones by the OMCT by at least .044 if $k = 6$ and .106 if $k = 9$. The OMCT mean heights are smaller than their counterparts of the OSRT. The larger the difference between j and i, the larger the difference will be between the two mean heights. The reduction in the mean height by the OMCT over that of OSRT can be as large as .24 (13.5%). The Pittman efficiency for the mean height is the ratio of squared mean heights as stated in Schoenfeld (1986). The ratio of the OSRT mean height squared compared to that of the OMCT can reach 106% for the linear regression curve and 124% for the step regression function when $k = 6$. It can reach 113% for linear regression curve and 134% for the step regression curve when $k = 9$. The OMCT procedure is generally preferable to the OSRT procedure when k is large and the dose-response curve increases moderately.

5.3.3 Comparing Two Categories of Dosage Levels

By (5.4) and equation (1.4) in Hayter (1990), the probability of detecting the difference between a mean response μ_{rs} of the dosage levels from r to s and a mean response μ_{pq} of the dosage levels from p to q without the assumption of the monotone regression curve is $\Phi[(\mu_{rs} - \mu_{pq})/\|C\| - l^{.05}_{k,\infty}]$ for the OMCT procedure and $\Phi[(\mu_{rs} - \mu_{pq} - h^{.05}_{k,\infty})/\|C\|]$ for the OSRT procedure where $\|C\|^2 = (r - s + 1)^{-1} + (q - p + 1)^{-1}$. If any of the two categories consists of more than one dosage level then the former probability is larger than the latter. The difference may be quite large. For example, when $k = 9, \Phi(\mu_{89} - \mu_{12} - l^{.05}_{9,\infty}) = .697$ and $\Phi(\mu_{89} - \mu_{12} - h^{.05}_{9,\infty}) = .311$ for Case I, and they are .338 and .077 respectively for OMCT and OSRT for Case II. A simulation study is conducted to investigate their behaviors when the regression curve is monotone. Three types of comparisons, $\mu_j - \mu_{12}, \mu_j - \mu_{13}$ and $\mu_{j-1,j} - \mu_{12}$, are considered for $k = 6$ and 9 for Case I, the linear regression curve, and Case II, the step regression function, when $\Delta = 4$. The results are provided in Table 5.10.

The probabilities are much larger than the ones without the monotone assumption, particularly when $k = 9$, Case II, and by the OMCT procedure. The OMCT procedure performs overwhelmingly better than the OSRT procedure except for the few occasions when probabilities are extremely low. The difference in probabilities can be as large as .283. These probabilities are bounded from above by the corresponding ones for $\mu_j - \mu_1$ in Table 5.8, and the probabilities by the OMCT are uniformly closer to their upper bounds than the ones by the OSRT. They are bounded from below by the corresponding probabilities for $\mu_j - \mu_2, \mu_j - \mu_3$ and $\mu_{j-1} - \mu_2$ respectively in Table 5.10 for the three types of comparisons.

The probability of detecting the difference between μ_{rs} and μ_{pq} by the OMCT may exceed the one by the LSD. For the step regression function in Case II, they include the comparisons of $\mu_5 - \mu_{12}, \mu_6 - \mu_{12}, \mu_6 - \mu_{13}$ when $k = 6$, and $\mu_7 - \mu_{12}, \mu_8 - \mu_{12}, \mu_9 - \mu_{12}, \mu_7 - \mu_{13}, \mu_8 - \mu_{13}, \mu_9 - \mu_{13}, \mu_{78} - \mu_{12}$ and $\mu_{89} - \mu_{12}$ when $k = 9$.

When the probability of detecting the difference between μ_{rs} and μ_{pq} is at least 60% by both OMCT and OSRT, the mean heights of their simultaneous confidence lower bounds were computed. The results are provided in Table 5.11. Comparing these probabilities with the corresponding ones in Table 5.10, it is found that these probabilities are less than the ones by the OSRT by no more than .008, but they are less than the ones by the OMCT by at least .092. The mean heights of 95% simultaneous confidence lower bound by the OMCT are uniformly shorter than those by the OSRT. The reduction in mean height by the OMCT over the OSRT can be as large as .37(12.9%). The Pittman efficiency for the mean height of the OSRT compared to that of the OMCT can reach 133% for the linear regression curve and 139% for the step regression curve when $k = 6$; and they are 132% and 144% respectively, when $k = 9$.

5.4 Extension to Simultaneous Confidence Intervals

The ideas behind the multiple comparison procedure can also be used to construct simultaneous confidence intervals. An extension of the OMCT procedure to simultaneous confidence interval is as follows. The test statistic

$$M = \max_{1\le p\le q<r\le s\le k} \frac{|\bar{Y}_{rs} - \bar{Y}_{pq}|}{S\sqrt{(\sum_{j=r}^{s} n_j)^{-1} + (\sum_{i=p}^{q} n_i)^{-1}}}$$

is used. Let $m^{\alpha}_{\mathbf{n},k,\nu}$ be the critical value of M. The exact $100(1-\alpha)\%$ simultaneous confidence intervals for the multiple comparison $\mu_{rs} - \mu_{pq}$ are

$$\begin{aligned} & \bar{Y}_{rs} - \bar{Y}_{pq} - m^{\alpha}_{\mathbf{n},k,\nu} S\sqrt{(\sum_{j=r}^{s} n_j)^{-1} + (\sum_{i=p}^{q} n_i)^{-1}} \\ \le\ & \mu_{rs} - \mu_{pq} \\ \le\ & \bar{Y}_{rs} - \bar{Y}_{pq} + m^{\alpha}_{\mathbf{n},k,\nu} S\sqrt{(\sum_{j=r}^{s} n_j)^{-1} + (\sum_{i=p}^{q} n_i)^{-1}}. \end{aligned}$$

The $100(1-\alpha)\%$ simultaneous confidence intervals for $\mu_j - \mu_i$ by Tukey-Kramer (TK) procedure are

$$P\{\mu_j - \mu_i \in [\bar{Y}_j - \bar{Y}_i \pm q^{\alpha}_{k,\nu} S\sqrt{\frac{1}{2}(\frac{1}{n_j} + \frac{1}{n_i})}];\ 1 \le i,j \le k\} \ge 1 - \alpha, \qquad (5.9)$$

where $q^{\alpha}_{k,\nu}$ is the critical value of the studentized range statistic (see Hayter 1986). It also can be generalized to more complicated nonnegative contrasts.

If the common variance σ^2 is known and the sample sizes are equal, the ratio of the mean lengths (i.e, the difference of the confidence upper bound and the confidence lower bound) of the generalized OMCT confidence intervals and TK confidence intervals is

$$m^{\alpha}_{\mathbf{n},k,\infty}\sqrt{(s-r+1)^{-1} + (q-p+1)^{-1}}/q^{\alpha}_{k,\infty}.$$

When $k = 9$, we study the same four types of contrasts $\mu_j - \mu_i, \mu_j - \mu_{i,i+1}, \mu_j - \mu_{i,i+2}$ and $\mu_{j-1,j} - \mu_{i,i+1}$ as in Section 5.2.2. The corresponding ratios are provided in Table 5.12. The more complicated the contrasts are, the more reduction we obtain by the generalized OMCT procedure.

If we utilize the prior knowledge that $\mu_1 \le \mu_2 \le \cdots \le \mu_k$, the $100(1-\alpha)\%$ simultaneous confidence intervals for $\mu_j - \mu_i, 1 \le i < j \le k$, can be improved to

$$\begin{aligned} & \max_{i\le p\le q<r\le s\le j} \{\bar{Y}_{rs} - \bar{Y}_{pq} - m^{\alpha}_{\mathbf{n},k,\nu} S \sqrt{(\sum_{j=r}^{s} n_j)^{-1} + (\sum_{i=p}^{q} n_i)^{-1}}\} \\ \le \; & \mu_j - \mu_i \\ \le \; & \min_{p\le q\le i<j\le r\le s} \{\bar{Y}_{rs} - \bar{Y}_{pq} + m^{\alpha}_{\mathbf{n},k,\nu} S \sqrt{(\sum_{j=r}^{s} n_j)^{-1} + (\sum_{i=p}^{q} n_i)^{-1}}\}. \end{aligned}$$

As in Section 5.3, the $100(1-\alpha)\%$ simultaneous confidence intervals for the multiple comparisons $\mu_{rs} - \mu_{pq}$ are

$$\begin{aligned} & \max_{p\le p'\le q'<r'\le s'\le s, q\le q', r'\le r} \{\bar{Y}_{r's'} - \bar{Y}_{p'q'} - m^{\alpha}_{\mathbf{n},k,\nu} S \sqrt{(\sum_{j=r'}^{s'} n_j)^{-1} + (\sum_{i=p'}^{q'} n_i)^{-1}}\} \\ \le \; & \mu_{rs} - \mu_{pq} \\ \le \; & \min_{p'\le q'\le q<r\le r'\le s', p'\le p, s'\le s} \{\bar{Y}_{r's'} - \bar{Y}_{p'q'} + m^{\alpha}_{\mathbf{n},k,\nu} S \sqrt{(\sum_{j=r'}^{s'} n_j)^{-1} + (\sum_{i=p'}^{q'} n_i)^{-1}}\} \end{aligned}$$

The critical value $m^{\alpha}_{\mathbf{n},k,\nu}$ is a little larger than the corresponding $l^{\alpha}_{\mathbf{n},k,\nu}$. For example $m^{0.05}_{\mathbf{n},9,15} = 3.82$, where $\mathbf{n} = (2,2,4,2,3,3,2,4,2)$, whereas $l^{0.05}_{\mathbf{n},9,15} = 3.52$.

One may use a conservative two-sided simultaneous confidence interval procedure as in Berk and Marcus (1996),

$$P\{\mu_{rs} - \mu_{pq} \in \bar{Y}_{rs} - \bar{Y}_{pq} \pm l^{\alpha/2}_{\mathbf{n},k,\nu} S \sqrt{(\sum_{j=r'}^{s'} n_j)^{-1} + (\sum_{i=p'}^{q'} n_i)^{-1}}\} \ge 1 - \alpha. \qquad (5.10)$$

By comparing the table values $l_{\mathbf{n},9,15}^{0.05} = 3.52$ and $m_{\mathbf{n},9,15}^{0.10} = 3.40$, we can see our generalized OMCT approach is more efficient.

5.5 A Numerical Example

The data given in Table 3.1 from a binding inhibition assay which was described fully by Kanowith-Klein, Vitetta, Korn, and Ashman (1979) will be studied here. In this set of data, there are $k = 9$ different dilutions of one antiserum and 24 observations were made. The pooled estimate of variance, S^2, is 86.48 with $\nu = 15$ degrees of freedom. To test the null hypothesis $H_0 : \mu_1 = \cdots = \mu_9$ against all alternatives, the usual overall F-test statistic is $F = 7.40$ and it has a p-value 0.0005. The null hypothesis is rejected and the means $\mu_1, \cdots, \mu_9$ are not all equal.

The behavior of these means without the assumption of monotone regression curve is of considerable interest. The scatterplot in Figure 5.1 indicates that there are no differences among the six levels, level 4, 5, 6, 7, 8 and 9, of high doses. The upper percentage points are $l_{\mathbf{n},9,15}^{.05} = 3.52$ and $l_{\mathbf{n},9,15}^{.10} = 3.11$. Hayter's OSRT procedure applies only to the equal sample size case. One can generalize it by using the statistic

$$\max_{1 \le i < j \le k} \sqrt{2} \frac{\bar{Y}_j - \bar{Y}_i}{S\sqrt{\frac{1}{n_i} + \frac{1}{n_j}}}$$

with critical values $h_{\mathbf{n},9,15}^{.05} = 4.68$ and $h_{\mathbf{n},9,15}^{.10} = 4.13$. Both procedures detect the difference between the group of levels 1, 2, and 3 and the group of levels 4, 5, 6, 7, 8, and 9. The 95% OMCT simultaneous confidence lower bound for

$\mu_{49} - \mu_{13}$ is 13.20 while the corresponding OSRT simultaneous confidence lower bound is 2.09, a difference of 11.11. The OMCT also detects the difference between the group of levels 2 and 3 and the group of levels 4, 5, 6, 7, 8 and 9, while the OSRT fails to do so. The 95% OMCT simultaneous confidence lower bound for $\mu_{49} - \mu_{23}$ is 5.33. Furthermore, the OMCT detects the difference between level 1 and the group of levels 2 and 3, but again the OSRT fails to do so. The 90% OMCT simultaneous confidence lower bound for $\mu_{23} - \mu_1$ is 1.89. Marcus' simultaneous confidence lower bound can be computed similarly as (5.4) with $l^{\alpha}_{\mathbf{n},k,\nu}$ replaced by $\sqrt{S^{\alpha}_{12,\mathbf{n},k,\nu}}$, where $\sqrt{S^{.05}_{12,\mathbf{n},9,15}} = 4.16$ and $\sqrt{S^{.10}_{12,\mathbf{n},9,15}} = 3.69$. Marcus' simultaneous confidence lower bounds are always less efficient than those of the OMCT.

Consider the one-sided test of $H_0 : \mu_1 = \cdots = \mu_9$ against $H_1 : \mu_1 \leq \cdots \leq \mu_9$ with at least one strict inequality. The OSRT test statistic is $H = \max_{1 \leq i < j \leq k} \sqrt{2}(\bar{Y}_j - \bar{Y}_i)/(S\sqrt{1/n_j + 1/n_i})$, with the maximum occurring at $i = 1$ and $j = 8$ and its value is $\sqrt{2}(\bar{Y}_8 - \bar{Y}_1)/(S\sqrt{1/2 + 1/4}) = 8.47$. The p-value is 0.0003. Utilizing the OMCT statistic (5.1), the maximum of the test statistic $L = \max_{1 \leq p \leq q < r \leq s \leq k}(\bar{Y}_{rs} - \bar{Y}_{pq})/[S\sqrt{(\sum_{j=r}^{s} n_j)^{-1} + (\sum_{i=p}^{q} n_i)^{-1}}]$ occurs at $p = 1, q = 3, r = 4$ and $s = 9$ and its value is $(\bar{Y}_{49} - \bar{Y}_{13})/(S\sqrt{1/16 + 1/8}) = 6.80$. Its corresponding p-value is 0.0001. The value of the LRT test statistic is 54.16 with p-value 0.0000. The null hypothesis is rejected at significant level $\alpha = 0.05$ by all three tests. The LRT is the best, and the OMCT procedure is more powerful than the OSRT.

From the scatterplot in Figure 5.1, one can see that percentage inhibition is monotone in the levels of dilution. Based on a monotone regression curve, the

95% OMCT simultaneous confidence lower bounds for $\mu_j - \mu_i, i < j, i = 1, 2, 3$, and those of OSRT and Marcus' (1978) $\max F_R$ are provided in Table 5.14. There are no significant differences between any two levels of the high dosage categories, levels 4 to 9.

It is found that the OMCT is the most efficient test in comparing μ_j to μ_i. The differences $\mu_5 - \mu_2, \mu_8 - \mu_3$ and $\mu_9 - \mu_3$ can be detected by the OMCT procedure, while they failed to be detected by the OSRT and Marcus' (1978) procedures. However, the difference $\mu_3 - \mu_1$ can be detected by the OSRT, but neither the OMCT nor Marcus' (1978) procedures could detect this difference. It is noted that the OMCT simultaneous confidence lower bounds are larger than those of Marcus (1978). The latter is a Scheffé-type procedure which is known to be less efficient for pairwise comparisons.

The efficiency of the OMCT simultaneous confidence lower bounds can also be examined by comparing to the LSD one-sided confidence lower bounds. The OMCT simultaneous confidence lower bound for $\mu_7 - \mu_2$ is 3.18, which is the simultaneous confidence lower bound for $\mu_{45} - \mu_{23}$ without the monotone assumption, while the LSD confidence lower bound for $\mu_7 - \mu_2$ is $\bar{Y}_7 - \bar{Y}_2 - t_{15}^{0.05} S / \sqrt{\frac{1}{2} + \frac{1}{2}} = 3.20$. The OMCT simultaneous confidence lower bound in this case is approximately the same as the corresponding LSD confidence lower bound.

The OMCT procedure indicates that in general the dilution levels can be classified into a low inhibition percentage category (level 1) , a high inhibition percentage category (level 4, 5, 6, 7, 8, and 9) and an in-between inhibition percentage category (level 2 and 3). However, there is no significant difference between the means of the six levels in the high inhibition percentage category and there is also no significant difference between the means of the two levels within the in-between percentage category.

With the monotone assumption, the generalized OMCT simultaneous confidence intervals for the numerical example is given in Table 5.15 where the critical point $m^{0.05}_{\mathbf{n},9,15} = 3.82$. The 95% generalized OMCT simultaneous confidence intervals for pairwise mean differences also show the difference in high dose levels from the low and in-between dose levels. In addition, the difference between the low dose level and in-between level can not be detected by the lower bound of contrast $\mu_{23} - \mu_1$ which is -3.51. However, the lower bounds for $\mu_{49} - \mu_{23}$ and $\mu_{49} - \mu_{13}$ are 3.99 and 11.99 respectively. The differences between the high dose levels and low, in-between dose levels are detectable by the generalized OMCT procedure as well.

5.6 Proof of Theorem 5.1.1 and Theorem 5.3.1

Proof of Theorem 5.3.1:

It suffices to consider the case that

$$\nu_i = \mu_i, i = 1, 2, ..., t \text{ and } \nu_i = \mu_i + \delta, i = t + 1, ..., k$$

for an index $t, 1 \leq t < k$, and for an arbitrary positive real number δ. Let $\bar{X}_1, ..., \bar{X}_k$ have identical distributions as $\bar{Y}_1, ..., \bar{Y}_k$ and let $\bar{X}_i^\delta = \bar{X}_i, i = 1, ..., t$, $\bar{X}_i^\delta = \bar{X}_i + \delta, i = t + 1, ..., k$. For each $p \leq q < r \leq s$, let $\gamma = (\nu_{rs} - \nu_{pq}) - (\mu_{rs} - \mu_{pq})$. We shall establish that

$$\gamma \geq L_{pqrs}(\bar{X}^\delta) - L_{pqrs}(\bar{X}) \tag{5.11}$$

It follows that

$$\nu_{rs} - \nu_{pq} - L_{pqrs}(\bar{X}^\delta) \geq \mu_{rs} - \mu_{pq} - L_{pqrs}(\bar{X})$$

and hence

$$\begin{aligned} C(\nu) &= P(\nu_{rs} - \nu_{pq} \geq L_{pqrs}(\bar{X}^\delta) \quad \text{for all} \quad 1 \leq p \leq q < r \leq s \leq k) \\ &\geq P(\mu_{rs} - \mu_{pq} \geq L_{pqrs}(\bar{X}) \quad \text{for all} \quad 1 \leq p \leq q < r \leq s \leq k) \\ &= P(\mu_{rs} - \mu_{pq} \geq L_{pqrs}(\bar{Y}) \quad \text{for all} \quad 1 \leq p \leq q < r \leq s \leq k) \\ &= C(\mu) \end{aligned}$$

The inequality (5.11) is implied by

$$l_{p'q'r's'}(\bar{X}^\delta) - l_{p'q'r's'}(\bar{X}) \leq \gamma \tag{5.12}$$

for all $p' \leq q' < r' \leq s'$ with $p \leq p', q \leq q', r' \leq r, s' \leq s$. This is because

$$\begin{aligned} L_{pqrs}(\bar{X}^\delta) - L_{pqrs}(\bar{X}) &= \max_{p \leq p' \leq q' < r' \leq s' \leq s, q \leq q', r' \leq r} l_{p'q'r's'}(\bar{X}^\delta) \\ &\quad - \max_{p \leq p' \leq q' < r' \leq s' \leq s, q \leq q', r' \leq r} l_{p'q'r's'}(\bar{X}) \\ &= l_{p_0q_0r_0s_0}(\bar{X}^\delta) - \max_{p \leq p' \leq q' < r' \leq s' \leq s, q \leq q', r' \leq r} l_{p'q'r's'}(\bar{X}) \\ &\leq l_{p_0q_0r_0s_0}(\bar{X}^\delta) - l_{p_0q_0r_0s_0}(\bar{X}) \\ &\leq \max_{p \leq p' \leq q' < r' \leq s' \leq s, q \leq q', r' \leq r} [l_{p'q'r's'}(\bar{X}^\delta) - l_{p'q'r's'}(\bar{X})] \end{aligned}$$

where $p_0 \leq q_0 < r_0 \leq s_0$ are such that

$$l_{p_0q_0r_0s_0}(\bar{X}^\delta) = \max_{p \leq p' \leq q' < r' \leq s' \leq s, q \leq q', r' \leq r} l_{p'q'r's'}(\bar{X}^\delta).$$

Consider the following four cases.

(I) $t < p$ or $s \leq t$:

Here, $\gamma = (\nu_{rs} - \nu_{pq}) - (\mu_{rs} - \mu_{pq}) = 0$ and so does

$$l_{p'q'r's'}(\bar{X}^{\delta}) - l_{p'q'r's'}(\bar{X}) = 0 \leq \gamma.$$

(II) $p \leq t < q$:

Here, $\gamma = (\nu_{rs} - \nu_{pq}) - (\mu_{rs} - \mu_{pq}) = \delta - \frac{\sum_{i=t+1}^{q} n_i}{\sum_{i=p}^{q} n_i}\delta = \frac{\sum_{i=p}^{t} n_i}{\sum_{i=p}^{q} n_i}\delta$

If $t < p'$ then $l_{p'q'r's'}(\bar{X}^{\delta}) - l_{p'q'r's'}(\bar{X}) = 0 \leq \gamma$.

Otherwise, $p' \leq t$ and

$$l_{p'q'r's'}(\bar{X}^{\delta}) - l_{p'q'r's'}(\bar{X}) = \frac{\sum_{i=p'}^{t} n_i}{\sum_{i=p'}^{q'} n_i}\delta.$$

However,

$$\frac{\sum_{i=p'}^{t} n_i}{\sum_{i=p'}^{q'} n_i} \leq \frac{\sum_{i=p'}^{t} n_i}{\sum_{i=p'}^{q} n_i} \leq \frac{\sum_{i=p}^{t} n_i}{\sum_{i=p}^{q} n_i}$$

and (5.12) is satisfied.

(III) $q \leq t < r$:

Here, $\gamma = (\nu_{rs} - \nu_{pq}) - (\mu_{rs} - \mu_{pq}) = \delta$. It is trivial that

$$l_{p'q'r's'}(\bar{X}^{\delta}) \leq l_{p'q'r's'}(\bar{X}) + \delta$$

and hence (5.12) is satisfied.

(IV) $r \leq t < s$:

Here, $\gamma = (\nu_{rs} - \nu_{pq}) - (\mu_{rs} - \mu_{pq}) = \frac{\sum_{i=t+1}^{s} n_i}{\sum_{i=r}^{s} n_i}\delta$. If $t < s'$, then

$$l_{p'q'r's'}(\bar{X}^{\delta}) = l_{p'q'r's'}(\bar{X}) + \frac{\sum_{i=t+1}^{s'} n_i}{\sum_{i=r'}^{s'} n_i}\delta.$$

However,

$$\frac{\sum_{i=t+1}^{s'} n_i}{\sum_{i=r'}^{s'} n_i} \leq \frac{\sum_{i=t+1}^{s'} n_i}{\sum_{i=r}^{s'} n_i} \leq \frac{\sum_{i=t+1}^{s} n_i}{\sum_{i=r}^{s} n_i}$$

and (5.12) is satisfied. Otherwise, $s' \leq t$ in which case $l_{p'q'r's'}(\bar{X}^\delta) = l_{p'q'r's'}(\bar{X})$ and this completes the proof. □

Proof of Theorem 5.1.1*:*

By the proof of Theorem 5.3.1,

$$\begin{aligned} & P_\mu(\mu_{rs} - \mu_{pq} \geq \max_{p \leq p' \leq q' < r' \leq s' \leq s, q \leq q', r' \leq r} l_{p'q'r's'}(\bar{Y}) \quad \text{for all } 1 \leq p \leq q < r \leq s \leq k) \\ \leq \; & P_\nu(\nu_{rs} - \nu_{pq} \geq \max_{p \leq p' \leq q' < r' \leq s' \leq s, q \leq q', r' \leq r} l_{p'q'r's'}(\bar{X}^\delta) \quad \text{for all } 1 \leq p \leq q < r \leq s \leq k) \end{aligned}$$

if $\mu_{i+1} - \mu_i \leq \nu_{i+1} - \nu_i, i = 1, ..., k-1$. Since the event on the left hand side of the above inequality is equivalent to the event

$$[\mu_{rs} - \mu_{pq} \geq l_{pqrs}(\bar{Y}) \quad \text{for all } 1 \leq p \leq q < r \leq s \leq k]$$

under the monotone nondecreasing regression function, and that event is equivalent to the event

$$[\max_{1 \leq p \leq q < r \leq s \leq k} \frac{\bar{Y}_{rs} - \bar{Y}_{pq}}{S\sqrt{(\sum_{i=p}^{q} n_i)^{-1} + (\sum_{j=r}^{s} n_j)^{-1}}} \geq l^\alpha_{\mathbf{n},k,\nu}],$$

the monotonicity of the power function (5.3) is established. Consequently,

$$\begin{aligned} & \inf_\mu P_\mu(\max_{1 \leq p \leq q < r \leq s \leq k} \frac{\bar{Y}_{rs} - \bar{Y}_{pq}}{S\sqrt{(\sum_{j=r'}^{s'} n_j)^{-1} + (\sum_{i=p'}^{q'} n_i)^{-1}}} \geq l^\alpha_{\mathbf{n},k,\nu}) \\ = \; & P_{H_0}(\max_{1 \leq p \leq q < r \leq s \leq k} \frac{\bar{Y}_{rs} - \bar{Y}_{pq}}{S\sqrt{(\sum_{j=r}^{s} n_j)^{-1} + (\sum_{i=p}^{q} n_i)^{-1}}} \geq l^\alpha_{\mathbf{n},k,\nu}) \\ = \; & \alpha \end{aligned}$$

and the test is unbiased.

Let $\Delta = \max_{1\le p\le q<r\le s\le k} \frac{\mu_{rs}-\mu_{pq}}{\sigma\sqrt{(\sum_{j=r}^{s} n_j)^{-1}+(\sum_{i=p}^{q} n_i)^{-1}}}$. If $\mu \in H_1 - H_0$ then there exist $p \le q < r \le s$ such that $\frac{\mu_{rs}-\mu_{pq}}{\sigma\sqrt{(\sum_{j=r}^{s} n_j)^{-1}+(\sum_{i=p}^{q} n_i)^{-1}}} > 0$ and hence $\Delta > 0$. Since the one-sided t-test

$$\frac{\bar{Y}_{rs} - \bar{Y}_{pq}}{S\sqrt{(\sum_{j=r}^{s} n_j)^{-1} + (\sum_{i=p}^{q} n_i)^{-1}}} \ge c$$

is consistent, so is our OMCT statistic. □

5.7 Conclusion

If experimenters have a prior reason to believe that the regression curve is monotone nondecreasing, then a test procedure can be chosen to have good power properties under this ordered alternative. The inversion of the test procedure results in a set of simultaneous confidence intervals for various contrasts of the means (Hayter 1990).

The multiple comparison procedure proposed in this chapter is a simple and effective method for constructing one-sided simultaneous confidence lower bounds for multiple comparisons. The OMCT simultaneous confidence lower bounds are compared favorably to those of the OSRT simultaneous confidence lower bounds as the latter does not fully utilize all the observed information. When differences between the means $\mu_i \le \cdots \le \mu_j$ are small, it is advantageous to use weighted average means $\sum_{\alpha=i}^{j} n_\alpha \bar{Y}_{\alpha.} / \sum_{\alpha=i}^{j} n_\alpha$ in the inference procedure, see Wright (1982). The OMCT procedure is most advantageous when the regression curve $f(x)$ does not increase rapidly in one or more intervals of dosage levels. Without the prior knowledge of the monotonicity of the response curves, the OMCT lower bound is the most effective method to categorize the dosage levels into different response groups as shown in the above numerical example. Applied to the dose-response curves, the OMCT procedure tends to have sharper confidence lower bounds than the OSRT procedure for the pairwise mean differences $\mu_j - \mu_i$ when $j - i$ is large. It must be stressed that these confidence lower bounds are valid only when the ordering is specified prior to observations of the data and hence is independent of the data.

Table 5.1: The Powers of the OMCT, the OSRT and the LRT at $\alpha = 0.05$ and $\nu = \infty$

		$k=4$			$k=6$			$k=9$		
Configuration	Δ	OSRT	OMCT	LRT	OSRT	OMCT	LRT	OSRT	OMCT	LRT
	1	.173	.186	.239	.143	.163	.234	.121	.144	.230
I^a	2	.455	.487	.594	.365	.428	.586	.289	.377	.578
	3	.785	.814	.885	.679	.758	.879	.561	.702	.874
	4	.957	.967	.985	.907	.947	.983	.822	.924	.983
	1	.167	.184	.212	.138	.162	.200	.117	.143	.192
II^b	2	.440	.489	.545	.350	.429	.515	.276	.379	.493
	3	.761	.814	.856	.647	.758	.832	.529	.703	.812
	4	.945	.967	.979	.881	.947	.972	.784	.922	.966

[a] I: $(1,2,3,4)\Delta/\sqrt{5}$ for $k=4$, $(1,2,3,4,5,6)\Delta/\sqrt{35/2}$ for $k=6$ and $(1,2,3,4,5,6,7,8,9)\Delta/\sqrt{60}$ for $k=9$.

[b] II: $(0,0,1,1)\Delta$ for $k=4$, $(0,0,0,1,1,1)\Delta/\sqrt{3/2}$ for $k=6$ and $(0,0,0,0,1,1,1,1,1)\Delta/\sqrt{20/9}$ for $k=9$.

Table 5.2: Upper Percentage Points for One-Sided Multiple Comparison Test

		k										
ν	α	3	4	5	6	7	8	9	10	12	15	20
5	.10	2.20	2.65	3.00	3.27	3.49	3.68	3.86	4.00	4.25	4.55	4.92
	.05	2.79	3.30	3.69	4.01	4.25	4.47	4.68	4.83	5.11	5.46	5.88
	.01	4.36	5.02	5.54	5.97	6.29	6.60	6.89	7.10	7.49	7.96	8.55
10	.10	1.98	2.35	2.63	2.84	3.01	3.16	3.29	3.41	3.60	3.83	4.12
	.05	2.42	2.81	3.09	3.31	3.49	3.65	3.79	3.91	4.11	4.36	4.66
	.01	3.41	3.83	4.15	4.40	4.61	4.78	4.93	5.08	5.30	5.59	5.95
15	.10	1.91	2.26	2.51	2.71	2.87	3.01	3.13	3.23	3.41	3.62	3.87
	.05	2.32	2.66	2.91	3.11	3.28	3.42	3.53	3.64	3.82	4.04	4.31
	.01	3.16	3.51	3.78	4.00	4.16	4.31	4.44	4.55	4.75	4.99	5.25
20	.10	1.88	2.22	2.46	2.65	2.81	2.93	3.05	3.15	3.32	3.51	3.76
	.05	2.27	2.59	2.83	3.03	3.18	3.30	3.42	3.52	3.69	3.89	4.14
	.01	3.05	3.38	3.63	3.81	3.97	4.10	4.21	4.32	4.50	4.70	4.97
25	.10	1.86	2.20	2.43	2.62	2.77	2.89	3.00	3.10	3.26	3.45	3.69
	.05	2.23	2.56	2.79	2.97	3.11	3.24	3.35	3.45	3.61	3.80	4.04
	.01	2.99	3.30	3.54	3.70	3.84	3.98	4.09	4.19	4.35	4.55	4.79
30	.10	1.85	2.18	2.41	2.59	2.74	2.86	2.93	3.06	3.22	3.41	3.64
	.05	2.20	2.53	2.76	2.94	3.08	3.20	3.30	3.40	3.56	3.74	3.97
	.01	2.95	3.24	3.47	3.64	3.79	3.90	4.00	4.11	4.25	4.44	4.67
40	.10	1.84	2.16	2.39	2.56	2.71	2.83	2.93	3.03	3.18	3.36	3.58
	.05	2.19	2.50	2.73	2.89	3.03	3.15	3.25	3.34	3.49	3.67	3.89
	.01	2.90	3.18	3.40	3.56	3.69	3.81	3.91	3.99	4.15	4.32	4.53
60	.10	1.83	2.14	2.36	2.54	2.68	2.79	2.90	2.98	3.13	3.31	3.52
	.05	2.17	2.47	2.69	2.85	2.99	3.10	3.20	3.28	3.43	3.60	3.81
	.01	2.84	3.12	3.32	3.48	3.61	3.72	3.81	3.89	4.03	4.20	4.40
∞	.10	1.80	2.10	2.32	2.48	2.61	2.72	2.82	2.90	3.04	3.21	3.41
	.05	2.12	2.41	2.61	2.77	2.90	3.00	3.09	3.17	3.31	3.46	3.65
	.01	2.75	3.01	3.19	3.30	3.44	3.55	3.63	3.70	3.82	3.96	4.14

Table 5.3: Differences of Upper Percentage Points Between OMCT and OSRT/$\sqrt{2}$

		k								
ν	α	3	4	5	6	7	8	9	10	12
5	.10	.05	.08	.13	.16	.18	.20	.24	.25	.28
	.05	.05	.10	.15	.20	.22	.25	.30	.30	.34
	.01	.09	.15	.22	.29	.31	.36	.43	.44	.49
10	.10	.04	.07	.11	.14	.16	.18	.20	.22	.24
	.05	.05	.10	.13	.16	.18	.21	.23	.25	.28
	.01	.07	.13	.18	.22	.25	.27	.29	.32	.36
15	.10	.03	.07	.10	.13	.15	.17	.19	.20	.24
	.05	.05	.09	.12	.15	.18	.20	.21	.23	.26
	.01	.06	.11	.15	.20	.21	.24	.26	.28	.32
20	.10	.03	.07	.10	.13	.15	.16	.19	.20	.23
	.05	.05	.08	.11	.15	.17	.18	.21	.22	.25
	.01	.06	.11	.16	.18	.21	.23	.24	.27	.30
25	.10	.03	.08	.10	.13	.15	.17	.18	.20	.23
	.05	.04	.09	.12	.14	.16	.18	.20	.22	.25
	.01	.07	.11	.16	.17	.19	.22	.24	.26	.29
30	.10	.03	.07	.10	.12	.15	.16	.18	.19	.22
	.05	.03	.08	.12	.15	.17	.18	.19	.22	.25
	.01	.07	.10	.15	.17	.21	.22	.23	.27	.28
40	.10	.04	.07	.10	.12	.15	.17	.18	.20	.22
	.05	.04	.08	.12	.14	.16	.18	.20	.21	.24
	.01	.07	.10	.15	.17	.19	.22	.23	.24	.28
60	.10	.04	.07	.10	.13	.15	.16	.18	.19	.21
	.05	.05	.08	.12	.14	.16	.18	.20	.21	.23
	.01	.05	.10	.13	.16	.19	.21	.22	.24	.26
∞	.10	.04	.07	.10	.12	.14	.15	.17	.18	.21
	.05	.04	.08	.11	.14	.16	.17	.19	.20	.23
	.01	.06	.10	.13	.15	.17	.20	.22	.23	.25

Table 5.4: Differences of Upper Percentage Points Between $\sqrt{S_{12}}$ and OMCT

		k								
ν	α	3	4	5	6	7	8	9	10	12
5	.10	.01	.09	.19	.32	.46	.60	.74	.90	1.20
	.05	.02	.11	.23	.36	.54	.71	.87	1.07	1.43
	.01	.03	.17	.34	.55	.81	1.04	1.27	1.54	2.07
10	.10	.01	.08	.16	.27	.39	.51	.63	.75	.99
	.05	.02	.08	.18	.30	.43	.56	.69	.83	1.11
	.01	.02	.11	.23	.38	.53	.71	.88	1.03	1.39
15	.10	.02	.08	.16	.26	.37	.47	.59	.70	.93
	.05	.01	.08	.18	.29	.40	.52	.65	.77	1.02
	.01	.02	.11	.21	.34	.48	.62	.76	.91	1.20
20	.10	.02	.07	.16	.25	.35	.47	.57	.68	.89
	.05	.01	.09	.18	.27	.38	.51	.62	.73	.96
	.01	.02	.10	.19	.32	.45	.58	.72	.84	1.10
30	.10	.02	.07	.15	.25	.34	.45	.55	.66	.87
	.05	.03	.08	.16	.26	.37	.48	.60	.70	.92
	.01	.02	.10	.19	.30	.41	.54	.67	.77	1.03
40	.10	.01	.07	.15	.24	.34	.44	.54	.64	.85
	.05	.02	.08	.15	.26	.37	.47	.58	.69	.90
	.01	.01	.11	.18	.29	.41	.52	.64	.76	.98
60	.10	.01	.07	.15	.23	.33	.44	.53	.64	.84
	.05	.01	.08	.15	.26	.35	.46	.56	.67	.87
	.01	.02	.09	.19	.29	.39	.50	.62	.73	.95
∞	.10	.01	.07	.14	.23	.33	.42	.52	.62	.81
	.05	.02	.07	.16	.25	.34	.44	.54	.64	.83
	.01	.02	.08	.17	.30	.37	.46	.57	.67	.87

Table 5.5: Ratios of Upper Percentage Points of OSRT/$\sqrt{2}$ to OMCT

		k								
ν	α	3	4	5	6	7	8	9	10	12
5	.10	.978	.970	.958	.952	.949	.945	.938	.937	.932
	.05	.981	.969	.958	.949	.948	.943	.937	.936	.934
	.01	.980	.969	.960	.952	.951	.945	.938	.938	.934
10	.10	.979	.968	.957	.951	.948	.944	.940	.936	.932
	.05	.980	.965	.957	.951	.947	.942	.938	.936	.932
	.01	.979	.967	.957	.951	.945	.943	.941	.936	.935
15	.10	.983	.969	.960	.952	.948	.942	.938	.937	.931
	.05	.977	.967	.960	.953	.946	.942	.941	.937	.933
	.01	.981	.970	.959	.950	.949	.944	.941	.939	.933
20	.10	.982	.967	.959	.952	.945	.944	.938	.935	.930
	.05	.977	.969	.960	.949	.945	.943	.939	.936	.932
	.01	.980	.967	.956	.952	.947	.944	.942	.938	.932
25	.10	.983	.965	.959	.950	.945	.943	.939	.935	.931
	.05	.982	.965	.958	.951	.949	.943	.940	.936	.932
	.01	.978	.967	.955	.954	.951	.944	.941	.937	.933
30	.10	.982	.967	.959	.952	.946	.943	.939	.937	.932
	.05	.987	.967	.957	.950	.946	.943	.941	.936	.931
	.01	.977	.969	.958	.952	.946	.945	.942	.935	.935
40	.10	.980	.967	.957	.953	.945	.941	.939	.934	.930
	.05	.981	.967	.955	.952	.947	.943	.940	.937	.933
	.01	.977	.968	.957	.952	.949	.944	.940	.939	.932
60	.10	.978	.967	.959	.950	.945	.943	.938	.936	.932
	.05	.979	.966	.956	.951	.945	.942	.938	.936	.932
	.01	.981	.968	.960	.953	.948	.943	.941	.939	.935
∞	.10	.979	.967	.956	.951	.947	.943	.938	.937	.932
	.05	.982	.967	.959	.951	.945	.943	.940	.937	.931
	.01	.980	.966	.958	.962	.950	.942	.940	.939	.935

Table 5.6: Ratios of Upper Percentage Points of OMCT to $\sqrt{S_{12}}$

		k								
ν	α	3	4	5	6	7	8	9	10	12
5	.10	.997	.967	.941	.912	.884	.859	.839	.817	.780
	.05	.993	.969	.942	.917	.887	.862	.843	.819	.781
	.01	.994	.968	.942	.916	.886	.863	.845	.821	.784
10	.10	.994	.968	.943	.914	.886	.862	.840	.821	.784
	.05	.993	.972	.944	.916	.889	.866	.845	.825	.787
	.01	.993	.971	.947	.921	.896	.871	.849	.831	.793
15	.10	.991	.968	.939	.913	.887	.864	.842	.821	.786
	.05	.996	.969	.941	.915	.892	.869	.845	.825	.790
	.01	.992	.969	.945	.923	.896	.874	.853	.833	.800
20	.10	.991	.969	.940	.913	.890	.863	.843	.823	.788
	.05	.996	.968	.942	.919	.893	.867	.847	.828	.793
	.01	.993	.972	.949	.922	.899	.876	.854	.836	.803
30	.10	.991	.969	.940	.913	.889	.864	.843	.823	.788
	.05	.986	.970	.944	.919	.893	.869	.847	.829	.795
	.01	.995	.970	.948	.923	.902	.877	.856	.841	.804
40	.10	.993	.969	.942	.913	.890	.866	.843	.826	.790
	.05	.992	.970	.947	.917	.892	.870	.847	.830	.795
	.01	.995	.966	.949	.924	.900	.880	.860	.840	.809
60	.10	.996	.969	.940	.916	.891	.865	.846	.824	.789
	.05	.994	.970	.946	.918	.895	.871	.851	.830	.797
	.01	.991	.971	.947	.924	.902	.881	.861	.842	.809
∞	.10	.995	.969	.943	.915	.888	.865	.845	.825	.790
	.05	.991	.970	.943	.918	.896	.872	.851	.832	.800
	.01	.993	.973	.949	.917	.902	.885	.865	.847	.814

Table 5.7: Ratios of the Height of the OSRT Simultaneous Lower Bound to the Height of the OMCT Simultaneous Lower Bound for Various Contrast **C** and $\alpha = .05$

	k								
C	3	4	5	6	7	8	9	10	12
$\mu_j - \mu_i$	.982	.967	.959	.951	.945	.943	.940	.937	.931
$\mu_j - \mu_{i,i+1}$	1.133	1.116	1.107	1.098	1.091	1.089	1.085	1.082	1.075
$\mu_j - \mu_{i,i+2}$	NA	1.184	1.174	1.165	1.157	1.155	1.151	1.147	1.140
$\mu_{j-1,j} - \mu_{i,i+1}$	NA	1.367	1.356	1.345	1.336	1.333	1.329	1.325	1.316

Table 5.8: Probabilities of Detecting the Difference Between μ_j and μ_i by 95% One-Sided Simultaneous Confidence Lower Bounds

k		Δ	i		j							
					2	3	4	5	6	7	8	9
6	I	2.16	1	OSRT	.012	.046	.118	.240	.412			
				OMCT	.008	.040	.120	.269	.482			
		4	1	OSRT	.025	.128	.364	.680	.907			
				OMCT	.018	.116	.380	.734	.947			
	II	3.46	1	OSRT	.004	.012	.436	.624	.727			
				OMCT	.003	.009	.436	.707	.831			
			2	OSRT		.004	.354	.524	.624			
				OMCT		.003	.340	.579	.707			
		4	1	OSRT	.004	.012	.620	.802	.881			
				OMCT	.003	.009	.624	.872	.947			
			2	OSRT		.004	.530	.713	.802			
				OMCT		.003	.518	.772	.872			
9	I	4	1	OSRT	.006	.024	.067	.151	.285	.465	.659	.822
				OMCT	.003	.018	.062	.161	.334	.564	.782	.924
	II	4	1	OSRT	.002	.005	.010	.372	.554	.663	.734	.784
				OMCT	.001	.004	.008	.361	.655	.805	.881	.922
			2	OSRT		.002	.005	.320	.487	.592	.663	.715
				OMCT		.001	.004	.303	.569	.720	.805	.856
			3	OSRT			.002	.253	.394	.487	.555	.605
				OMCT			.001	.225	.436	.569	.655	.713

Table 5.9: Probabilities of Detecting the Difference by Both OMCT and OSRT and Their Mean Heights of 95% One-Sided Simultaneous Confidence Lower Bounds for $\mu_j - \mu_i$ when $\Delta = 4$

				Mean Height	
μ	k	Contrast	Probability	OSRT	OMCT
I	6	$\mu_6 - \mu_1$	.903	3.17	3.08
		$\mu_6 - \mu_2, \mu_5 - \mu_1$	.665	3.00	2.97
	9	$\mu_9 - \mu_1$	.818	3.09	2.91
		$\mu_9 - \mu_2, \mu_8 - \mu_1$	.646	2.90	2.79
II	6	$\mu_6 - \mu_1$	.878	1.97	1.77
		$\mu_6 - \mu_2, \mu_5 - \mu_1$	.796	2.19	2.06
		$\mu_5 - \mu_2$	.702	2.38	2.31
	9	$\mu_9 - \mu_1$	.780	1.78	1.54
		$\mu_8 - \mu_1$	.728	1.87	1.66
		$\mu_9 - \mu_2$	.707	1.91	1.72
		$\mu_8 - \mu_2, \mu_7 - \mu_1$	.654	1.99	1.82

Table 5.10: Probabilities of Detecting the Difference $\mu_{rs} - \mu_{pq}$ by 95% One-Sided Simultaneous Confidence Lower Bounds for Various Comparisons when $\Delta = 4$

Comparison	k	μ		j					
				4	5	6	7	8	9
$\mu_j - \mu_{12}$	6	I	OSRT	.204	.525	.838			
			OMCT	.294	.676	.933			
		II	OSRT	.576	.766	.853			
			OMCT	.607	.865	.943			
	9	I	OSRT	.031	.087	.195	.363	.569	.763
			OMCT	.042	.130	.297	.531	.763	.916
		II	OSRT	.006	.343	.520	.629	.703	.756
			OMCT	.005	.350	.646	.799	.877	.920
$\mu_j - \mu_{13}$	6	I	OSRT	.080	.337	.707			
			OMCT	.136	.534	.887			
		II	OSRT	.500	.694	.793			
			OMCT	.568	.844	.934			
	9	I	OSRT	.011	.044	.122	.262	.462	.678
			OMCT	.016	.079	.227	.462	.717	.896
		II	OSRT	.002	.305	.473	.582	.657	.712
			OMCT	.002	.328	.627	.787	.869	.914
$\mu_{j-1j} - \mu_{12}$	6	I	OSRT	.068	.330	.722			
			OMCT	.208	.613	.915			
		II	OSRT	.051	.679	.820			
			OMCT	.149	.851	.940			
	9	I	OSRT	.009	.039	.115	.257	.463	.685
			OMCT	.025	.101	.263	.501	.743	.908
		II	OSRT	.002	.035	.416	.576	.670	.734
			OMCT	.003	.087	.627	.793	.874	.918

Table 5.11: Probabilities of Detecting the Difference by Both OMCT and OSRT and Their Mean Heights of 95% One-Sided Simultaneous Confidence Lower Bounds for $\mu_{rs} - \mu_{pq}$ when $\Delta = 4$

				Mean Height	
μ	k	Contrast	Probability	OSRT	OMCT
I	6	$\mu_6 - \mu_{12}$	.836	3.07	2.80
		$\mu_6 - \mu_{13}$	.704	2.91	2.66
		$\mu_{56} - \mu_{12}$	.721	2.93	2.54
	9	$\mu_9 - \mu_{12}$	.760	2.95	2.67
		$\mu_9 - \mu_{13}$	.674	2.87	2.56
		$\mu_{89} - \mu_{12}$	.683	2.87	2.50
II	6	$\mu_5 - \mu_{12}$	.761	2.28	2.10
		$\mu_6 - \mu_{12}$	.851	2.07	1.81
		$\mu_5 - \mu_{13}$	.691	2.42	2.19
		$\mu_6 - \mu_{13}$	.792	2.23	1.90
		$\mu_{45} - \mu_{12}$	.676	2.45	2.17
		$\mu_{56} - \mu_{12}$	.818	2.18	1.85
	9	$\mu_7 - \mu_{12}$	.621	2.04	1.84
		$\mu_8 - \mu_{12}$	.698	1.93	1.68
		$\mu_9 - \mu_{12}$	.753	1.84	1.56
		$\mu_8 - \mu_{13}$	.653	2.00	1.72
		$\mu_9 - \mu_{13}$	.709	1.92	1.60
		$\mu_{78} - \mu_{12}$	.666	1.98	1.70
		$\mu_{89} - \mu_{12}$	.731	1.89	1.58

Table 5.12: Ratio of the Heights of the OMCT Simultaneous Lower Bounds to the Heights of the TK Simultaneous Lower Bounds for Various Contrast **C** with $\alpha = .05$

C	k									
	3	4	5	6	7	8	10	12	15	20
$\mu_j - \mu_i$	1.020	1.035	1.045	1.053	1.058	1.062	1.068	1.074	1.079	1.084
$\mu_j - \mu_{i,i+1}$	0.883	0.897	0.905	0.912	0.916	0.920	0.925	0.930	0.935	0.938
$\mu_j - \mu_{i,i+2}$	0.833	0.845	0.853	0.860	0.864	0.868	0.872	0.877	0.881	0.885
$\mu_{j-1,j} - \mu_{i,i+1}$	0.721	0.732	0.739	0.744	0.748	0.751	0.755	0.759	0.763	0.766

Table 5.13: The 95% $\max F_R$, OMCT and OSRT Simultaneous Confidence Lower Bounds for $\mu_j - \mu_i, i < j$

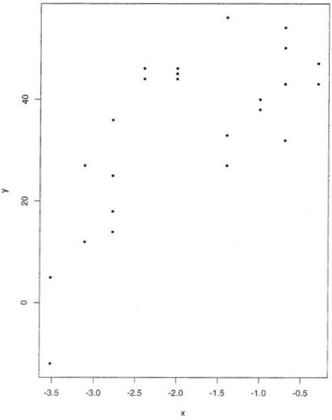
i		j							
		2	3	4	5	6	7	8	9
1	$\max F_R$	n[a]	n	9.79	16.11	16.31	16.41	18.18	18.79
	OSRT	n	0.10	17.73	20.41	20.41	20.41	21.60	21.60
	OMCT	n	n	15.77	21.11	21.11	21.11	21.47	21.95
2	$\max F_R$		n	n	n	0.38	0.69	2.72	3.40
	OSRT		n	n	n	n	n	n	n
	OMCT		n	n	3.18	3.18	3.18	4.74	5.33
3	$\max F_R$			n	n	n	n	n	n
	OSRT			n	n	n	n	n	n
	OMCT			n	n	n	n	0.91	1.45

[a] n represents the difference can not be detected

Table 5.14: 95% Two-Sided OMCT Simultaneous Confidence Intervals for $\mu_j - \mu_1, 1 \leq i < j \leq 9$

	Upper Bound									
	1	2	3	4	5	6	7	8	9	j/i
1		54.51	57.51	72.02	72.02	72.02	74.96	77.34	84.02	1
2	-12.52		34.51	49.02	49.02	49.02	51.96	54.34	61.02	2
3	-3.51	-27.01		38.01	38.05	38.12	40.56	43.34	52.01	3
4	12.98	-6.01	-9.01		23.52	23.52	26.46	28.84	35.52	4
5	18.78	1.49	-2.08	-32.43		16.25	18.63	21.34	29.72	5
6	18.78	1.49	-2.08	-32.17	-35.34		18.51	21.34	29.72	6
7	18.78	1.56	-2.08	-28.67	-31.84	-32.10		21.28	29.72	7
8	19.36	3.38	-0.68	-17.88	-17.88	-17.88	-25.01		29.14	8
9	19.86	3.99	-0.11	-15.41	-15.48	-15.48	-23.17	-30.51		
j/i	1	2	3	4	5	6	7	8		
	Lower bound									

Figure 5.1: Scatterplot of the Data of Binding Inhibition Assay



Chapter 6

Simultaneous Statistical Inference with a Control

6.1 Introduction

In drug development studies, several increasing dose levels of a substance are usually compared with the zero-dose control to investigate the effect of the substance. For this purpose, a dose-response experiment is often conducted in which the doses of the substance under consideration are administered to separate groups of subjects. There are many applications when the dose-response curve is monotone. Our first concern is whether there exists one response mean which is better than the zero-dose control mean. If so, we will be interested in identifying the lowest dose level that produces a desirable effect over that of the zero-dose control.

Specifically, we assume that we have the responses $Y_{ij}(i = 0, 1, ..., k, j = 1, ..., n_i)$ from k dose levels and a control $(i = 0)$. The sample means $\bar{Y}_0, ..., \bar{Y}_k$ are normally distributed with means μ_i and variances σ^2/n_i. For our first concern, as we know that the response means $\mu_i, i = 1, ..., k$, are at least as effective as the control mean μ_0, and a natural strategy in the statistical analysis is to test the hypothesis $H_0 : \mu_0 = \mu_1 = \cdots = \mu_k$ against the one-sided alternative that at least one response mean μ_i is better than the control, i.e., $H_1^{st} : \mu_0 \leq \mu_i, i = 1, ..., k$, with at least one inequality. This one-sided alternative is a well known simple tree order restriction. A variety of test procedures have been proposed and the majority are based on one or more contrasts among the sample means. The best known is Dunnett's (1955) multiple comparison procedure. Dunnett's approach has the advantage of providing confidence limits for the differences between the response mean and the control mean, but the case of unequal sample sizes prevents the use of the existing table values of Dunnett's test statistic. There is no basic theoretical reason requiring the number of observations in each of $k + 1$ dose levels to be equal. In fact, it would be more appealing to allow the control to have more observations than the other k dose levels. An alternative to Dunnett's test is the likelihood ratio test (LRT) by Bartholomew (1959a, 1959b, 1961a, 1961b). As the null distribution of the LRT also depends on the sample sizes, implementing this test is difficult in practice. Abelson and Tukey (1963) and Schaafsma and Smid (1966) developed the single contrast tests with high power at the center of the alternative region but a very low power at the edge of this region that is generally far below that of the LRT (Robertson, Wright and Dykstra 1988). Mukerjee, Robertson and Wright (1987) proposed a family of orthogonal contrasts which includes Dunnett's and the aforementioned single contrast as special cases. Tang and Lin (1997) proposed a LRT based on an orthant approximation and the generalizations of the orthogonal contrast test of Mukerjee, Robertson and

Wright (1987) was recently studied by McDermott (1999).

Usually, a more restrictive order, a simple order, is considered in dose-response studies when prior knowledge indicates that the response means are monotone nondecreasing with the dose levels and are at least as effective as the control, i.e. $H_1 : \mu_0 \leq \mu_1 \leq \cdots \leq \mu_k$. The related tests of equality of μ_i against H_1 can be found in Section 2.2. For the monotone dose-response means, we are also interested in identifying the dose level i such that any other dose levels higher than i will be more efficacious than the control simultaneously. The difference of the response mean with that of the control is evaluated by the interval estimate. With the monotone assumption, the lower bound for $\mu_j - \mu_0$ will be nonnegative. A positive lower bound for $\mu_j - \mu_0$ indicates that the response mean μ_j is larger than the control mean μ_0. By the LRT statistic for the simple order alternative, Marcus and Peritz (1976) obtained explicit one-sided simultaneous confidence intervals for monotone contrasts, $\sum_{i=0}^{k} n_i c_i \mu_i$, for which $\sum_{i=0}^{k} n_i c_i = 0$ and $c_0 \leq c_1 \leq \cdots \leq c_k$. Utilizing the properties of the dual cone of the simple order cone, Marcus (1978) studied the confidence lower bounds for nonnegative combinations of pairwise mean comparisons with the application to both the simple order and the simple tree ordering assumptions. Berk and Marcus (1996) gave a review of the work of the simultaneous bounds for partially ordered means.

In this chapter, we will propose a new procedure which outperforms its predecessors and is invariant with respect to sample sizes. In Section 6.2, we introduce a simultaneous inference procedure that will be used in our study. In Section 6.3, a new test statistic will be presented and power comparisons are conducted. In Section 6.4, a one-sided optimal simultaneous confidence lower bound for pairwise mean differences $\mu_j - \mu_0$ is proposed. Also included are an algorithm to compute this optimal simultaneous lower bound and a numerical example. Technical results can be found in Section 6.5. A discussion is given in Section 6.6.

6.2 Simultaneous Inference Procedure

Let $\mu_0, \mu_1, \cdots, \mu_k$ be dose-response means at dose level i with level 0 as the control. We assume that $\mu_0 \leq \mu_1 \leq \cdots \leq \mu_k$. In order to identify the minimum dose level which has a desirable effect, we consider the null hypothesis H_0 : $\mu_0 = \mu_1 = \cdots = \mu_k$ against the alternative hypothesis $H_1 : \mu_0 \leq \mu_1 \leq \cdots \leq \mu_k$ with at least one strict inequality. If H_0 is rejected, we conclude that that $\mu_k > \mu_0$. It is of considerable interest to identify the smallest dose level j such that $\mu_r > \mu_0$, $r \geq j$, simultaneously. For example, when the response means satisfy $\mu_0 = \mu_1 = \mu_2 < \mu_3 \leq \mu_4$, one would like to identify simultaneously $\mu_3 > \mu_0$ and $\mu_4 > \mu_0$. This can be achieved by simultaneous tests and the simultaneous confidence lower bound for $\mu_j - \mu_0$.

6.2.1 Dunnett's Procedure

Dunnett (1955) proposed the test statistic

$$D_k = \max_{1 \leq i \leq k} \{(\bar{Y}_i - \bar{Y}_0)/\{s(n_0^{-1} + n_i^{-1})^{1/2}\}$$

for testing H_0 against $H_1^{st} : \mu_0 \leq \mu_i, i = 1, ..., k$ with at least one strict inequality, where $s^2 = \sum_{i,j}(Y_{ij} - \bar{Y}_i)^2/\nu$ and $\nu = \sum_{i=0}^{k} n_i - (k+1) > 0$. The critical value for D_k is denoted by $d_{\alpha,k,\nu}$. If H_0 is rejected for large values of D_k, one concludes that there exists a level $i \le k$ such that $\mu_i > \mu_0$. Incorporating the prior knowledge that $\mu_0, \mu_1, ..., \mu_k$ are monotone, one would also conclude that $\mu_k > \mu_0$. The smallest level j such that $\mu_r > \mu_0$ for any $r \ge j$ can be found by testing $H_{0j} : \mu_0 = \mu_1 = \cdots = \mu_j$ against $H_{1j}^{st} : \mu_0 \le \mu_i (i = 1, ..., j)$ with at least one strict inequality, $j = 1, ..., k$, simultaneously. As D_j has the property that $D_1 \le D_2 \le \cdots \le D_k$, if $D_j > d_{\alpha,k,\nu}$ where

$$D_j = \max_{1 \le i \le j} \{(\bar{Y}_i - \bar{Y}_0)/\{s(n_0^{-1} + n_i^{-1})^{1/2}\},$$

one rejects H_{0j} and concludes that $\mu_r > \mu_j$ for all $r \ge j$. With the assumption that $\mu_0 \le \mu_1 \le \cdots \le \mu_k$, the one-sided simultaneous confidence lower bound for $\mu_j - \mu_0$ is constructed as

$$L^d(\mu_j - \mu_0) = \max_{1 \le i \le j} \{\bar{Y}_i - \bar{Y}_0 - d_{\alpha,k,\nu} s(n_0^{-1} + n_i^{-1})^{1/2}\}. \tag{6.1}$$

Note that $L^d(\mu_j - \mu_0) > 0$ implies $L^d(\mu_r - \mu_0) > 0$ for any $r \ge j$. Furthermore, for a given α, $L^d(\mu_j - \mu_0) > 0$ is equivalent to $D_j > d_{\alpha,k,\nu}$.

6.2.2 Modified Likelihood Ratio Test for the Simple Tree Alternative

An excellent alternative to Dunnett's procedure is the modified likelihood ratio test (MLRT) considered by Wright (1988) for testing H_0 against H_1^{st}. The MLRT T_k^{st} rejects H_0 for large values of

$$T_k^{st} = \{\sum_{i=0}^{k} n_i(\mu_i^{st} - \hat{\mu}_i)^2/s^2\}^{1/2}$$

where $\hat{\mu} = \sum_{i=0}^k n_i \bar{Y}_i / \sum_{i=0}^k n_i$ is the MLE of the common population mean under H_0 and $\mu_i^{st}(i = 0, ..., j)$ are the restricted MLE of μ_i's under the simple tree order alternative. For the simple tree alternative, Thompson's minimum-violator algorithm provides a convenient method for computing the estimate μ_i^* (Thompson 1962). Hogg (1965) discussed the relationship between the likelihood ratio function and the class of linear functions of the sample mean $\bar{Y}_i$. It follows that

$$T_k^{st} = \max_{c_0 \le c_i}\{\sum_{i=0}^{k} n_i c_i \bar{Y}_i / (s^2 \sum_{i=0}^{k} n_i c_i^2)^{1/2}\}.$$

If H_0 is rejected for large values of T_k^{st}, one concludes that $\mu_k > \mu_0$. By testing H_{0j} against H_{1j}^{st} simultaneously, we conclude $\mu_j > \mu_0$ if H_{0j} is rejected. That is, if

$$T_j^{st} = \max_{c_0 \le c_i}\{\sum_{i=0}^{j} n_i c_i \bar{Y}_i / (s^2 \sum_{i=0}^{j} n_i c_i^2)^{1/2}\} > t_{\alpha,k,\nu}^{st},$$

where $t_{\alpha,k,\nu}^{st}$ is the critical value for T_k^{st}. Since $T_1^{st} \le T_2^{st} \le \cdots \le T_k^{st}$, one concludes that $\mu_r > \mu_0$ for all $r \ge j$. The simultaneous confidence lower bound for $\mu_j - \mu_0$ is constructed as

$$L^{st}(\mu_j - \mu_0) = \max_{c_0 \le c_i, \sum_{i=0}^{j} n_i c_i \mu_i \le \mu_j - \mu_0} \{\sum_{i=0}^{j} n_i c_i \bar{Y}_i - t_{\alpha,k,\nu}^{st} s (\sum_{i=0}^{j} n_i c_i^2)^{1/2}\}. \quad (6.2)$$

We have noticed that the test procedures by D_k and T_k^{st} are designed to test the homogeneity of the response means against the simple tree order alternative, however they do not fully utilize the prior knowledge that $\mu_i, i = 0, ..., k$, are monotone nondecreasing.

6.2.3 Modified Likelihood Ratio Test for the Simple Order Alternative

Wright (1988) also proposed the MLRT T_k^{so} to test H_0 against $H_1 : \mu_0 \le \mu_1 \le \cdots \le \mu_k$ with at least one strict inequality. The null hypothesis H_0 is rejected for large values of

$$T_k^{so} = \{\sum_{i=0}^{k} n_i(\mu_i^{so} - \hat{\mu}_i)^2/s^2\}^{1/2}.$$

Here $\mu_i^{so}(i = 0, ..., k)$ are the restricted MLE of μ_i under the simple order alternative which can be computed by the pool-adjacent-violator algorithm (see Section 2.1). In a similar manner as T_k^{st}, the statistic T_k^{so} can be formatted as

$$T_k^{so} = \max_{\mathbf{c}\in\mathbf{C}_k}\{\sum_{i=0}^{k} n_i c_i \bar{Y}_i/(s^2\sum_{i=0}^{k} n_i c_i^2)^{1/2}\},$$

where $\mathbf{C}_k = \{\mathbf{c} = (c_0, c_1, ..., c_k)' \in R^{k+1} : \sum_{i=0}^{k} n_i c_i = 0, c_0 \leq c_1 \leq \cdots \leq c_k\}$.

Let $t_{\alpha,k,\nu}^{so}$ be the critical value of T_k^{so} and let

$$T_j^{so} = \max_{\mathbf{c}\in\mathbf{C}_j}\{\sum_{i=0}^{j} n_i c_i \bar{Y}_i/(s^2\sum_{i=0}^{j} n_i c_i^2)^{1/2}\},$$

where $\mathbf{C}_j = \{\mathbf{c} \in R^{k+1} : \sum_{i=0}^{k} n_i c_i = 0, c_0 \leq c_1 \leq \cdots \leq c_j, c_{j+1} = \cdots = c_k = 0\}$. When $T_j^{so} > t_{\alpha,k,\nu}^{so}$, one rejects H_{0j} in favor of $H_{1j} : \mu_0 \leq \mu_1 \leq \cdots \leq \mu_j$ with at least one strict inequality. Note that T_j^{so} fails to satisfy the property that $T_1^{so} \leq T_2^{so} \leq \cdots \leq T_k^{so}$. In order to make a simultaneous inference, one applies the Bonferroni inequality so that H_{0j} is rejected if $T_j^{so} > t_{\alpha/k,j,\nu}^{so}$. The corresponding simultaneous confidence lower bound for $\mu_j - \mu_0$ is

$$L^{so}(\mu_j - \mu_0) = \max_{\mathbf{c}\in\mathbf{C}_j, \sum_{i=0}^{j} n_i c_i \mu_i \leq \mu_j - \mu_0}\{\sum_{i=0}^{j} n_i c_i \bar{Y}_i - t_{\alpha/k,j,\nu}^{so} s(\sum_{i=0}^{j} n_i c_i^2)^{1/2}\}. \quad (6.3)$$

6.3 Orthant Test

The hypothesis $H_{0j} : \mu_0 = \mu_1 = \cdots = \mu_j$ satisfies $H_{01} \supset H_{02} \supset \cdots \supset H_{0k}$ where $H_0 = H_{0k}$. Consider the rejection region $R_j = \{y : T_j \geq t\}$ for the test of H_{0j}. If the test statistic T_j is monotone nondecreasing, then the rejection region for the union-intersection test of H_0 is $R = \cup_{j=1}^{k} R_j$ which is $\{y : T_k \geq t\}$. The test statistic for testing H_0 is $T_k = \max_{1 \leq j \leq k} T_j$. The Dunnett's test $D_k = \max_{1 \leq j \leq k} D_j$ and the MLRT $T_k^{st} = \max_{1 \leq j \leq k} T_j^{st}$ are both union-intersection tests; however, T_k^{so} is not. In the following subsection, we will propose a new test statistic which is a union-intersection test based on T_j^{so}.

6.3.1 Orthant Test Statistic

Consider the union-intersection test based on the statistic T_j^{so} and the rejection region for testing H_0 against H_1 is $\{y : \max_{1 \leq j \leq k} T_j^{so} \geq c\}$. Therefore, we have

$$\begin{aligned}\max_{1 \leq j \leq k} T_j^{so} &= \max_{1 \leq j \leq k} \max_{\mathbf{c} \in \mathbf{C}_j} \{\sum_{i=0}^{k} n_i c_i \bar{Y}_i / (s^2 \sum_{i=0}^{j} n_i c_i^2)^{1/2}\} \\ &= \max_{\mathbf{c} \in \cup_{j=1}^{k} \mathbf{C}_j} \{\sum_{i=0}^{k} n_i c_i \bar{Y}_i / (s^2 \sum_{i=0}^{j} n_i c_i^2)^{1/2}\}.\end{aligned}$$

However, the set $\cup_{j=1}^{k} \mathbf{C}_j$ is not convex, hence it is difficult to compute its critical value. Let

$$\mathbf{O}_k = \{\mathbf{c} \in \mathbf{R}^{k+1} : \sum_{i=0}^{k} n_i c_i = 0, c_0 \leq \bar{c}_{01} \leq \cdots \leq \bar{c}_{0k}\}$$

where $\bar{c}_{0j} = \sum_{i=0}^{j} n_i c_i / \sum_{i=0}^{j} n_i$. The convex set $\mathbf{O}_k$ is an orthant. It is also known as upper-starshaped (Robertson, Wright and Dykstra 1988).

Lemma 6.3.1 *The set $\mathbf{O}_k$ is a convex hull of $\cup_{j=1}^{k} \mathbf{C}_j$.*

Proof. Let $\mathbf{x} = [x_0, ..., x_k]'$ and $\mathbf{y} = [y_0, ..., y_k]'$ be two vectors in $\mathbf{O}_k$ and $\mathbf{z} = [z_0, ..., z_k]' = \lambda \mathbf{x} + (1-\lambda)\mathbf{y}$ where $0 \leq \lambda \leq 1$. Let $\bar{x}_{0j} = \sum_{i=0}^{j} n_i x_i / \sum_{i=0}^{j} n_i$, $\bar{y}_{0j} = \sum_{i=0}^{j} n_i y_i / \sum_{i=0}^{j} n_i$ and $\bar{z}_{0j} = \sum_{i=0}^{j} n_i z_i / \sum_{i=0}^{j} n_i$. It is trivial that

$$\bar{z}_{0,i-1} = \lambda \bar{x}_{0,i-1} + (1-\lambda)\bar{y}_{0,i-1}$$

$$\begin{aligned} &\leq \lambda \bar{x}_{0i} + (1-\lambda)\bar{y}_{0i} \\ &= \bar{z}_{0i} \end{aligned}$$

and hence $\mathbf{O}_k$ is convex. It is obvious that $\mathbf{C}_j \subset \mathbf{O}_k, j = 1, ..., k$. Let $co(\cup_{j=1}^k \mathbf{C}_j)$ denote the convex hull of $\cup_{j=1}^k \mathbf{C}_j$. Therefore, we have $\mathbf{O}_k \supset co(\cup_{j=1}^k \mathbf{C}_j)$. On the other hand, the generators of the set $\mathbf{O}_k$ are $\{\mathbf{e}_i\}_{i=1,...,k}$ where $\mathbf{e}_j = [-n_{0,j-1}^{-1}, ..., -n_{0,j-1}^{-1}, n_j^{-1}, 0, ..., 0]'$ with the jth entry $e_{jj} = n_j^{-1}$ and $n_{0j} = \sum_{i=0}^j n_i$. Since $\mathbf{e}_j$ is in $\mathbf{C}_j$, we have $\mathbf{O}_k \subset co(\cup_{j=1}^k \mathbf{C}_j)$. □

Let

$$T_k^o = \max_{\mathbf{c} \in \mathbf{O}_k} \{ \sum_{i=0}^k n_i c_i \bar{Y}_i / (s^2 \sum_{i=0}^k n_i c_i^2)^{1/2} \}.$$

The T_k^o is a modified union-intersection test statistic obtained by fully utilizing the prior knowledge $\mu_0 \leq \mu_1 \leq \cdots \leq \mu_k$. The T_k^o is slightly greater than the union-intersection test statistic $\max_{1 \leq j \leq k} T_j^{so}$; however, the difference is small. For example, when $\nu = \infty$, the upper 5% critical points for T_k^o are 2.057, 2.331 and 2.549 for $k = 2, 3$ and 4, respectively. They are the upper critical points of the statistic $\max_{1 \leq j \leq k} T_j^{so}$ at the levels 4.9%, 4.6% and 4.1%, respectively The statistic T_k^o can be formatted as

$$\begin{aligned} T_k^o &= \{ \sum_{i=0}^k n_i \mu_i^{o2} / s^2 \}^{1/2} \\ &= \{ \sum_{i=1}^k \frac{n_i n_{0,i-1}}{n_{0i}} [\max(0, \bar{Y}_i - \bar{Y}_{0,i-1})]^2 / s^2 \}^{1/2} \end{aligned}$$

where $\mu^o = (\mu_1^o, ..., \mu_k^o)$ is the weighted least square projection of $(\bar{Y}_0, ..., \bar{Y}_k)$ onto $\mathbf{O}_k$ with the weights $n_0, n_1, ..., n_k$, and $\bar{Y}_{0j} = \sum_{i=0}^j n_i \bar{Y}_i / \sum_{i=0}^j n_i$. The derivation of the last equality is seen in Section 6.5. The statistic T_k^o is used to test H_0 against $H_1^o : \mu_0 \le \bar{\mu}_{01} \le \cdots \le \bar{\mu}_{0k}$ with at least one strict inequality, where $\bar{\mu}_{0j} = \sum_{i=0}^{j} n_i\mu_i / \sum_{i=0}^{j} n_i$. One rejects H_0 in favor of H_1^o if $T_k^o > t_{\alpha,k,\nu}^o$. Define

$$T_j^o = \max_{\mathbf{c}\in\mathbf{O}_j}\{\sum_{i=0}^{j} n_i c_i \bar{Y}_i / (s^2 \sum_{i=0}^{j} n_i c_i^2)^{1/2}\},$$

where

$$\mathbf{O}_j = \{\mathbf{c} \in \mathbf{R}^{k+1} : \sum_{i=0}^{k} n_i c_i = 0, c_0 \le \bar{c}_{01} \le \cdots \le \bar{c}_{0j}, c_{j+1} = \cdots = c_k = 0\}.$$

The statistic T_j^o has the property that $T_1^o \le T_2^o \le \cdots \le T_k^o$. It will be demonstrated in Section 6.3.3 that this new test statistic is more powerful than the aforementioned test statistics for testing H_{0j} against H_{1j}. The simultaneous confidence lower bound for pairwise mean difference $\mu_j - \mu_0$ corresponding to T_k^o will be discussed in Section 6.4.

6.3.2 The Null Distribution of T_k^o.

The critical value $t_{\alpha,k,\nu}^o$ for T_k^o is given by

$$P_\mu\{\max_{\mathbf{c}\in\mathbf{O}_k} \sum_{i=0}^{k} n_i c_i(\bar{Y}_i - \mu_i)/s(\sum_{i=0}^{k} n_i c_i^2)^{1/2} \le t_{\alpha,k,\nu}^o, \forall \mu \in \mathbf{R}^{k+1}\} = 1 - \alpha. \quad (6.4)$$

The left hand side can be rewritten as

$$\begin{aligned} & P_0\{\max_{\mathbf{c}\in\mathbf{O}_k} \sum_{i=0}^{k} n_i c_i \bar{Y}_i / s(\sum_{i=0}^{k} n_i c_i^2)^{1/2} \le t_{\alpha,k,\nu}^o\} \\ = \quad & P_0\{\sum_{i=0}^{k} n_i \mu_i^{o2}/s^2 \le (t_{\alpha,k,\nu}^o)^2\}. \end{aligned}$$

The p-value of $T_k^o = t^o$ is given by

$$P(T_k^o \ge t^o) = \sum_{l=2}^{k+1} \binom{k}{l-1} 2^{-k} P(F_{l-1,\nu} \ge \frac{t^o}{l-1}), \quad (6.5)$$

while the corresponding one for $T_k^{st} = t^{st}$ is given by

$$P(T_k^{st} \geq t^{st}) = \sum_{l=2}^{k+1} P_{st}(l, k+1; \mathbf{w}) P(F_{l-1,\nu} \geq \frac{t^{st}}{l-1}) \tag{6.6}$$

(Wright 1988) where $\mathbf{w} = [w_0, w_1, ..., w_k]'$ with $w_i = n_i/\sigma^2$ and $P_{st}(l, k+1; \mathbf{w})$ denotes the level probability that there are exactly l distinct values (levels) for the MLE satisfying the simple tree order (Robertson, Wright and Dykstra 1988). The $P_{st}(l, k+1; \mathbf{w})$'s depend on the sample sizes and the population variances through the weights w_i. Robertson, Wright and Dykstra (1988) discussed that $P_{st}(l, k+1; \mathbf{w})$'s converge to the binomial probabilities $\binom{k}{l-1}(\frac{1}{2})^k$ with k trials and the probability of success equals 1/2 when the weight at the control $w_0 \to \infty$.

On the other hand, the p-value of $T_k^{so} = t^{so}$ has the same format as (6.6) except that one replaces $P_{st}(l, k+1; \mathbf{w})$ by $P_S(l, k+1; \mathbf{w})$ which denotes the level probability for the simple order restriction (see Section 2.2). Robertson and Wright (1982) discussed that $P_S(l, k+1; \mathbf{w})$'s converge to the binomial probabilities with k trials and probability of success equal to 1/2 for a particular sequence $\mathbf{w}$. These particular limiting distributions of T_k^{st} and T_k^{so} correspond to that of T_k^o. Therefore, T_k^o will have the same distribution as the limiting distribution of T_k^{so} and T_k^{st}, where $T_k^{so} \leq T_k^o \leq T_k^{st}$. The critical value of $t^o_{\alpha,k,\nu}$ can be found in Table A.9 when $w_0 = \infty$ by Robertson, Wright and Dysktra (1988).

6.3.3 Power Comparisons

The power functions of simultaneous tests for null hypotheses $H_{0j} : \mu_0 = \mu_1 = \cdots = \mu_j$ against $H_{1j} : \mu_0 \leq \mu_1 \leq \cdots \leq \mu_j$ with at least one strict inequality are studied for D_j, T_j^{so} (with Bonferroni inequality applied), T_j^{st} and $T_j^o, j = 1, ..., k$. For simplicity, we consider the equal sample size case $n_i = n$ with $\sigma^2/n = 1, \alpha = 0.05$. In dose-response studies, the logistic function is one of the most popular dose-response curves. The logistic function considered here is $f(x) = E\{1 - [1 + (x/C)^5]^{-1}\}$ where x is the dose level and $f(x)$ is the corresponding dose-response mean with $f(5)$ fixed at 4 (Ruberg 1995). We study five cases with $C = 1.5, 2.0, 2.5, 3.0, 3.5$.

As Dunnett's test D_j, T_j^{st} and the new proposed test T_j^o have the property that $D_1 \le D_2 \le \cdots \le D_k$, $T_1^{st} \le T_2^{st} \le \cdots \le T_k^{st}$ and $T_1^o \le T_2^o \le \cdots \le T_k^o$, they can be used to detect the difference between μ_j and μ_0. However, as the statistic T_j^{so} does not satisfy $T_1^{so} \le T_2^{so} \le \cdots \le T_k^{so}$, we apply the Bonferroni inequality to obtain a conservative simultaneous test such that we reject H_{0j} in favor of H_{1j} for large values of T_j^{so}. With the prior assumption of monotonicity, if H_{0j} is rejected, the lower bound for $\mu_j - \mu_0$ will be positive. The power for testing H_{0j} against H_{1j} is actually the probability of obtaining the positive simultaneous lower bound for $\mu_j - \mu_0$. The simulation results for $k = 5$ are given in Table 6.1.

Table 6.1 indicates that if there are significant differences between the dose levels and the zero-dose control level, the orthant test statistic T_j^o is much more powerful in detecting this difference than the other three procedures. When $C = 2.5$, the orthant test T_j^o has the largest power among the four tests for detecting the differences between μ_j and μ_0 for $j = 3, 4$ and 5. Even though D_j and T_j^{so} have larger powers to detect the differences between μ_j and μ_0 for $j = 1$ and 2, they gain little over the orthant test statistic. The maximum gains of the statistic T_j^o over D_j, T_j^{so} and T_j^{st} are 11.7%, 4.4% and 10.9%; however, the maximum gains of the statistics D_j and T_j^{so} over T_j^o are only 3.1% and 1.6%, respectively. Similar results apply in the other four cases. In general, when the probabilities to detect the significant difference between μ_j and μ_0 by the four procedures are all above 50%, the gains of the orthant test statistic T_j^o over the other three tests D_j, T_j^{so} and T_j^{st} can reach 15.0%, 7.8% and 10.8%, respectively. When the difference between the dose-response mean and the control mean is detected, the new statistic T_j^o is the one to use.

6.4 Simultaneous Confidence Lower Bounds for Pairwise Mean Differences

6.4.1 The Optimization Problem

In order to assess the size of the difference between the response mean at level j and the zero-dose control mean, one needs to construct a corresponding simultaneous confidence lower bound. According to (6.4), a $100(1-\alpha)\%$ simultaneous confidence lower bound for $\mu_j - \mu_0$ can be inverted by the orthant test and is given by

$$l(\mu_j - \mu_0) = \bar{Y}_j - \bar{Y}_0 - t^o_{\alpha,k,\nu} s (n_j^{-1} + n_0^{-1})^{1/2}. \tag{6.7}$$

For more general contrasts, $\sum_{i=0}^k n_i c_i \mu_i$, $\mathbf{c} = (c_0, c_1, ..., c_k) \in \mathbf{O}_k$, the $100(1-\alpha)\%$ simultaneous confidence lower bound can be constructed as

$$l(\sum_{i=0}^k n_i c_i \mu_i) = \sum_{i=0}^k n_i c_i \bar{Y}_i - t^o_{\alpha,k,\nu} s (\sum_{i=0}^k n_i c_i^2)^{1/2}. \tag{6.8}$$

If one rejects H_0, there exists at least one contrast $\sum_{i=0}^k n_i c_i \mu_i$ that has a positive lower bound. Specifically, if $T_j^o > t^o_{\alpha,k,\nu}$, one rejects H_{0j} in favor of H_{1j} and there exists a contrast $\sum_{i=0}^{j} n_i c_i \mu_i \leq \mu_j - \mu_0, \mathbf{c} \in \mathbf{O}_j$ such that $l(\sum_{i=0}^{j} n_i c_i \mu_i) > 0$. It suffices to consider the confidence lower bound for $\mu_k - \mu_0$ under the assumption $\mu_0 \leq \mu_1 \leq \cdots \leq \mu_k$. The result for $\mu_j - \mu_0$ follows similarly. The lower bound for $\mu_k - \mu_0$ can be improved to

$$L^o(\mu_k - \mu_0) = \max_{\mathbf{c} \in \mathbf{O}_k, \sum_{i=0}^{k} n_i c_i \mu_i \leq \mu_k - \mu_0} l(\sum_{i=0}^{k} n_i c_i \mu_i). \tag{6.9}$$

The positive lower bound for $\mu_k - \mu_0$ indicates the difference between the dose level k and the control. We have the following lemma and its proof is straightforward.

Lemma 6.4.1 *$T_k^o > t_{\alpha,k,\nu}^o$ if and only if $L^o(\mu_k - \mu_0) > 0$.*

We shall restrict our attention to the case $l(\sum_{i=0}^{k} n_i c_i \mu_i) > 0$ for some $\mathbf{c} \in \mathbf{O}_k$, i.e., when $T_k^o > t_{\alpha,k,\nu}^o$. The value of the lower bound $L^o(\mu_k - \mu_0)$ indicates the size of the difference between μ_k and μ_0. We can assess a minimum dosage level which has the desired difference from the zero-dose control mean.

Let $n_{pq} = \sum_{j=p}^{q} n_j$ if $p \leq q$ and $n_{pq} = 0$ if $p > q$. The evaluation of the lower bound $L^o(\mu_k - \mu_0)$ in (6.9) is an optimization problem. In order to solve this rather complicated concave programming problem and seek an efficient algorithm to compute this improved lower bound, we consider the transformation $z_i = \bar{Y}_i - \bar{Y}_{0,i-1}, \delta_i = \mu_i - \bar{\mu}_{0,i-1}, a_i = n_{0,i-1} n_i (c_i - \bar{c}_{0,i-1})/n_{0i}$. Then $z_1, ..., z_k$ are normally distributed with means δ_i and covariance matrix $\sigma^2 \Sigma = \sigma^2 [\sigma_{ij}]$ where $\sigma_{ii} = n_{0,i-1}^{-1} + n_i^{-1}$, $\sigma_{ij} = 0$ if $i \neq j$ and $\sum_{i=0}^{k} n_i c_i \mu_i = \sum_{i=1}^{k} a_i \delta_i$. Let $A = [\alpha_{ij}]$ with $\alpha_{ij} = n_{0,i-1}/n_{0,j-1}$ if $i \leq j$, and 0 otherwise. The constraint $\mathbf{c} \in \mathbf{O}_k$, i.e, $\bar{c}_{0,j-1} \leq c_j (j = 1, ..., k)$, is equivalent to $a_j \geq 0$.

In addition, with the prior knowledge $\mu_0 \leq \mu_1 \leq \cdots \leq \mu_k$, the constraint $\sum_{i=0}^{k} n_i c_i \mu_i \leq \mu_k - \mu_0$ is equivalent to $\sum_{j=i}^{k} n_j c_j \leq 1, i = 1, ..., k$, which can also be shown equivalent to $\sum_{j=i}^{k} a_j n_{0,i-1}/n_{0,j-1} \leq 1$. Let $\mathbf{a} = [a_1, ..., a_k]'$. The problem (6.9) becomes

$$\max_{\mathbf{a} \geq 0, \mathbf{A}\mathbf{a} \leq \mathbf{1}} l(\sum_{i=1}^{k} a_i \delta_i) = \max_{\mathbf{a} \geq 0, \mathbf{A}\mathbf{a} \leq \mathbf{1}} \sum_{i=1}^{k} \{a_i z_i - t^o_{\alpha,k,\nu} s (\sum_{i=1}^{k} a_i^2 \sigma_{ii})^{1/2}\} \tag{6.10}$$

where $\mathbf{1} = [1, 1, ..., 1]'_{k \times 1}$. Let $\mathbf{a}^o$ be the optimal solution to the problem (6.10). Note that $\mathbf{a}^o$ has the following property:

Lemma 6.4.2 *Suppose that the maximum of* $l(\sum_{i=1}^{k} a_i \delta_i)$ *subject to* $\mathbf{a} \geq \mathbf{0}, \mathbf{A}\mathbf{a} \leq \mathbf{1}$ *is attained at* $\mathbf{a}^o$, *then* $z_i \leq 0$ *implies that* $a_i^o = 0$.

Proof. Suppose there exists $z_j < 0$ and $a_j^o > 0$. Let $d_j = 0$ and $d_i = a_i^o$ if $i \neq j$. Then we have

$$\sum_{i=1}^{k} a_i^o z_i < \sum_{i=1}^{k} d_i z_i$$

and

$$\sum_{i=1}^{k} a^{o2}_i \sigma_{ii} > \sum_{i=1}^{k} d_i^2 \sigma_{ii}.$$

Therefore,

$$\sum_{i=1}^{k} a_i^o z_i - t^o_{\alpha,k,\nu} s (\sum_{i=1}^{k} a^{o2}_i \sigma_{ii})^{1/2} < \sum_{i=1}^{k} d_i z_i - t^o_{\alpha,k,\nu} s (\sum_{i=1}^{k} d_i^2 \sigma_{ii})^{1/2}$$

which contradicts the assumption. The proof is complete. □

Let $\mathbf{w} = [w_0, w_1, \cdots, w_k]'$ be the vector of weights where $w_i = n_i / n_{0k} (i = 0, 1, ..., k)$. If $\mathbf{x} = [x_0, x_1, ..., x_k]'$ and $\mathbf{y} = [y_0, y_1, ..., y_k]'$ are in $\mathbf{R^{k+1}}$, then the inner product and the norm are defined respectively by

$$< \mathbf{x}, \mathbf{y} >_w = \sum_{i=0}^{k} w_i x_i y_i$$

$$||\mathbf{x}||_w^2 = \sum_{i=0}^{k} w_i x_i^2.$$

Let $\mathbf{e}_i = [-n_{0,i-1}^{-1}, \cdots, -n_{0,i-1}^{-1}, n_i{}^{-1}, 0, \cdots, 0]'$ with $e_{ii} = n_i{}^{-1}$. Let $P(\bar{\mathbf{Y}}|\mathbf{O}_k)$ be the vector $\mathbf{v} \in \mathbf{O}_k$ minimizing $||\bar{\mathbf{Y}} - \mathbf{v}||_{\mathbf{w}}$. It can be shown that $P(\bar{\mathbf{Y}}|\mathbf{O}_k)$ can be expressed by $\sum_{i=1}^{k} < \bar{\mathbf{Y}}, \mathbf{e}_i >_{\mathbf{w}}^{+} \mathbf{e}_i/||\mathbf{e}_i||_{\mathbf{w}}^2$ where $c^+ = \max\{c, 0\}$. Lemma 6.4.2 guarantees that $\bar{\mathbf{Y}}$ and $P(\bar{\mathbf{Y}}|\mathbf{O}_k)$ will lead to the same optimal lower bound for $\mu_k - \mu_0$.

Let $R = \{i : a_i > 0, [A\mathbf{a}]_i = 1\}, S = \{i : a_i = 0, [A\mathbf{a}]_i < 1\}$ and $T = \{i : a_i > 0, [A\mathbf{a}]_i < 1\}$ where the notation $[A\mathbf{a}]_i$ denotes the ith component of the vector $A\mathbf{a}$. Since $[A\mathbf{a}]_i = a_i + (n_{0,i-1}/n_{0i})[A\mathbf{a}]_{i+1}$, $a_i = 0$ implies that $[A\mathbf{a}]_i < 1$. Therefore, R, S and T form a partition of $\{1, ..., k\}$. Let $\mathbf{a} = [a_1, ..., a_k]', \mathbf{Z} = [z_1, ..., z_k]'$ and $\mathbf{1}$ be partitioned as $\mathbf{a} = [\mathbf{a}'_R, \mathbf{a}'_S, \mathbf{a}'_T]', \mathbf{Z} = [\mathbf{Z}'_R, \mathbf{Z}'_S, \mathbf{Z}'_T]'$ and $\mathbf{1} = [\mathbf{1}'_R, \mathbf{1}'_S, \mathbf{1}'_T]'$. The same partition applies to A and Σ. A necessary and sufficient condition for the optimal solution to (6.10) is given by the following theorem.

Theorem 6.4.1 *The maximum of* $l(\sum_{i=1}^{k} a_i \delta_i)$ *subject to* $\mathbf{a} \geq 0, A\mathbf{a} \leq \mathbf{1}$ *is attained at* $\mathbf{a}^o$ *if and only if* $\mathbf{a}^o$ *satisfies*

$$\mathbf{a}_T^o = \Delta_{T.R}^{-1}\{A'_{RT}A'^{-1}_{RR}\Sigma_{RR}A_{RR}^{-1}\mathbf{1}_R + b(\mathbf{Z}_T - A'_{RT}A'^{-1}_{RR}\mathbf{Z}_R)\}; \tag{6.11}$$

$$\mathbf{a}_R^o = A_{RR}^{-1}(\mathbf{1}_R - A_{RT}\mathbf{a}_T^o); \tag{6.12}$$

$$A'^{-1}_{RR}(\mathbf{Z}_R - b^{-1}\Sigma_{RR}\mathbf{a}_R^o) \geq 0; \tag{6.13}$$

$$A'_{RS}A'^{-1}_{RR}(\mathbf{Z}_R - b^{-1}\Sigma_{RR}\mathbf{a}^o_R) \geq \mathbf{Z}_S \tag{6.14}$$

where $b = (\mathbf{a}^{o\prime}\Sigma\mathbf{a}^o)^{1/2}/(t^o_{\alpha,k,\nu}s)$ *and* $\Delta_{T.R} = \Sigma_{TT} + A'_{RT}A'^{-1}_{RR}\Sigma_{RR}A^{-1}_{RR}A_{RT}$. *When* $T = \emptyset$, (6.11) *does not apply and* (6.12) *becomes* $\mathbf{a}_R = A^{-1}_{RR}\mathbf{1}_R$.

Proof. Consider the problem

$$\max_{\mathbf{a}\geq\mathbf{0},\mathbf{Aa}\leq\mathbf{1}}\{\mathbf{a}'\mathbf{Z} - t^o_{\alpha,k,\nu}s(\mathbf{a}'\Sigma\mathbf{a})^{1/2}\}. \tag{6.15}$$

Let $\phi(\mathbf{a},\mathbf{u}) = l(\sum_{i=1}^k a_i\delta_i) + \mathbf{u}'(\mathbf{1} - A\mathbf{a})$ and let $\frac{\partial\phi}{\partial\mathbf{a}^o}$ denote the partial derivatives evaluated at the point $\mathbf{a}^o$ and $\mathbf{u}^o$. It can be shown that $l(\sum_{i=1}^k a_i\delta_i)$ is concave. By the Kuhn-Tucker equivalence theorem (Kuhn and Tucker 1951), $\mathbf{a}^o$ is the solution to the problem in (6.15) if and only if

(i) $\frac{\partial\phi}{\partial\mathbf{a}^o} \leq 0, (\frac{\partial\phi}{\partial\mathbf{a}^o})'\mathbf{a}^o = 0$ and $\mathbf{a}^o \geq 0$,

(ii) $\mathbf{1} - A\mathbf{a}^o \geq 0, (\mathbf{1} - A\mathbf{a}^o)'\mathbf{u}^o = 0$ and $\mathbf{u}^o \geq 0$.

Let $\mathbf{a}^o$ be the optimal solution and let $\mathbf{u}$ have the same partition $\mathbf{u} = [\mathbf{u}'_R, \mathbf{u}'_S, \mathbf{u}'_T]'$. Therefore, $\phi(\mathbf{a},\mathbf{u})$ can be written as

$$\begin{aligned}\phi(\mathbf{a},\mathbf{u}) \;=\;& \mathbf{a}'_R\mathbf{Z}_R + \mathbf{a}'_S\mathbf{Z}_S + \mathbf{a}'_T\mathbf{Z}_T \\ & -t^o_{\alpha,k,\nu}s(\mathbf{a}'_R\Sigma_{RR}\mathbf{a}_R + \mathbf{a}'_S\Sigma_{SS}\mathbf{a}_S + \mathbf{a}'_T\Sigma_{TT}\mathbf{a}_T)^{1/2} \\ & +\mathbf{u}'_R\{\mathbf{1}_R - (A_{RR}\mathbf{a}_R + A_{RS}\mathbf{a}_S + A_{RT}\mathbf{a}_T)\} \\ & +\mathbf{u}'_S\{\mathbf{1}_S - (A_{SR}\mathbf{a}_R + A_{SS}\mathbf{a}_S + A_{ST}\mathbf{a}_T\} \\ & +\mathbf{u}'_T\{\mathbf{1}_T - (A_{TR}\mathbf{a}_R + A_{TS}\mathbf{a}_S + A_{TT}\mathbf{a}_T)\}.\end{aligned}$$

Condition (ii) implies that $\mathbf{u}^o_R \geq 0, \mathbf{u}^o_S = 0$, and $\mathbf{u}^o_T = 0$. Condition (i) becomes

$$\frac{\partial\phi}{\partial\mathbf{a}^o_R} = \mathbf{Z}_R - b^{-1}\Sigma_{RR}\mathbf{a}^o_R - A'_{RR}\mathbf{u}^o_R = 0,$$

$$\frac{\partial \phi}{\partial \mathbf{a}_S^o} = \mathbf{Z}_S - b^{-1}\Sigma_{SS}\mathbf{a}_S^o - A'_{RS}\mathbf{u}_R^o \leq 0,$$

and

$$\frac{\partial \phi}{\partial \mathbf{a}_T^o} = \mathbf{Z}_T - b^{-1}\Sigma_{TT}\mathbf{a}_T^o - A'_{RT}\mathbf{u}_R^o = 0,$$

where $b = (\mathbf{a}^{o'}\Sigma\mathbf{a}^o)^{1/2}/(t^o_{\alpha,k,\nu}s)$.

It follows that

$$\mathbf{u}_R^o = A'^{-1}_{RR}(\mathbf{Z}_R - b^{-1}\Sigma_{RR}\mathbf{a}_R^o) \geq 0,$$

$$A'_{RS}\mathbf{u}_R^o \geq \mathbf{Z}_S,$$

$$\Sigma_{TT}\mathbf{a}_T^o - A'_{RT}A'^{-1}_{RR}\Sigma_{RR}\mathbf{a}_R^o = b(\mathbf{Z}_T - A'_{RT}A'^{-1}_{RR}\mathbf{Z}_R).$$

The condition $[A\mathbf{a}]_i = 1$ for any $i \in R$ is equivalent to

$$A_{RR}\mathbf{a^o}_R + A_{RT}\mathbf{a^o}_T = \mathbf{1}_R.$$

The last two identities lead to the expressions (6.11) and (6.12), while the last two inequalities are equivalent to expressions (6.13) and (6.14). □

It can be shown that if $l(\sum_{i=1}^k a_i^o\delta_i) > 0$, then R is not an empty set.

6.4.2 Simplified Formulas

The computation for $\mathbf{a}^o$ and the conditions in Theorem 6.4.1 can be simplified. Let $R = \{r_1, ..., r_m\}$ with the convention $r_0 = 0$ and $r_{m+1} = k+1$ and let t, p and q be three consecutive indices in $R \cup \{0, k+1\}$. Let $\tau_{p,q} = n_{p,q-1}/(n_{0,p-1}n_{0,q-1})$ with the convention $\tau_{p,k+1} = n^{-1}_{0,p-1}$ and $\tau_{0,p} \equiv 0$. Let $\tau_{p,q,S} = \sum_{j=p,j\notin S}^{q-1}(n^{-1}_{0,j-1} - n^{-1}_{0j})$ with the convention that $\tau^{-1}_{0,p,S} \equiv 0$. Note that if there does not exist any index i, $p < i < q$ such that $i \in S$, then $\tau_{p,q,S} = \tau_{p,q}$ for $p \neq 0, q \neq k+1$. Let $\eta_{p,q,S} = \sum_{j=p,j\notin S}^{q-1} (n_j/n_{0j}) z_j$ with the convention that $\eta_{0,p,S} \equiv 0$. The expressions (6.11) and (6.12) become

$$a_i^o = (n_i/n_{0i})\{\tau_{p,q}\tau_{p,q,S}^{-1} + b(n_{0,i-1}z_i - \eta_{p,q,S}\tau_{p,q,S}^{-1})\}, \text{ for } p \leq i < q \text{ and } i \in T \cup R. \tag{6.16}$$

Conditions (6.13) and (6.14) become

$$n_{0,p-1}^{-1}\{\tau_{p,q,S}^{-1}\eta_{p,q,S} - \tau_{t,p,S}^{-1}\eta_{t,p,S} - b^{-1}(\tau_{p,q}\tau_{p,q,S}^{-1} - \tau_{t,p}\tau_{t,p,S}^{-1})\} \geq 0, \text{ for } p \in R; \tag{6.17}$$

and

$$n_{0,i-1}^{-1}\tau_{p,q,S}^{-1}(\eta_{p,q,S} - b^{-1}\tau_{p,q}) \geq z_i, \text{ for } i \in S, \tag{6.18}$$

respectively. The coefficient b can be obtained by

$$b^2 = \sum_{p\in R} \tau_{p,q}^2 \tau_{p,q,S}^{-1} / \{(t_{\alpha,k,\nu}^o s)^2 - \sum_{i=1,i\notin S}^{k} (n_i n_{0,i-1}/n_{0i}) z_i^2 + \sum_{p\in R} \eta_{p,q,S}^2 \tau_{p,q,S}^{-1}\}. \tag{6.19}$$

The constraint $[A\mathbf{a}] \leq 1$ becomes

$$\tau_{p,q}\tau_{p,q,S}^{-1}\tau_{i,q,S} + \tau_{q,k+1} + b(\eta_{i,q,S} - \tau_{p,q,S}^{-1}\eta_{p,q,S}\tau_{i,q,S}) \leq n_{0,i-1}^{-1}, \text{ for } p \leq i < q. \tag{6.20}$$

The simplified formulas (6.16) to (6.20) determine whether the partition R, S and T is optimal. The number of feasible partitions for R, S and T is $3^k - 2^k$, a large number even for a moderate k. It is important to have an efficient algorithm to compute the optimal confidence lower bound for $\mu_k - \mu_0$. The following algorithm provides optimal partitions R, S and T for different confidence level $1-\alpha$, starting from $1-p$, where p is the p-value of the test statistic T_k^o.

6.4.3 Computation Algorithm

Without loss of generality, it suffices to consider the optimal lower bound for $\mu_k - \mu_0$. For simplicity, we use t_α to denote the critical value $t^o_{\alpha,k,\nu}$ in the remainder of this chapter. Let $t_0 = \{\sum_{i=1}^k \frac{n_i n_{0,i-1}}{n_{0i}}[\max(0, \bar{Y}_i - \bar{Y}_{0,i-1})]^2/s^2\}^{1/2}$. If $t_\alpha \geq t_0$, we have $L^o(\mu_k - \mu_0) = 0$. We assume that $t_0 > t_\alpha$.

(0) Let $M = \max_{1 \leq i \leq k} \sum_{j=i}^k n_i P(\bar{\mathbf{Y}}|\mathbf{O}_k)_j$. The initial $c_i^{(0)} = n_i P(\bar{Y}|\mathbf{O}_k)_i/M, i = 0, 1, ..., k$. Let $R^1 = \{i : [A\mathbf{a}^{(0)}]_i = \sum_{j=i}^k n_j c_j^{(0)} = 1\}, S^1 = \{i : a_i^{(0)} = c_i^{(0)} - \bar{c}_{0,i-1}^{(0)} = 0\}$ and $T^1 = \{1, ..., k\} - (R^1 \cup S^1)$. Set $r = 1$.

(1) Let p and q be consecutive indices in R^r. Compute

$$a_i^{(r)} = (n_i/n_{0i})\{\tau_{p,q}\tau_{p,q,S^r}^{-1} + b(n_{0,i-1}z_i - \eta_{p,q,S^r}\tau_{p,q,S^r}^{-1})\},\ \text{for } p \leq i < q \text{ and } i \in T^r \cup R^r,$$

$$b_r = \min\{b > 0 : [A\mathbf{a}^{(r)}]_i = 1 \text{ or } a_i^{(r)} = 0, i \in T^r\},$$

and

$$t_r = \{\sum_{p \in R^r} \tau_{p,q}^2 \tau_{p,q,S^r}^{-1}/b_r^2 + \sum_{j=1, j \notin S^r}^k (n_j n_{0,j-1}/n_{0j}) z_j^2 - \sum_{p \in R^r} \eta_{p,q,S^r}^2 \tau_{p,q,S^r}^{-1}\}^{1/2}/s.$$

If $t_\alpha > t_r$, stop. Otherwise, go to the next step.

(2) If there exists an index $h \in T^r$ such that $b = b_r$ and $[A\mathbf{a}^{(r)}]_h = 1$, define $R^{r+1} = R^r \cup \{h\}, S^{r+1} = S^r$ and $T^{r+1} = T^r - \{h\}$. On the other hand, if there exists an index $h \in T^r$ such that $b = b_r$, and $a_h^{(r)} = 0$, define $R^{r+1} = R^r, S^{r+1} = S^r \cup \{h\}$ and $T^{r+1} = T^r - \{h\}$. Set $r = r + 1$, go to step (1).

6.4.4 Application of the Algorithm

Let $\bar{Y}_0 = 2, \bar{Y}_1 = 4, \bar{Y}_2 = 0, \bar{Y}_3 = 10, \bar{Y}_4 = 14, \bar{Y}_5 = 12, n_0 = n_1 = \cdots = n_5 = 6$ and $s^2 = 35.4$. The computation of $L^o(\mu_5 - \mu_0)$ is illustrated as follows:

(0) Since $\bar{\mathbf{Y}} = (2, 4, 0, 10, 14, 12) \notin \mathbf{O}_k$, the projection $P(\bar{Y}|\mathbf{O}_k)$ is (-6, -4, -5, 3, 7, 5). Compute $t_0 = \{\sum_{i=1}^{k} \frac{n_i n_{0,i-1}}{n_{0i}} [\max(0, \bar{Y}_i - \bar{Y}_{0,i-1})]^2/s^2\}^{1/2} = 5.21$. The p-value of the test statistic T_k^o is 0.0002. We have $M = \max_i \sum_{j=i}^{k} n_j P(\bar{Y}|\mathbf{O})_j = 90$. The initial $\mathbf{c}^{(0)} = (-\frac{6}{15}, -\frac{4}{15}, -\frac{5}{15}, \frac{3}{15}, \frac{7}{15}, \frac{5}{15})$.

(1) Set $r = 1$ and $R^1 = \{3\}, S^1 = \{2\}$ and $T^1 = \{1, 4, 5\}$. Compute

$$\mathbf{a}^{(1)} = (6b, 0, \frac{1}{2} - 9b, \frac{2}{5} + 12b, \frac{1}{3}).$$

We have $b_1 = \min\{\frac{1}{18}, \frac{1}{9}, \frac{1}{36}\} = \frac{1}{36}$ and $t_1 = 2.40$. The R^1, S^1 and T^1 form the optimal partition for confidence level between 99.98% and 87.8%.

(2) Since $b_1 = 1/36$ occurs at the index $h = 4$ such that $[A\mathbf{a}^{(1)}]_4 = 1$, define $R^2 = \{3, 4\}, S^2 = \{2\}$ and $T^2 = \{1, 5\}$. Compute

$$\mathbf{a}^{(2)} = (6b, 0, \frac{1}{4}, \frac{3}{5} + \frac{24}{5}b, \frac{1}{2} - 6b).$$

We have $b_2 = \min\{\frac{1}{12}, \frac{1}{9}\} = \frac{1}{12}$ and $t_2 = 1.11$. The partition is optimal for confidence level between 87.8% and 38.7%.

(3) Since $b_2 = 1/12$ occurs at the index $h = 5$ such that $\mathbf{a}_5^{(2)} = 0$, define $R^3 = \{3, 4\}, S^3 = \{2, 5\}$ and $T^3 = \{1\}$. Compute

$$\mathbf{a}^{(3)} = (6b, 0, \frac{1}{4}, 1, 0).$$

We have $b_3 = \frac{1}{9}$ and $t_3 = 0.92$. The partition is optimal for confidence level between 38.7% and 30.0%.

(4) Since $b_3 = 1/9$ occurs at the index $h = 1$ such that $[A\mathbf{a}^{(3)}]_1 = 1$, define $R^4 = \{1, 3, 4\}, S^4 = \{2, 5\}$ and $T^4 = \emptyset$. The partition is optimal for confidence level less than 30.0%.

When $\alpha = 0.05$, the critical value with $k = 5$ and $\nu = 30$ is $t_{.05} = 2.88$. The 95% simultaneous confidence lower bound $L^o(\mu_5 - \mu_0) = 4.91$ can be obtained at Step (1) with $n\mathbf{c}^o = (-0.077, -0.034, -0.056, 0.012, 0.099, 0.056)'$.

Similarly, we have $L^o(\mu_4 - \mu_0) = 4.30$ and $L^o(\mu_3 - \mu_0) = 0.09$. Comparing to the Dunnett's procedure, with the critical value $d_{.05,5,30} = 2.33$ we note that $L^d(\mu_5 - \mu_0) = 4.00, L^d(\mu_4 - \mu_0) = 4.00$ and $L^d(\mu_3 - \mu_0) = 0$. This example demonstrates that lower bounds obtained by the new procedure are sharper than those of Dunnett's.

6.5 Technical Results

6.5.1 Simplification of the Optimal Solution

The following lemma will be used to simplify the computation procedure and its proof is straightforward.

Lemma 6.5.1 *The inverse matrix of* $I + \mathbf{w}\mathbf{w}'$ *is* $I - \lambda\mathbf{w}\mathbf{w}'$ *where* $\mathbf{w}$ *is a* $k \times 1$ *vector and the scalar* $\lambda = (1 + \mathbf{w}'\mathbf{w})^{-1}$.

Proof. We will show when $\lambda = 1/(1 + \mathbf{w}'\mathbf{w})$, $(I + \mathbf{w}'\mathbf{w})(I - \lambda\mathbf{w}'\mathbf{w}) = I$

$$(I + \mathbf{w}'\mathbf{w})(I - \lambda\mathbf{w}'\mathbf{w}) = I - \lambda\mathbf{w}'\mathbf{w} + \mathbf{w}'\mathbf{w} - \lambda\mathbf{w}'\mathbf{w}\mathbf{w}'\mathbf{w}$$

$$= I + (-\lambda + 1 - \lambda \mathbf{w}\mathbf{w}')\mathbf{w}'\mathbf{w}$$
$$= I.$$

It completes the proof. □

Let $R = \{r_1, ..., r_m\}$ with the convention $r_0 = 0$ and $r_{m+1} = k+1$ and let t, p and q be three consecutive indices in $R \cup \{0, k+1\}$. Then we have that

$$[A_{RR}^{-1}]_{r_i r_j} = \begin{cases} 1 & \text{if } r_i = r_j = p; \\ -n_{0,p-1}/n_{0,q-1} & \text{if } r_i = p, r_j = q; \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, $[A_{RR}^{-1}\mathbf{1}_R]_p = n_{0,p-1}\tau_{p,q}$. We also have

$$A_{RR}^{-1}A_{RT} = \begin{bmatrix} 0 & \mathbf{v}'_{r_1} & 0 & \cdots & 0 & 0 \\ 0 & 0 & \mathbf{v}'_{r_2} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \mathbf{v}'_{r_m} \end{bmatrix},$$

where $\mathbf{v}_p$ is a column vector with entries $\mathbf{v}_{pi} = n_{0,p-1}/n_{0,i-1}, p < i < q, i \in T$. Therefore,

$$\begin{aligned} [A'_{RT}A'^{-1}_{RR}\Sigma_{RR}A_{RR}^{-1}A_{RT}]_{ij} &= [\sigma_{pp}\mathbf{v}_p\mathbf{v}_p']_{ij} \\ &= \tau_{p,p+1}^{-1}n_{0,i-1}^{-1}n_{0,j-1}^{-1} \ \text{ for } \ p < i, j < q, i, j \in T \end{aligned}$$

with the convention $\sigma_{00} \equiv 0$. Then

$$\begin{aligned} \Delta_{T.R} &= \Sigma_{TT} + A'_{RT}A'^{-1}_{RR}\Sigma_{RR}A_{RR}^{-1}A_{RT} \\ &= \begin{bmatrix} \Sigma_{r_0} + \sigma_{r_0r_0}\mathbf{v}_{r_0}\mathbf{v}'_{r_0} & 0 & \cdots & 0 \\ 0 & \Sigma_{r_1} + \sigma_{r_1r_1}\mathbf{v}_{r_1}\mathbf{v}'_{r_1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Sigma_{r_m} + \sigma_{r_mr_m}\mathbf{v}_{r_m}\mathbf{v}'_{r_m} \end{bmatrix} \end{aligned}$$

where $\Sigma_p = [\sigma_{ij}]$ with $p < i, j < q, i, j \in T$.

We find $(\Sigma_p + \sigma_{pp}\mathbf{v}_p\mathbf{v'}_p)^{-1}$ by Lemma 6.5.1 as follows. Rewriting $\Sigma_p + \sigma_{pp}\mathbf{v}_p\mathbf{v'}_p$, we have

$$\Sigma_p + \sigma_{pp}\mathbf{v}_p\mathbf{v'}_p = \Sigma_p^{1/2}(I + \mathbf{w}_p\mathbf{w'}_p)\Sigma_p^{1/2}$$

where $\mathbf{w}_p = \sigma_{pp}^{1/2}\Sigma_p^{-1/2}\mathbf{v}_p$. Then $\mathbf{w'}_p\mathbf{w}_p = \sigma_{pp}\mathbf{v'}_p\Sigma_p^{-1}\mathbf{v}_p = \tau_{p,p+1}^{-1}\tau_{p+1,q,S}$ and $(1 + \mathbf{w'}_p\mathbf{w}_p)^{-1} = \tau_{p,p+1}\tau_{p,q,S}^{-1}$. Hence, we have that

$$\begin{aligned}
[\Delta_{T.R}]_{ij}^{-1} &= [\Sigma_p + \sigma_{pp}\mathbf{v}_p\mathbf{v'}_p]_{ij}^{-1} \\
&= [\Sigma_p^{-1} - (1 + \mathbf{w'}_p\mathbf{w}_p)^{-1}\Sigma_p^{-1}(\sigma_{pp}\mathbf{v}_p\mathbf{v'}_p)\Sigma_p^{-1}]_{ij} \\
&= \frac{n_i n_{0,i-1}}{n_{0i}}\delta_{ij} - \{\tau_{p,p+1}\tau_{p,q,S}^{-1}(\frac{n_{0i}}{n_{0,i-1}n_i})^{-1}(\tau_{p,p+1}^{-1}n_{0,i-1}^{-1}n_{0,j-1}^{-1})(\frac{n_{0j}}{n_{0,j-1}n_j})^{-1}\} \\
&= \frac{n_i n_{0,i-1}}{n_{0i}}\delta_{ij} - \tau_{p,q,S}^{-1}\frac{n_i n_j}{n_{0i}n_{0j}}
\end{aligned}$$

where $\delta_{ii} = 1$ and $\delta_{ij} = 0$ for $i \neq j$. Also we have

$$[A'_{RT}A'^{-1}_{RR}\Sigma_{RR}A_{RR}^{-1}\mathbf{1}_R]_j = \begin{bmatrix} \sigma_{r_0r_0}\mathbf{v}_{r_0}n_{0,r_0-1}\tau_{r_0,r_1} \\ \sigma_{r_1r_1}\mathbf{v}_{r_1}n_{0,r_1-1}\tau_{r_1,r_2} \\ \vdots \\ \sigma_{r_mr_m}\mathbf{v}_{r_m}n_{0r_m}\tau_{r_m,r_{m+1}} \end{bmatrix}_j = \tau_{p,q}\tau_{p,p+1}^{-1}n_{0,j-1}^{-1}$$

for $p < j < q, j \in T$. Hence,

$$\begin{aligned}
&[\Delta_{T.R}^{-1}A'_{RT}A'^{-1}_{RR}\Sigma_{RR}A_{RR}^{-1}\mathbf{1}_R]_i \\
=\ & (n_i n_{0,i-1}/n_{0i})\tau_{p,q}\tau_{p,p+1}^{-1}n_{0,i-1}^{-1} - \tau_{p,q,S}^{-1}(n_i/n_{0i})\tau_{p,q}\tau_{p,p+1}^{-1}\sum_{j=p+1, j\notin S}^{q-1}\{n_j/(n_{0j}n_{0,j-1})\} \\
=\ & (n_i/n_{0i})\tau_{p,q}\tau_{p,p+1}^{-1}(1 - \tau_{p+1,q,S}\tau_{p,q,S}^{-1}) \\
=\ & (n_i/n_{0i})\tau_{pq}\tau_{p,p+1}^{-1}\tau_{p,p+1}\tau_{p,q,S}^{-1} \\
=\ & (n_i/n_{0i})\tau_{p,q}\tau_{p,q,S}^{-1}.
\end{aligned}$$

We have

$$[\mathbf{Z}_T - A'_{RT}A'^{-1}_{RR}\mathbf{Z}_R]' = [\mathbf{Z}'_{r_0} - z_{r_0}\mathbf{v}'_{r_0}, \cdots, \mathbf{Z}'_{r_m} - z_{r_m}\mathbf{v}'_{r_m}]$$

where $\mathbf{Z}_p = [z_j]'$ for $p < j < q$ with the convention $z_0 \equiv 0$, then

$$\begin{aligned}
&[\Delta^{-1}_{T.R}(\mathbf{Z}_T - A'_{RT}A'^{-1}_{RR}\mathbf{Z}_R)]_i \\
= \;& (n_i n_{0,i-1}/n_{0i})z_i - \tau^{-1}_{p,q,S}(n_i/n_{0i}) \sum_{j=p+1,j\notin S}^{q-1} (n_j/n_{0j})z_j \\
& -\{(n_i n_{0,i-1}/n_{0i})(n_{0,p-1}/n_{0,i-1}) - \tau^{-1}_{p,q,S}(n_i/n_{0i})n_{0,p-1} \sum_{j=p+1,j\notin S}^{q-1} n_j/(n_{0j}n_{0,j-1})\}z_p \\
= \;& (n_i/n_{0i})\{n_{0,i-1}z_i - \tau^{-1}_{p,q,S} \sum_{j=p+1,j\notin S}^{q-1} (n_j/n_{0j})z_j - n_{0,p-1}(1 - \tau_{p+1,q,S}\tau^{-1}_{p,q,S})z_p\} \\
= \;& (n_i/n_{0i})\{n_{0,i-1}z_i - \tau^{-1}_{p,q,S} \sum_{j=p+1,j\notin S}^{q-1} (n_j/n_{0j})z_j - n_{0,p-1}\tau_{p,p+1}\tau^{-1}_{p,q,S}z_p\} \\
= \;& (n_i/n_{0i})\{n_{0,i-1}z_i - \tau^{-1}_{p,q,S} \sum_{j=p,j\notin S}^{q-1} (n_j/n_{0j})z_j\} \\
= \;& (n_i/n_{0i})\{n_{0,i-1}z_i - \tau^{-1}_{p,q,S}\eta_{p,q,S}\}.
\end{aligned}$$

It follows that

$$a^o_i = (n_i/n_{0i})\{\tau_{p,q}\tau^{-1}_{p,q,S} + b(n_{0,i-1}z_i - \eta_{p,q,S}\tau^{-1}_{p,q,S})\}, \quad \text{for } \; i \in T. \qquad (6.21)$$

For any $p \in R$,

$$\begin{aligned}
a^o_p \; = \;& [A^{-1}_{RR}\mathbf{1}_R - A^{-1}_{RR}A^{-1}_{RT}\mathbf{a}^o_T]_p \\
= \;& n_{0,p-1}\tau_{p,q} - \sum_{j=p+1,j\notin S}^{q-1} (n_{0,p-1}/n_{0,j-1})a^o_j \\
= \;& n_{0,p-1}\tau_{p,q} - \sum_{j=p+1,j\notin S}^{q-1} (n_{0,p-1}/n_{0,j-1})\{(n_j/n_{0j})[\tau_{p,q}\tau^{-1}_{p,q,S} + b(n_{0,j-1}z_j - \eta_{p,q,S}\tau^{-1}_{p,q,S})]\} \\
= \;& n_{0,p-1}\tau_{p,q}\{1 - \tau^{-1}_{p,q,S}\tau_{p+1,q,S}\} + bn_{0,p-1}\{(n_p/n_{0p})z_p - \eta_{p,q,S} + \tau^{-1}_{p,q,S}\tau_{p+1,q,S}\eta_{p,q,S}\}
\end{aligned}$$

$$
\begin{aligned}
&= n_{0,p-1}\tau_{p,q}\tau_{p,p+1}\tau_{p,q,S}^{-1} + bn_{0,p-1}\{(n_p/n_{0p})z_p - \eta_{p,q,S}\tau_{p,p+1}\tau_{p,q,S}^{-1}\} \\
&= n_p/n_{0p}\{\tau_{p,q}\tau_{p,q,S}^{-1} + b(n_{0,p-1}z_p - \eta_{p,q,S}\tau_{p,q,S}^{-1})\}.
\end{aligned}
$$

That is, (6.21) applies to all $i \in R \cup T$. Therefore, (6.16) follows.

For $A'^{-1}_{RR}(\mathbf{Z}_R - b^{-1}\Sigma_{RR}\mathbf{a}_R^o)$, if $p = r_1$, we have

$$
\begin{aligned}
[A'^{-1}_{RR}(\mathbf{Z}_R - b^{-1}\Sigma_{RR}\mathbf{a}_R^o)]_p &= z_p - b^{-1}n_{0p}/(n_p n_{0,p-1})a_p^o \\
&= n_{0,p-1}^{-1}\{\tau_{p,q,S}^{-1}\eta_{p,q,S} - b^{-1}\tau_{p,q}\tau_{p,q,S}^{-1}\};
\end{aligned}
$$

if $p \geq r_2$, we have

$$
\begin{aligned}
&[A'^{-1}_{RR}(\mathbf{Z}_R - b^{-1}\Sigma_{RR}\mathbf{a}_R^o)]_p \\
&= -(n_{0,t-1}/n_{0,p-1})[z_t - b^{-1}n_{0t}/(n_t n_{0,t-1})a_t^o] + z_p - b^{-1}n_{0p}/(n_p n_{0,p-1})a_p^o \\
&= n_{0,p-1}^{-1}\{-\tau_{t,p,S}^{-1}\eta_{t,p,S} + \tau_{p,q,S}^{-1}\eta_{p,q,S} - b^{-1}(-\tau_{t,p}\tau_{t,p,S}^{-1} + \tau_{p,q}\tau_{p,q,S}^{-1})\}.
\end{aligned}
$$

By the convention $\tau_{0,q} \equiv 0$ and $\tau_{0,q,S}^{-1} \equiv 0$, the condition (6.13) becomes

$$
n_{0,p-1}^{-1}\{-\tau_{t,p,S}^{-1}\eta_{t,p,S} + \tau_{p,q,S}^{-1}\eta_{p,q,S} - b^{-1}(\tau_{p,q}\tau_{p,q,S}^{-1} - \tau_{t,p}\tau_{t,p,S}^{-1})\} \geq 0 \qquad (6.22)
$$

for all $p \in R$.

Let t', p' and q' also be three consecutive indices in R. Consider the condition $A\mathbf{a} \leq 1$, for any $p \leq i < q$, we have

$$
\begin{aligned}
[A\mathbf{a}^o]_i &= \sum_{j=i}^{k} \alpha_{ij}a_j^o = \sum_{j=i}^{k}(n_{0,i-1}/n_{0,j-1})a_j^o \\
&= \sum_{j=i,j\notin S}^{q-1}(n_{0,i-1}/n_{0,j-1})a_j^o + \sum_{p'>i,p'\in R}\sum_{j=p',j\notin S}^{q'-1}(n_{0,i-1}/n_{0,j-1})a_j^o \\
&= \sum_{j=i,j\notin S}^{q-1}(n_{0,i-1}/n_{0,j-1})\{(n_j/n_{0j})[\tau_{p,q}\tau_{p,q,S}^{-1} + b(n_{0,j-1}z_j - \eta_{p,q,S}\tau_{p,q,S}^{-1})]\}
\end{aligned}
$$

$$+\sum_{p'>i,p'\in R}\sum_{j=p',j\notin S}^{q'-1}(n_{0,i-1}/n_{0,j-1})\{(n_j/n_{0j})[\tau_{p',q'}\tau_{p',q',S}^{-1}+b(n_{0,j-1}z_j-\eta_{p',q',S}\tau_{p',q',S}^{-1})]\}$$
$$= n_{0,i-1}\{\tau_{p,q}\tau_{p,q,S}^{-1}\tau_{i,q,S}+b\eta_{i,q,S}-b\eta_{p,q,S}\tau_{p,q,S}^{-1}\tau_{i,q,S}+\sum_{p'>i,p'\in R}\tau_{p',q'}\}$$
$$= n_{0,i-1}\{\tau_{p,q}\tau_{p,q,S}^{-1}\tau_{i,q,S}+\tau_{q,k+1}+b(\eta_{i,q,S}-\tau_{p,q,S}^{-1}\eta_{p,q,S}\tau_{i,q,S})\}.$$

The condition $A\mathbf{a}\leq 1$ becomes

$$\tau_{p,q}\tau_{p,q,S}^{-1}\tau_{i,q,S}+\tau_{q,k+1}+b(\eta_{i,q,S}-\tau_{p,q,S}^{-1}\eta_{p,q,S}\tau_{i,q,S})\leq n_{0,i-1}^{-1}. \tag{6.23}$$

If $i\in R$, (6.23) is an equality, otherwise, it is an inequality.

The condition (6.14) can be simplified as follows. For any $p<i<q, i\in S$,

$$[A'_{RS}A_{RR}^{-1}(\mathbf{Z}_R-b^{-1}\Sigma_{RR}\mathbf{a}_R^o)]_i$$
$$= \sum_{p'<i}(n_{0,p'-1}/n_{0,i-1})[A_{RR}^{-1}(\mathbf{Z}_R-b^{-1}\Sigma_{RR}\mathbf{a}_R^o)]_{p'}$$
$$= \sum_{p'<i}(n_{0,p'-1}/n_{0,i-1})n_{0,p'-1}^{-1}\{\tau_{p',q',S}^{-1}\eta_{p',q',S}-\tau_{t',p',S}^{-1}\eta_{t',p',S}+b^{-1}(-\tau_{p',q'}\tau_{p',q',S}^{-1}+\tau_{t',p'}\tau_{t',p',S}^{-1})\}$$
$$= b^{-1}n_{0,i-1}^{-1}\{\sum_{q'=r_2,r_3,\ldots,p}\tau_{p',q'}\tau_{p',q',S}^{-1}-\sum_{q'=r_2,r_3,\ldots,q}\tau_{p',q'}\tau_{p',q',S}^{-1}\}$$
$$-n_{0,i-1}^{-1}\{\sum_{q'=r_2,r_3,\ldots,p}\tau_{p',q',S}^{-1}\eta_{p',q',S}-\sum_{q'=r_2,r_3,\ldots,q}\tau_{p',q',S}^{-1}\eta_{p',q',S}\}$$
$$= -n_{0,i-1}^{-1}\tau_{p,q,S}^{-1}\{b^{-1}\tau_{p,q}-\eta_{p,q,S}\}.$$

The condition (6.14) becomes for any $i\in S$

$$n_{0,i-1}^{-1}\tau_{p,q,S}^{-1}\{\eta_{p,q,S}-b^{-1}\tau_{p,q}\}\geq z_i. \tag{6.24}$$

Furthermore,

$$\mathbf{a}^{o'}\Sigma\mathbf{a}^o = \sum_{i=1}^{k}a_i^{o2}\sigma_{ii}$$
$$= \sum_{i=1,i\notin S}^{k}(n_i/n_{0i})^2\{\tau_{p,q}\tau_{p,q,S}^{-1}+b(n_{0,i-1}z_i-\eta_{p,q,S}\tau_{p,q,S}^{-1})\}^2 n_{0i}/(n_i n_{0,i-1})$$

$$= \sum_{i=1,i\notin S}^{k} (n_i/n_{0i})^2 \tau_{p,q}^2 \tau_{p,q,S}^{-2} n_{0i}/(n_i n_{0,i-1})$$

$$+b^2 \sum_{i=1,i\notin S}^{k} (n_i/n_{0i})^2 (n_{0,i-1}z_i - \eta_{p,q,S}\tau_{p,q,S}^{-1})^2 n_{0i}/(n_i n_{0,i-1})$$

$$+2b \sum_{i=1,i\notin S}^{k} (n_i/n_{0i})^2 \tau_{p,q}\tau_{p,q,S}^{-1}(n_{0,i-1}z_i - \eta_{p,q,S}\tau_{p,q,S}^{-1}) n_{0i}/(n_i n_{0,i-1})$$

$$= \sum_{p=r_0,\ldots,r_m} \{\tau_{p,q}^2 \tau_{p,q,S}^{-2}\tau_{p,q,S} + b^2 \sum_{i=p,i\notin S}^{q-1} (n_i n_{0,i-1}/n_{0i}) z_i^2$$

$$+b^2 \eta_{p,q,S}^2 \tau_{p,q,S}^{-2} \sum_{i=p,i\notin S}^{q-1} n_i/(n_{0,i-1}n_{0i}) - 2b^2 \eta_{p,q,S}\tau_{p,q,S}^{-1} \sum_{i=p,i\notin S}^{q-1} (n_i/n_{0i}) z_i\}$$

$$= \sum_{p=r_0,\ldots,r_m} \{\tau_{p,q}^2 \tau_{p,q,S}^{-1} + b^2 [\sum_{i=p,i\notin S}^{q-1} (n_i n_{0,i-1}/n_{0i}) z_i^2 - \eta_{p,q,S}^2 \tau_{p,q,S}^{-1}]\}.$$

Note the cross product term $\sum_{i=1}^{k}(n_i/n_{0i})^2\tau_{p,q}\tau_{p,q,S}^{-1}(n_{0,i-1}z_i - \eta_{p,q,S}\tau_{p,q,S}^{-1})^2 n_{0i}/(n_i n_{0,i-1})$ is equal to zero. Since $b = (\mathbf{a}^{o\prime}\Sigma\mathbf{a}^o)^{1/2}(t_\alpha s)^{-1}$, we have

$$b^2 = \sum_{p\in R} \tau_{p,q}^2 \tau_{p,q,S}^{-1} / \{(t_\alpha s)^2 - \sum_{i=1,i\notin S}^{k} (n_i n_{0,i-1}/n_{0i}) z_i^2 + \sum_{p\in R} \eta_{p,q,S}^2 \tau_{p,q,S}^{-1}\}. \quad (6.25)$$

6.5.2 Justification of the Algorithm

(A.) The R^1, S^1 and T^1 form an optimal solution.

First we shall show $R^1 = \{p\}, S^1 = \{i : z_i \leq 0\} = \{i : a_i^{(0)} = c_i^{(0)} - \bar{c}_{0,i-1}^{(0)} = 0\}$ and $T^1 = \{i : z_i > 0, i \neq p\}$ is the optimal solution. Let $M = \sum_{j=p,j\notin S^1}^{k} n_j P(\bar{\mathbf{Y}}|\mathbf{O}_k)_j = n_{0,p-1}\sum_{j=p,j\notin S^1}^{k}(n_j/n_{0,j-1})z_j$. For $i \neq p$ we have

$$\sum_{j=i}^{k} n_j P(\bar{\mathbf{Y}}|\mathbf{O}_k)_j = n_{0,i-1} \sum_{j=i,j\notin S^1}^{k} (n_j/n_{0,j-1}) z_j < M.$$

For $\tau_{p,k+1} = 1/n_{0,p-1}, \eta_{p,k+1,S^1} = \sum_{j=p,j\notin S^1}^{k}(n_j/n_{0,j-1})z_j = M\tau_{p,k+1}$ and

$$\tau_{p,k+1,S^1} = \sum_{j=p,j\notin S^1}^{k} (\frac{1}{n_{0,j-1}} - \frac{1}{n_{0j}}) \leq \frac{1}{n_{0,p-1}} - \frac{1}{n_{0k}} < \tau_{p,k+1},$$

we have

$$\mathbf{a}^{(0)\prime}\Sigma\mathbf{a}^{(0)} = \tau^2_{p,k+1}\tau^{-1}_{p,k+1,S^1} + b^2(t_0^2 s^2 - \eta^2_{p,k+1,S^1}\tau^{-1}_{p,k+1,S^1}) = b^2 {t_\alpha}^2 s^2$$

where $t_0 = \{\sum_{j=0}^k n_j P(\bar{\mathbf{Y}}|\mathbf{O}_k)_j^2\}^{1/2}/s = (\mathbf{Z}^{+\prime}\Sigma^{-1}\mathbf{Z}^{+})^{1/2}/s$.

Consider $t_0 - \epsilon < t_\alpha < t_0$, as

$$b^2 = \tau^2_{p,k+1}\tau^{-1}_{p,k+1,S^1}/\{t_\alpha^2 s^2 - (t_0^2 s^2 - M^2\tau^2_{p,k+1}\tau^{-1}_{p,k+1,S^1})\},$$

we have $M^{-1} < b < \{M^2 + \epsilon(\epsilon - 2t_0)/(\tau^2_{p,k+1}\tau^{-1}_{p,k+1,S^1})\}^{-1/2}$. We denote the right hand side of the inequality as $M^{-1} + \delta$.

If $i \in T^1, i < p$, we have $a_i^{(0)} = (n_i/n_{0i})n_{0,i-1}z_i b > 0$. For $i \in R^1 \cup T^1, i \geq p$, we have

$$\begin{aligned} a_i^{(0)} &= \frac{n_i}{n_{0i}}\{\tau_{p,k+1}\tau^{-1}_{p,k+1,S^1} + b(n_{0,i-1}z_i - \eta_{p,k+1,S^1}\tau^{-1}_{p,k+1,S^1})\} \\ &= \frac{n_i}{n_{0i}}\{n_{0,i-1}z_i b - (bM-1)\tau_{p,k+1}\tau^{-1}_{p,k+1,S^1}\} > 0. \end{aligned}$$

Consider the condition $A\mathbf{a} \leq 1$, for the index $p \in R^1$, we have

$$\begin{aligned} [A\mathbf{a}^{(0)}]_p &= \sum_{j=p}^k n_{0,p-1}a_j^{(0)}/n_{0,j-1} \\ &= \sum_{j=p, j\notin S^1}^k (n_{0,p-1}/n_{0,j-1})(n_j/n_{0j})n_{0,j-1}z_j b \\ &\quad - \sum_{j=p, j\notin S}^k (n_{0,p-1}/n_{0,j-1})(n_j/n_{0j})(bM-1)\tau_{p,k+1}\tau^{-1}_{p,k+1,S^1} \\ &= n_{0,p-1}\eta_{p,k+1,S^1}b - n_{0,p-1}\tau_{p,k+1,S^1}(bM-1)\tau_{p,k+1}\tau^{-1}_{p,k+1,S^1} \\ &= 1. \end{aligned}$$

For $i > p$, when $b = M^{-1}$, we have

$$\begin{aligned}[A\mathbf{a}^{(0)}]_i &= \sum_{j=i,j\notin S^1}^{k} (n_{0,i-1}/n_{0,j-1})(n_j/n_{0j})n_{0,j-1}z_j b \\ &\quad - \sum_{j=i,j\notin S^1}^{k} (n_{0,i-1}/n_{0,j-1})(n_j/n_{0j})(bM-1)\tau_{p,k+1}\tau_{p,k+1,S^1}^{-1} < 1.\end{aligned}$$

Hence, $[A\mathbf{a}^{(0)}]_i < 1$ if $M^{-1} < b < M^{-1} + \delta$.

For $i < p$, when $b = M^{-1}$, we have

$$\begin{aligned}[A\mathbf{a}^{(0)}]_i &= \sum_{j=i,j\notin S^1}^{p-1} (n_{0,i-1}/n_{0,j-1})a_j^{(0)} + n_{0,i-1}/n_{0,p-1}[A\mathbf{a}^{(0)}]_p \\ &= \sum_{j=i,j\notin S^1}^{p-1} (n_{0,i-1}/n_{0,j-1})(n_j/n_{0j})n_{0,j-1}z_j b + n_{0,i-1}/n_{0,p-1} \\ &< 1.\end{aligned}$$

Thus, $[A\mathbf{a}^{(0)}]_i < 1$ if $M^{-1} < b < M^{-1} + \delta$.

For condition (6.13), we have

$$n_{0,p-1}^{-1}(\eta_{p,k+1,S^1}\tau_{p,k+1,S^1}^{-1} - b^{-1}\tau_{p,k+1}\tau_{p,k+1,S^1}^{-1}) = M\tau_{p,k+1}(b - M^{-1})/(bn_{0,p-1}\tau_{p,k+1,S^1}) > 0.$$

The above inequality is also equivalent to the condition (6.14). Therefore, the initial R^1, S^1 and T^1 satisfy the sufficient and necessary condition in Theorem 6.4.1.

(B.) We will show that $R^{r+1} \supseteq R^r$ and $S^{r+1} \supseteq S^r$.

Let t, p and q be three consecutive indices in R^r. We prove $S^r \subseteq S^{r+1}$ first. Let $j \in S^r$ and $p < j < q$. Then by (6.24), we have that

$$\tau_{p,q,S^r}^{-1}(-b^{-1}\tau_{p,q} + \eta_{p,q,S^r}) \geq n_{0,j-1}z_j.$$

Therefore,

$$-n_{0,j-1}z_j + \eta_{p,q,S^r}\tau^{-1}_{p,q,S^r} \geq b^{-1}\tau_{p,q}\tau^{-1}_{p,q,S^r} \geq 0.$$

Hence we have

$$b \geq -\tau_{p,q}\tau^{-1}_{p,q,S^r}(n_{0,j-1}z_j - \eta_{p,q,S^r}\tau^{-1}_{p,q,S^r})^{-1}$$

Suppose that at Step $r+1, j \in T^{r+1}$, from (6.21), we have

$$\begin{aligned} a^o_j &= (n_j/n_{0j})\{\tau_{p,q}\tau^{-1}_{p,q,S^{r+1}} + b(n_{0,j-1}z_j - \eta_{p,q,S^{r+1}}\tau^{-1}_{p,q,S^{r+1}})\} \\ &\leq (n_j/n_{0j})\{\tau_{p,q}\tau^{-1}_{p,q,S^{r+1}} - \tau_{p,q}\tau^{-1}_{p,q,S^{r+1}}(n_{0,j-1}z_j - \eta_{p,q,S^{r+1}}\tau^{-1}_{p,q,S^{r+1}})^{-1} \\ &\quad \cdot(n_{0,j-1}z_j - \eta_{p,q,S^{r+1}}\tau^{-1}_{p,q,S^{r+1}})\} \\ &= 0. \end{aligned}$$

This contradicts that $j \in T^{r+1}$. It follows that $S^r \subseteq S^{r+1}$.

Let $p \in R^r$, from (6.22), we have that

$$-b(\tau^{-1}_{p,q,S^r}\eta_{p,q,S^r} - \tau^{-1}_{t,p,S^r}\eta_{t,p,S^r}) \leq (\tau_{t,p}\tau^{-1}_{t,p,S^r} - \tau_{p,q}\tau^{-1}_{p,q,S^r}). \tag{6.26}$$

Suppose that at Step $r+1, p \in T^{r+1}$ and without loss of generality we assume that $S^r = S^{r+1}$. By $[A\mathbf{a}^o]_p < 1$ and (6.23), we have

$$\begin{aligned} b(\eta_{p,q,S^r} - \tau^{-1}_{t,q,S^r}\tau_{p,q,S^r}\eta_{t,q,S^r}) &< n^{-1}_{0,p-1} - \tau_{t,q}\tau^{-1}_{t,q,S^r}\tau_{p,q,S^r} - \tau_{q,k+1} \\ &= \tau_{p,q} - \tau_{t,q}\tau^{-1}_{t,q,S^r}\tau_{p,q,S^r}. \end{aligned}$$

Multiplying it by $-\tau^{-1}_{p,q,S^r}$, we have

$$-b(\tau^{-1}_{p,q,S^r}\eta_{p,q,S^r} - \tau^{-1}_{t,q,S^r}\eta_{t,q,S^r}) > \tau_{t,q}\tau^{-1}_{t,q,S^r} - \tau_{p,q}\tau^{-1}_{p,q,S^r}. \tag{6.27}$$

Also assume $t \in R^{r+1}$, by

$$[A\mathbf{a}^o]_t = 1, \quad [A\mathbf{a}^o]_p < 1$$

and

$$[A\mathbf{a}^o]_t = \sum_{j=t,j\notin S^t}^{p-1} (n_{0,t-1}/n_{0,j-1})a_j^o + (n_{0,t-1}/n_{0,p-1})[A\mathbf{a}^o]_p,$$

we have that

$$\begin{aligned}\sum_{j=t,j\notin S^t}^{p-1} n_{0,j-1}^{-1}a_j^o &= n_{0,t-1}^{-1}\{1-(n_{0,t-1}/n_{0,p-1})[A\mathbf{a}^o]_p\} \\ &> n_{0,t-1}^{-1}\{1-(n_{0,t-1}/n_{0,p-1})\} = \tau_{t,p}.\end{aligned}$$

That is

$$\begin{aligned}&\sum_{j=t,j\notin S^r}^{p-1} n_{0,j-1}^{-1}(n_j/n_{0j})\{\tau_{t,q}\tau_{t,q,S^r}^{-1} + b(n_{0,j-1}z_j - \eta_{t,q,S^r}\tau_{t,q,S^r}^{-1})\} \\ = \; &\tau_{t,q}\tau_{t,q,S^r}^{-1}\tau_{t,p,S^r} + b(\eta_{t,p,S^r} - \eta_{t,q,S^r}\tau_{t,q,S^r}^{-1}\tau_{t,p,S^r}) > \tau_{t,p}.\end{aligned}$$

Therefore, we have

$$-b(\eta_{t,q,S^r}\tau_{t,q,S^r}^{-1} - \eta_{t,p,S^r}\tau_{t,p,S^r}^{-1}) > \tau_{t,p}\tau_{t,p,S^r}^{-1} - \tau_{t,q}\tau_{t,q,S^r}^{-1}. \tag{6.28}$$

Summing the inequalities (6.27) and (6.28), we have

$$-b(\tau_{p,q,S^r}^{-1}\eta_{p,q,S^r} - \tau_{t,p,S^r}^{-1}\eta_{t,p,S^r}) > (\tau_{t,p}\tau_{t,p,S^r}^{-1} - \tau_{p,q}\tau_{p,q,S^r}^{-1}),$$

which contradicts (6.26). It follows that $R^r \subseteq R^{r+1}$.

Therefore, the algorithm terminates at no more than k steps.

Let R^r, S^r and T^r be the optimal partition satisfying (6.11) – (6.14) of Theorem 6.5.1 for a given $t_\alpha > t_r$ where t_r corresponds to the confidence level α_r. As α decreases, the optimal solution holds at $t_\alpha \geq t_r$ until either

(I) there exists a $p \in T^r$ so that $R^{r+1} = R^r \cup \{p\}, S^{r+1} = S^r$ and $T^{r+1} = T^r - \{p\}$ is the optimal partition for $t_\alpha < t_r$, or

(II) there exists an $h \in T^r$ so that $R^{r+1} = R^r, S^{r+1} = S^r \cup \{h\}$ and $T^{r+1} = T^r - \{h\}$ is the optimal partition for $t_\alpha < t_r$.

(C.) Continuity of b.

We have that $b_r^2 = A_r/(t_r^2 s^2 - B_r)$ where

$$A_r = \sum_{p'=r_0,\ldots,r_m} \tau^{-1}_{p'q',S^r}\tau^2_{p',q'}$$

and

$$B_r = \sum_{p'=r_0,\ldots,r_m} \{ \sum_{i=p',i\notin S^r}^{q'-1} (n_i n_{0,i-1}/n_{0i}) z_i^2 - \eta^2_{p',q',S^r}\tau^{-1}_{p'q',S^r} \}.$$

For Case (I), we assume $t < p < q$ and $t, q \in R^r$. Then, by $[A\mathbf{a}]_p < 1$ and (6.23), we have

$$b_r(\tau^{-1}_{p,q,S^r}\eta_{p,q,S^r} - \tau^{-1}_{t,q,S^r}\eta_{t,q,S^r}) < \tau_{p,q}\tau^{-1}_{p,q,S^r} - \tau_{t,q}\tau^{-1}_{t,q,S^r}. \tag{6.29}$$

For $t_\alpha < t_r$, we have that $b^2(\alpha) = A_{r+1}/(t_\alpha^2 s^2 - B_{r+1})$ where $A_{r+1} = A_r - b_r^2\Delta$ and $B_{r+1} = B_r + \Delta$ with

$$\Delta = \eta^2_{t,q,S^r}\tau^{-1}_{t,q,S^r} - \eta^2_{t,p,S^r}\tau^{-1}_{t,p,S^r} - \eta^2_{p,q,S^r}\tau^{-1}_{p,q,S^r}.$$

Therefore, $\lim_{\alpha\to\alpha_r+} b^2(\alpha) = A_{r+1}/(t_r^2 s^2 - B_{r+1})$.

Firstly, we will show that $A_{r+1} = A_r - b_r^2\Delta$. As $A_{r+1} = A_r - (\tau^{-1}_{t,q,S^r}\tau^2_{t,q} - \tau^{-1}_{p,q,S^r}\tau^2_{p,q} - \tau^{-1}_{t,p,S^r}\tau^2_{t,p})$, then we only need to show that

$$b_r^2\Delta = \tau^{-1}_{t,q,S^r}\tau^2_{t,q} - \tau^{-1}_{p,q,S^r}\tau^2_{p,q} - \tau^{-1}_{t,p,S^r}\tau^2_{t,p}. \tag{6.30}$$

By (6.29), we show that

$$\{\tau^{-1}_{t,q,S^r}\tau^2_{t,q} - \tau^{-1}_{p,q,S^r}\tau^2_{p,q} - \tau^{-1}_{t,p,S^r}\tau^2_{t,p}\}(\tau^{-1}_{p,q,S^r}\eta_{p,q,S^r} - \tau^{-1}_{t,q,S^r}\eta_{t,q,S^r})^2 \tag{6.31}$$

and

$$(\eta^2_{t,q,S^r}\tau^{-1}_{t,q,S^r} - \eta^2_{t,p,S^r}\tau^{-1}_{t,p,S^r} - \eta^2_{p,q,S^r}\tau^{-1}_{p,q,S^r})(\tau_{t,q}\tau^{-1}_{t,q,S^r} - \tau_{p,q}\tau^{-1}_{p,q,S^r})^2 \qquad (6.32)$$

are equal.

By $\eta_{t,p,S} = \eta_{t,q,S} - \eta_{p,q,S}$ and $\tau_{t,q,S} = \tau_{t,p,S} + \tau_{p,q,S}$, the expression (6.32) can be rewritten and expanded as

$$\begin{aligned}
&\{\eta^2_{t,q,S^r}(\tau^{-1}_{t,q,S^r} - \tau^{-1}_{t,p,S^r}) + 2\eta_{t,q,S^r}\eta_{p,q,S^r}\tau^{-1}_{t,p,S^r} - \eta^2_{p,q,S^r}(\tau^{-1}_{p,q,S^r} + \tau^{-1}_{t,p,S^r})\} \\
&\quad \cdot (\tau_{t,q}\tau^{-1}_{t,q,S^r} - \tau_{p,q}\tau^{-1}_{p,q,S^r})^2 \\
= &\{-\eta^2_{t,q,S^r}\tau^{-1}_{t,q,S^r}\tau_{p,q,S^r}\tau^{-1}_{t,p,S^r} - \eta^2_{p,q,S^r}\tau^{-1}_{p,q,S^r}\tau_{t,q,S^r}\tau^{-1}_{t,p,S^r} \\
&+2\eta_{t,q,S^r}\eta_{p,q,S^r}\tau^{-1}_{t,p,S^r}\}(\tau_{t,q}\tau^{-1}_{t,q,S^r} - \tau_{p,q}\tau^{-1}_{p,q,S^r})^2 \\
= &\eta^2_{t,q,S^r}\{-\tau^2_{t,q}\tau^{-3}_{t,q,S^r}\tau_{p,q,S^r}\tau^{-1}_{t,p,S^r} - \tau^2_{p,q}\tau^{-1}_{p,q,S^r}\tau^{-1}_{t,q,S^r}\tau^{-1}_{t,p,S^r} + 2\tau_{t,q}\tau_{p,q}\tau^{-2}_{t,q,S^r}\tau^{-1}_{t,p,S^r}\} \\
&\eta^2_{p,q,S^r}\{-\tau^2_{t,q}\tau^{-1}_{t,q,S^r}\tau^{-1}_{t,p,S^r}\tau^{-1}_{p,q,S^r} - \tau^2_{p,q}\tau^{-3}_{p,q,S^r}\tau^{-1}_{t,p,S^r}\tau_{t,q,S^r} + 2\tau_{t,q}\tau_{p,q}\tau^{-1}_{t,p,S^r}\tau^{-2}_{p,q,S^r}\} \\
&+2\eta_{t,q,S^r}\eta_{p,q,S^r}\tau^{-1}_{t,p,S^r}\{\tau^2_{t,q}\tau^{-2}_{t,q,S^r} + \tau^2_{p,q}\tau^{-2}_{p,q,S^r} - 2\tau_{t,q}\tau_{p,q}\tau^{-1}_{t,q,S^r}\tau^{-1}_{p,q,S^r}\}
\end{aligned}$$

Consider the coefficient for η^2_{t,q,S^r} in the expression (6.31). By $\tau_{t,p} = \tau_{t,q} - \tau_{p,q}$, we have

$$\begin{aligned}
&\tau^{-2}_{t,q,S^r}(\tau^{-1}_{t,q,S^r}\tau^2_{t,q} - \tau^{-1}_{p,q,S^r}\tau^2_{p,q} - \tau^{-1}_{t,p,S^r}\tau^2_{t,p}) \\
= &\tau^{-2}_{t,q,S^r}(\tau^{-1}_{t,q,S^r}\tau^2_{t,q} - \tau^{-1}_{p,q,S^r}\tau^2_{p,q} - \tau^{-1}_{t,p,S^r}\tau^2_{t,q} - \tau^{-1}_{t,p,S^r}\tau^2_{p,q} + 2\tau^{-1}_{t,p,S^r}\tau_{t,q}\tau_{p,q}) \\
= &\tau^2_{t,q}\tau^{-2}_{t,q,S^r}(\tau^{-1}_{t,q,S^r} - \tau^{-1}_{t,p,S^r}) - \tau^2_{p,q}\tau^{-2}_{t,q,S^r}(\tau^{-1}_{p,q,S^r} + \tau^{-1}_{t,p,S^r}) + 2\tau_{t,q}\tau_{p,q}\tau^{-2}_{t,q,S^r}\tau^{-1}_{t,p,S^r} \\
= &-\tau^2_{t,q}\tau^{-3}_{t,q,S^r}\tau^{-1}_{t,p,S^r}\tau_{p,q,S^r} - \tau^2_{p,q}\tau^{-1}_{t,q,S^r}\tau^{-1}_{t,p,S^r}\tau^{-1}_{p,q,S^r} + 2\tau_{t,q}\tau_{p,q}\tau^{-2}_{t,q,S^r}\tau^{-1}_{t,p,S^r}
\end{aligned}$$

which is the same as the coefficient for η^2_{t,q,S^r} in the expression (6.32). Similarly, the corresponding coefficients for η^2_{p,q,S^r} and $2\eta_{t,q,S^t}\eta_{p,q,S^t}$ in (6.31) and (6.32) are equal, that is (6.30) is proven. However,

$$\frac{A_{r+1}}{s^2t_r^2 - B_{r+1}} = \frac{A_r - b_r^2\Delta}{s^2t_r^2 - B_r - \Delta} = \frac{A_r - b_r^2\Delta}{A_r/b_r^2 - \Delta} = b_r^2.$$

It follows that $\lim_{\alpha\to\alpha_r-} b(\alpha) = b_r$. Hence, the coefficient $b(\alpha)$ is a continuous function of α.

For Case (II), we assume $p < h < q$ and $p, q \in R^r$. By the condition $a_h^o < 0$, we have $A_{r+1} = A_r - b_r^2\Delta$ and $B_{r+1} = B_r - \Delta$ where

$$b_r = -\tau_{p,q}\tau_{p,q,S^r}^{-1}/(n_{0,h-1}z_h - \eta_{p,q,S^r}\tau_{p,q,S^r}^{-1})$$

and

$$\Delta = -(n_h n_{0,h-1}/n_{0h})z_h^2 + \eta_{p,q,S^r}^2\tau_{p,q,S^r}^{-1} - \eta_{p,q,S^{r+1}}^2\tau_{p,q,S^{r+1}}^{-1}.$$

Note that $\tau_{p,q,S^{r+1}} = \tau_{p,q,S^r} + n_h/(n_{0,h-1}n_{0h})$, $\eta_{p,q,S^{r+1}} = \eta_{p,q,S^r} + (n_h/n_{0h})z_h$ and $\tau_{p,q.S^{r+1}} = \tau_{p,q,S^r} - n_h/(n_{0,h-1}n_{0h})$. Therefore, we have

$$\begin{aligned} A_{r+1} &= A_r - (\tau_{p,q,S^r}^{-1}\tau_{p,q}^2 - \tau_{p,q,S^{r+1}}^{-1}\tau_{p,q}^2) \\ &= A_r + \tau_{p,q}^2(-\tau_{p,q,S^r}^{-1} + \tau_{p,q,S^{r+1}}^{-1}) \\ &= A_r + \tau_{p,q}^2\tau_{p,q,S^{r+1}}^{-1}\tau_{p,q,S^r}^{-1}n_h/(n_{0,h-1}n_{0h}). \end{aligned}$$

Next, we show that

$$b_r^2\Delta = -\tau_{p,q}^2\tau_{p,q,S^{r+1}}^{-1}\tau_{p,q,S^r}^{-1}n_h/(n_{0,h-1}n_{0h}).$$

That is,

$$\{\tau_{p,q}^2\tau_{p,q,S^{r+1}}^{-1}\tau_{p,q,S^r}^{-1}n_h/(n_{0,h-1}n_{0h})\}(n_{0,h-1}z_h - \eta_{p,q,S^r}\tau_{p,q,S^r}^{-1})^2 \tag{6.33}$$

and

$$\{(n_h n_{0,h-1}/n_{0h})z_h^2 - \eta_{p,q,S^r}^2\tau_{p,q,S^r}^{-1} + \eta_{p,q,S^{r+1}}^2\tau_{p,q,S^{r+1}}^{-1}\}(\tau_{p,q}\tau_{p,q,S^r}^{-1})^2 \tag{6.34}$$

should be equal. The expression (6.33) is equivalent to

$$z_h^2\tau_{p,q}^2\tau_{p,q,S^{r+1}}^{-1}\tau_{p,q,S^r}^{-1}(n_hn_{0,h-1}/n_{0h})+\eta_{p,q,S^r}^2\tau_{p,q,S^{r+1}}^{-1}n_h/(n_{0,h-1}n_{0h})\tau_{p,q,S^r}^{-3}\tau_{p,q}^2$$
$$+\,2\eta_{p,q,S^r}z_h(n_h/n_{0h})\tau_{p,q,S^{r+1}}^{-1}\tau_{p,q,S^r}^{-2}\tau_{p,q}^2.$$

The expression (6.34) is equivalent to

$$\begin{aligned}
&\{(n_hn_{0,h-1}/n_{0h})z_h^2-\eta_{p,q,S^r}^2\tau_{p,q,S^r}^{-1}+[\eta_{p,q,S^r}-(n_h/n_{0h})z_h]^2\tau_{p,q,S^{r+1}}^{-1}\}(\tau_{p,q}\tau_{p,q,S^r}^{-1})^2\\
=\;&\{[(n_hn_{0,h-1}/n_{0h})+(n_h/n_{0h})^2\tau_{p,q,S^{r+1}}^{-1}]z_h^2+\eta_{p,q,S^r}^2(-\tau_{p,q,S^r}^{-1}+\tau_{p,q,S^{r+1}}^{-1})\\
&-2\eta_{p,q,S^r}z_h(n_h/n_{0h})\tau_{p,q,S^{r+1}}^{-1}\}(\tau_{p,q}\tau_{p,q,S^r}^{-1})^2\\
=\;&\eta_{p,q,S^r}^2\tau_{p,q,S^{r+1}}^{-1}n_h/(n_{0,h-1}n_{0h})\tau_{p,q,S^r}^{-3}\tau_{p,q}^2\\
&+z_h^2\{(n_hn_{0,h-1}/n_{0h})+(n_h/n_{0h})^2\tau_{p,q,S^{r+1}}^{-1}\}\tau_{p,q,S^r}^{-2}\tau_{p,q}^2\\
&-2\eta_{p,q,S^r}z_h(n_h/n_{0h})\tau_{p,q,S^{r+1}}^{-1}\tau_{p,q,S^r}^{-2}\tau_{p,q}^2\\
=\;&\eta_{p,q,S^r}^2\tau_{p,q,S^{r+1}}^{-1}n_h/(n_{0,h-1}n_{0h})\tau_{p,q,S^r}^{-3}\tau_{p,q}^2\\
&+z_h^2\tau_{p,q,S^r}^{-2}\tau_{p,q}^2\tau_{p,q,S^{r+1}}^{-1}\{(n_hn_{0,h-1}/n_{0h})\tau_{p,q,S^{r+1}}+(n_h/n_{0h})^2\}\\
&-2\eta_{p,q,S^r}z_h(n_h/n_{0h})\tau_{p,q,S^{r+1}}^{-1}\tau_{p,q,S^r}^{-2}\tau_{p,q}^2\\
=\;&\eta_{p,q,S^r}^2\tau_{p,q,S^{r+1}}^{-1}n_h/(n_{0,h-1}n_{0h})\tau_{p,q,S^r}^{-3}\tau_{p,q}^2\\
&+z_h^2\tau_{p,q,S^r}^{-2}\tau_{p,q}^2\tau_{p,q,S^{r+1}}^{-1}\{(n_hn_{0,h-1}/n_{0h})[\tau_{p,q,S^r}-n_h/(n_{0h}n_{0,h-1})]+(n_h/n_{0h})^2\}\\
&-2\eta_{p,q,S^r}z_h(n_h/n_{0h})\tau_{p,q,S^{r+1}}^{-1}\tau_{p,q,S^r}^{-2}\tau_{p,q}^2\\
=\;&+\eta_{p,q,S^r}^2\tau_{p,q,S^{r+1}}^{-1}n_h/(n_{0,h-1}n_{0h})\tau_{p,q,S^r}^{-3}\tau_{p,q}^2\\
&+z_h^2\tau_{p,q,S^r}^{-1}\tau_{p,q}^2\tau_{p,q,S^{r+1}}^{-1}(n_hn_{0,h-1}/n_{0h})\\
&-2\eta_{p,q,S^r}z_h(n_h/n_{0h})\tau_{p,q,S^{r+1}}^{-1}\tau_{p,q,S^r}^{-2}\tau_{p,q}^2.
\end{aligned}$$

Therefore, expressions (6.33) and (6.34) are equal. By a similar discussion as Case (I), we prove that $b(\alpha)$ is a continuous function of α for Case (II). This completes the proof.

6.6 Discussion

If several dose response means are compared with the control mean and the experimenter has a prior knowledge that the response means are monotone nondecreasing, a test procedure is available that has good properties under this simple order alternative, hence improving confidence bounds. The orthant test T_k^o introduced in this article is an effective method for testing the equality of the response means against the simple order alternative and constructing one-sided simultaneous confidence lower bounds for $\mu_j - \mu_0$. The proposed test is easy to implement and its p-value is a mixture of F tail probabilities. Furthermore, an efficient algorithm is given to compute the confidence lower bound.

Table 6.1: Probabilities (in Percentage) of Detecting the Difference Between μ_j and μ_0 for $k = 5, \alpha = 0.05, \nu = \infty$

C		j				
		1	2	3	4	5
1.5	D_j	2.9	52.7	78.4	87.6	91.9
	T_j^{st}	0.4	34.5	73.8	89.2	95.2
	T_j^{so}	2.3	50.5	81.3	91.3	95.2
	T_j^{o}	0.8	45.3	82.1	93.6	97.5
2.0	D_j	1.6	21.5	64.3	82.1	89.4
	T_j^{st}	0.2	10.3	57.7	85.5	94.9
	T_j^{so}	1.2	19.7	66.9	87.9	95.0
	T_j^{o}	0.4	16.2	68.5	91.1	97.4
2.5	D_j	1.4	7.4	45.3	74.3	86.3
	T_j^{st}	0.2	2.3	36.3	78.9	94.3
	T_j^{so}	1.1	5.9	45.1	81.6	94.2
	T_j^{o}	0.3	4.3	47.2	86.0	96.9
3.0	D_j	1.3	4.1	25.2	63.3	82.7
	T_j^{st}	0.1	1.0	17.3	67.2	92.6
	T_j^{so}	1.0	3.0	23.4	70.1	92.3
	T_j^{o}	0.3	1.9	25.2	76.4	95.9
3.5	D_j	1.3	3.1	13.3	50.5	79.3
	T_j^{st}	0.1	0.7	7.5	52.2	90.4
	T_j^{so}	1.0	2.3	11.1	54.8	89.8
	T_j^{o}	0.3	1.4	12.0	62.6	94.3

Chapter 7

A Stepwise Multiple Test Procedure

We continue to consider the problem of identifying the lowest dose level for which the mean response differs from the zero dose level in the dose-response studies. Ruberg (1989) referred to this dose as the *minimum effective dose* (MED). However, test procedures only find the *minimum detectable dose* (MDD). In dose-response studies, the response means $\mu_1, ..., \mu_k$ correspond to increasing doses of a substance and μ_0 corresponds to the zero dose. It is desirable for a method to not declare a lower dose to be efficacious if it does not declare a higher dose to be efficacious. This can be achieved by testing the null hypothesis $H_{0j} : \mu_i = \mu_0, i = 1, ..., j$, against the alternative hypothesis $H_{1j}^{st} : \mu_i \geq \mu_0, i = 1, ..., j$, with at least one strict inequality in a stepwise fashion starting from $j = k$, continuing only while H_{0j} is rejected. Tamhane, Hochberg and Dunnett (1996) studied various stepwise procedures including Williams' (1971) procedure and a class of stepwise procedures based on contrasts. Only Williams' procedure utilized the monotonicity assumption of the response means. The stepwise confidence intervals based on a pairwise t test statistic can be found in Hsu and Berger (1999), and they used a fundamentally different confidence set-based justification by partitioning the parameter space naturally and using the principle that exactly one member of the partition contains the true parameter.

By incorporating the assumption that $\mu_0 \leq \mu_1 \leq \cdots \leq \mu_k$, we will consider both likelihood ratio test and multiple comparison tests in a stepwise procedure in this chapter. It will be demonstrated by a simulation study that the prior knowledge of a monotone trend will provide us with more efficient test procedures. In Section 7.1, the stepwise testing procedure will be proposed. The simulation study to compare the probabilities of detecting the MDD are given in Section 7.2.

7.1 A Stepwise Test Procedure

Denote a set of increasing dose levels by 0, 1, 2, ..., k, where 0 corresponds to the zero dose level. Consider a one-way layout setting in which n_i experimental units are tested at the ith dose level, $i = 0, 1, ..., k$. We assume that all observations Y_{ij} are mutually independent with $Y_{ij} \sim N(\mu_i, \sigma^2), i = 0, 1, ..., k$ and $j = 1, 2, ..., n_i$. Let $\bar{Y}_i \sim N(\mu_i, \sigma^2/n_i), i = 0, 1, ..., k$, be the sample means, and let $s^2 = \sum_{i=0}^{k} \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2/\nu$ be an unbiased estimate of the common variance σ^2 based on $\nu = \sum_{i=0}^{k} n_i - (k+1) > 0$ degrees of freedom and distributed as $\sigma^2\chi^2_\nu/\nu$, independent of the $\bar{Y}_i$. For simplicity, we restrict our study to the case when sample sizes of the non-zero dose levels are the same. We assume that the common sample size is n.

Suppose that a larger μ_i indicates a better average response and the response means are monotone nondecreasing. We define MED as the minimum dose i such that $\mu_i > \mu_0$. The problem of identifying the MED is reformatted as a sequence of hypothesis testing problems:

$$H_{0j} : \mu_0 = \mu_1 = \cdots = \mu_j \quad vs \quad H_{1j} : \mu_0 \leq \mu_1 \leq \cdots \leq \mu_j.$$

If j^* is the smallest value for which H_{0j} is rejected, then the j^*th dose is identified to be the MED, that is $\widehat{\text{MED}} = j^*$. As previously mentioned, the $\widehat{\text{MED}}$ found is simply the lowest dose that differs significantly from the zero dose. In this sense, the hypothesis testing procedures do not really identify the MED; rather, we find the so-called MDD.

Suppose that H_{0j} is rejected for large values of the test statistic T_j, with critical value $c_{\alpha,j,\nu}$. Under a one-way model, the stepwise method to detect the MDD takes the following form (Tamhane, Hochberg and Dunnett 1996):

Step 1:

If $T_k > c_{\alpha,k,\nu}$,

then assert $\mu_k > \mu_0$ and go to Step 2;

else assert that there is no dose level which is significantly better than the zero dose level and stop.

Step 2:

If $T_{k-1} > c_{\alpha,k-1,\nu}$,

then assert $\mu_{k-1} > \mu_0$ and go to Step 3;

else assert MDD $= k$ and stop.

$\vdots$

Step k:

If $T_1 > c_{\alpha,1,\nu}$,

then assert $\mu_1 > \mu_0$ and go to Step $k+1$;

else assert MDD = 2 and stop.

Step $k+1$:

Assert that every dose level is significantly better than the zero dose level and stop.

Let step j $(1 \leq j \leq k+1)$ be the step at which the stepwise method stops. If $j > 1$, then the stepwise method declares dose $k-j+2, ..., k$ to be efficacious. If $j < k+1$, then the stepwise method fails to declare doses $1, ..., k-j+1$ to be efficacious.

We consider this stepwise test based on the following testing procedures:

(i) *DR Procedure* (Hsu and Berger 1999):

Let

$$T_j = (\bar{Y}_j - \bar{Y}_0)/\{s(n^{-1} + n_0^{-1})^{1/2}\} \tag{7.1}$$

which is the pairwise t test.

(ii) *Williams' Procedures*

Williams' (1971, 1972) procedure does not use the $\bar{Y}_i$'s as the estimates of the μ_i's; instead, it uses the isotonic estimates (see Section 2.1):

$$\mu_i^* = \max_{1\leq s\leq i} \min_{i\leq t\leq k} \sum_{j=s}^{t} Y_j/(t-s+1), i = 1, ..., k.$$

The test statistic proposed by Williams (1971) is

$$W_j^{(1)} = (\mu_j^* - \bar{Y}_0)/\{s(n_0^{-1} + n^{-1})^{1/2}\}. \tag{7.2}$$

Williams (1971, 1977) discussed another test statistic

$$W_j^{(2)} = (\hat{\mu_j}^* - \hat{\mu_0}^*)/\{s(n_0^{-1} + n^{-1})^{1/2}\}$$

to test H_{0j} against H_{1j} where

$$\hat{\mu_i}^* = \max_{0\le s\le i} \min_{i\le t\le k} \sum_{j=s}^{t} Y_j/(t-s+1), i = 0, 1, ..., k.$$

When σ is known, Marcus (1976) gave the exact upper 5% and 1% quantiles for $k = 2, ..., 5$ and estimated upper 5% and 1% quantiles for $k = 6, ..., 11$. Williams (1977) tabled the approximate critical values of $W_j^{(2)}$ for different degrees of freedom. The approximate critical values given by Williams (1977) will result in a slight decrease in the true size and power of the test. We will use the table values given by Marcus (1976) for the simulation study.

For the procedures studied below, we suppose that

$$T_j = (n_0 c_{0j} \bar{Y}_0 + \sum_{i=1}^{k} n c_{ij} \bar{Y}_i)/\{s^2(n_0 c_{0j}^2 + \sum_{i=1}^{k} n c_{ij}^2)^{1/2}\}.$$

(iii) *Linear Contrast Procedure* (Rom, Costello and Connell 1994)

The general form of the linear contrasts is

$$c_{ij} = \begin{cases} -j & i = 0; \\ c_{i-1,j} + 2 & i = 1, ..., j; \\ 0 & i = j+1, ..., k. \end{cases}$$

(iv) *Helmert Contrasts* (Ruberg 1989)

The jth Helmert contrast compares the jth dose response mean with the average of all the lower dose response means (including the zero dose) .

It is defined by

$$c_{ij} = \begin{cases} -1 & i = 0, 1, ..., j-1; \\ j & i = j; \\ 0 & i = j+1, ..., k. \end{cases}$$

(v) *Reverse Helmert Contrasts*

The jth reverse Helmert contrast compares the average of the first j dose response means with the zero dose response mean. It is defined by

$$c_{ij} = \begin{cases} -j & i = 0; \\ 1 & i = 1, ..., j; \\ 0 & i = j+1, ..., k. \end{cases}$$

(vi) *LRT for simple order alternative*

The likelihood ratio test

$$S_{01} = \frac{\{n_0(\hat{\mu_0}^* - \hat{\mu})^2 + \sum_{i=1}^{k} n(\hat{\mu_i}^* - \hat{\mu})^2\}/\sigma^2}{\{n_0(\bar{Y}_0 - \hat{\mu_0}^*)^2 + \sum_{i=1}^{k} n(\bar{Y}_i - \hat{\mu_i}^*)^2\}/(\nu\sigma^2) + Q(\nu)/\nu}$$

for testing the homogeneity of the response means against the simple order alternative (see Section 2.1) is considered, where $\hat{\mu} = \sum_{i=0}^{k} \bar{Y}_i/(k+1)$ and $Q(\nu) = \nu s^2/\sigma^2$. As S_{01} utilizes the monotonicity assumption of the response means, it is a more powerful test statistic for testing against the simple order alternative.

7.2 A Simulation Study

The simulation studies are conducted to compare the behavior of the stepwise method based on the LRT with DR method and the methods based on linear contrasts (denoted by LIN), Helmert contrasts, reverse Helmert contrasts and Williams' tests $W_j^{(1)}$ (denoted by WI) and $W_j^{(2)}$ (denoted by WII). Without loss of generality, a common sample size n is assumed for each dose level including the zero dose and μ_0 is fixed at 0. The number of the non-zero dose levels (k) is fixed at 5, the degrees of freedom (ν) is fixed at 6, the error rate α is 0.05 and $\sigma/\sqrt{n} = 1$ for all the simulations. The five logistic functions that have been studied in Section 6.3 are considered. For each case, 10,000 iterations were made.

The probability of detecting the difference between μ_j and μ_0 is the percentage that H_{0j} was rejected in a stepwise fashion as described in Section 7.1. The methods based on Helmert and Reverse Helmert contrasts have much lower probabilities to detect the difference between μ_j and μ_0 than the other test procedures for most of the cases studied. For example, in Case 2, the probability of successfully detecting the difference between μ_5 and μ_0 is only 52.2% by the method based on Helmert contrasts, and is 71.1% by the method based on reverse Helmert contrasts. However, the probabilities of successfully detecting the difference between μ_5 and μ_0 by the other five procedures are all above 80%. Hence, normally we will not use the test procedures based on Helmert and Reverse Helmert contrasts when the dose-response curve is approximately a logistic function. The simulation results excluding Helmert and reverse Helmert methods are given in Table 7.1. From Table 7.1, we can see that the LRT method, which has high probabilities of detecting μ_j and μ_0 for all the cases, compares favorably to the other methods. The maximum gains of the LRT method over the DR method, WI, WII and the method based on linear contrasts can reach 24.2%, 20.9%, 20.7% and 7.6%, respectively. For the aforementioned stepwise testing procedures, only Williams' tests and

LRT take the prior knowledge that $\mu_0 \leq \mu_1 \leq \cdots \leq \mu_k$ into account. Since Williams' methods have low probabilities of detecting the difference $\mu_j - \mu_0$, they are not recommended. The LRT is generally the best procedure which determines MDD for monotone dose-response curves without a high risk to make an incorrect decision.

Table 7.1: Probabilities (in percentage) of Detecting the Difference Between μ_j and μ_0 by Five Stepwise Procedures for $k = 5$, $\nu = 6$ and $\alpha = 0.05$

C	j	Method				
		DR	WI	WII	LIN	LRT
1.5	5	80.5	84.5	81.8	94.9	96.6
	4	70.3	78.4	75.0	91.3	93.6
	3	63.6	72.1	67.8	84.9	88.0
	2	52.6	56.9	51.1	63.0	69.2
	1	8.6	8.8	5.2	8.6	9.0
2.0	5	80.5	83.2	82.1	96.6	96.9
	4	69.6	75.3	73.4	91.9	92.6
	3	59.5	63.4	60.8	78.8	81.1
	2	31.8	31.5	27.0	35.5	41.8
	1	5.2	5.5	3.1	5.2	5.8
2.5	5	80.5	81.6	81.5	96.7	96.9
	4	67.9	70.4	69.4	89.1	89.8
	3	49.8	49.7	47.4	61.4	66.9
	2	14.3	14.0	11.3	15.6	19.5
	1	3.2	3.8	2.1	3.2	4.4
3.0	5	80.5	79.8	80.7	95.8	96.1
	4	63.7	63.0	63.2	81.2	83.9
	3	34.3	32.5	30.7	39.2	46.8
	2	7.8	7.8	6.1	8.7	11.7
	1	2.3	2.9	1.6	2.2	3.5
3.5	5	80.5	78.3	79.9	94.0	95.3
	4	56.7	53.9	54.7	68.8	74.3
	3	20.8	19.4	17.5	23.4	30.4
	2	4.9	5.6	4.3	5.6	8.7
	1	1.8	2.4	1.3	1.7	3.0

Chapter 8

Summary

The problem of identifying the differences among the monotone dose-response means is considered extensively in this thesis. If several response means are to compared with one another and the prior knowledge indicates that the response means are simply ordered, then better inference procedures can be chosen to detect the differences among the means. Our study focuses on the interpretation of the testing hypotheses, on the duality of simultaneous confidence lower bounds and on the constrained optimization problems. Interval estimation for the response mean differences has received much attention in our study. Four different approaches to construct efficient simultaneous bounds for linear contrasts of the response means are proposed.

The max-min multiple comparison procedure takes the advantage of Tukey's procedure, which is effective to give upper and lower bounds for pairwise mean differences. The extended OMCT procedure discussed in Section 5.4 may in general give shorter confidence intervals for pairwise comparisons $\mu_j - \mu_i, j > i$ than the max-min procedure when $j - i$ is large.

Marcus' results (1978) are significantly improved by giving a necessary and sufficient condition for the optimal solution and an easy computational algorithm to search for the improved lower bound for nonnegative contrasts. The approach is a good way to obtain bounds; however, its corresponding statistic S_{01} is not useful for testing the homogeneity against the simple order alternative. The OMCT approach is an intuitive, simple procedure to categorized the dosage levels. It is more efficient than OSRT as well as Marcus (1978) when the response means does not increase rapidly in one or more intervals of the dosage levels. This also suggests that if the differences among the means $\mu_i \leq \cdots \leq \mu_j$ are small, it is advantageous to use weighted average means $\sum_{j=\alpha}^{\beta} n_\alpha \bar{Y}_\alpha / \sum_{j=\alpha}^{\beta} n_\alpha$ in the inference procedures. The OMCT is not a good testing procedure in comparison to LRT S_{01}. However, the latter can only provide the lower bound for the pairwise difference $\mu_k - \mu_1$. While the OMCT can deal with any pairwise comparisons.

With the assumption of simple ordering of response means in dose-response studies, many analyses commence with an interest to discover the lowest dose (MED) of which the response mean is more efficacious than the control mean. We propose a more efficient test statistic, orthant test, by fully utilizing the prior knowledge $\mu_0 \leq \mu_1 \leq \cdots \leq \mu_k$ to test H_{0j} against H_{1j} simultaneously. The minimum effect dosage can be identified by simultaneous lower bounds for pairwise difference between $\mu_j - \mu_0$. This procedure could not give the bounds for general pairwise comparisons $\mu_j - \mu_i, i \neq 0$. Stepwise multiple testing procedure studied in Chapter 7 is another approach to identify this MED.

The most challenging part of this thesis is the study on the constrained confidence bound through deriving an efficient computational algorithm. It is a new field in order restricted statistical inference. The approach used in Chapter 4 and 6 can be applied to other constrained optimization problems.

Appendices: Fortran77 Programs

1. Program for Computing the Max-Min Simultaneous Confidence Intervals:

```
C ***********************************************************
C *   MAIN              MAXMIN.FOR           2001/01/15/      *
C *                                                           *
C *   Purpose :                                               *
C *              To compute the max-min simultaneous          *
C *              confidence  intervals                        *
C *                                                           *
C *   Variables:                                              *
C *                                                           *
C *   B      - Constant                                       *
C *   K      - Constant                                       *
C *            The number of the populations                  *
C *   Y      - One dimension array                            *
C *            Y(I) is the  sample mean of the Ith population *
C *   S      - One dimension array                            *
C *            S(I) is the sample size of the Ith population  *
C *   TEMP   - One dimension array                            *
C *            Store temporary data                           *
C *   INVS   - One dimension array                            *
C *            INVS(I) is the inverse of S(I)                 *
C *   VAR    - Pooled variance                                *
C *   CVQ    - The critical value of the studentized range    *
C *            test                                           *
C *   L      - Two dimension array                            *
C *          - L(I,J) is the max-min simultaneous confidence  *
C *            lower bound                                    *
C *   U      - Two dimension array                            *
C *          - U(I,J) is the max-min simultaneous confidence  *
C *            upper bound                                    *
C ***********************************************************

      INTEGER B
      PARAMETER (B = 20)
      INTEGER K
      REAL Y(B), TEMP(B), S(B), INVS(B), C(B,B), CL(B,B)
      REAL CU(B,B), N
      REAL U(B,B),  L(B,B), VAR, CVQ
```

```
      OPEN (UNIT=2, FILE='data.in' , STATUS='OLD')
      OPEN (UNIT=3, FILE='data.out', STATUS='UNKNOWN')

      READ (2,*) K
      DO 10  I = 1, K
         READ (2,*)  Y(I)
         TEMP(I) = Y(I)
10    CONTINUE
      READ (2,*) CVQ, VAR
      CLOSE(2)

      WRITE(3,*) ' Finish inputting observations'
      WRITE(3,*)  (Y(I) , I= 1,K)
      WRITE(3,*)   CVQ, VAR

      OPEN (UNIT = 4, FILE= 'size.in', STATUS = 'OLD')
      DO 15 I= 1, K
         READ(4,*)  S(I)
      INVS(I) = 1/S(I)
15    CONTINUE
      CLOSE(4)

      WRITE(3,*) 'Finishing inputting the sample size'
      WRITE(3,150) (S(I), I= 1, K)

      DO 20  I = 1, K
           DO 30   J = 1, K
             C(I,J) = Y(I) - TEMP(J)
             N = CVQ*VAR**0.5*(0.5*(INVS(I) + INVS(J)))**0.5
             CL(I,J) = C(I,J) - N
             CU(I,J) = C(I,J) + N
30         CONTINUE
20    CONTINUE

      Write(3,*)  ' Mean difference'
      Write(3,200) ((C(I,J), J=1, K), I= 1, K)

      DO 40  I= 1, K
        DO 50  J= 1, K
              L(I,J) = CL(I,J)
```

```
            U(I,J) = CU(I,J)

            DO 60    M= 1, I
              DO 70  N= J, K

                     IF (M. NE. N) THEN
                         IF ( CL(M,N) .GT.  L(I,J) )  THEN
                             L(I,J) = CL(M,N)
                         ENDIF
                     ELSE
                         GO TO 70
                     ENDIF

70            CONTINUE
60          CONTINUE

            DO 80   M= I, K
               DO 90  N= 1,J

                     IF (M. NE. N) THEN
                        IF ( CU(M,N) .LT. U(I,J) )  THEN
                             U(I,J) = CU(M,N)
                        ENDIF
                     ELSE
                        GO TO 90
                     ENDIF

90            CONTINUE
80          CONTINUE

50       CONTINUE
40      CONTINUE

        WRITE(3,*)  'calculation end'

        WRITE(3,*)  'Max-min Lower Bound is'
        WRITE(3,200)  ( (L(I,J), J=1, K), I=1,K)

        WRITE(3,*)  'Max-min Upper Bound is'
        WRITE(3,200)  ( (U(I,J), J=1, K), I=1 ,K)
```

```
150     FORMAT(5X, 1F6.2)
200     FORMAT(5X, 9F8.3)

        CLOSE(3)
        STOP
        END
```

2. Program for Simulating the OMCT Critical Values:

```
C ***********************************************************
C *   MAIN               CVL.f          1998/08/27/           *
C *                                                           *
C *   Purpose :                                               *
C *             To generate the critical value for the OMCT  *
C *             statistic                                     *
C *                                                           *
C *   Variables:                                              *
C *                                                           *
C *  ISEED - Seed of the intrinsic uniform random generator, *
C *          usually a very large integer                     *
C *  Q     - Generated sample variance                        *
C *  K     - Number of population levels                      *
C *  DF    - Degrees of freedom                               *
C *  CHI   - Generated Chi-square statistic                   *
C *  NIT   - Number of iteration                              *
C *  S     - One dimension array                              *
C *          Sample size of each level                        *
C *  A     - Two dimension array                              *
C *          A(I,J) is the mean of the observation from level*
C *          I to level J                                     *
C *  Z     - One diemnsion array                              *
C *          Generated standard normal radom variable         *
C *  C     - One dimension array                              *
C *          Tentative critical point                         *
C *  P     - One dimension array                              *
C *          Percentage of the OMCT statistic greater than    *
C *          the tentative critical point C                   *
C *  SN    - Two dimension array                              *
C *          SN(I,J) is the sum of the sample size from       *
C *          level I to J                                     *
C *  WS    - Two dimension array                              *
C *          WS(I,J) is the inverse of SN(I,J)                *
C *                                                           *
C *  Subroutines: NORM01, CHISQ                               *
C ***********************************************************

        INTEGER       DF, ISEED, K, NIT
        REAL          Q, S1, U, AV, SN(20,20), WS(20,20), F, T
        REAL          Z(20), A(20,20) , P(30), C(30), CHI, S(10)
```

```
      OPEN(5, FILE='ldf.dat', STATUS='OLD')
      OPEN(6, FILE='ldf.out', STATUS='UNKNOWN')

      READ(5,*) ISEED, NIT, K, DF
      WRITE(6, 130) ISEED, NIT, K, DF

      READ(5,*) (S(I), I=1,K)
      WRITE(6,140) (S(I), I=1,K)

      DO 10 I = 1, 30
         READ(5,*,END=200) C(I)
10    CONTINUE

200   CLOSE(5)

      DO 20  I = 1, 30
      P(I) = 0.
20    CONTINUE

      DO 110 IT = 1, NIT

      CALL NORM01(ISEED, K, Z)

      CALL CHISQ(ISEED, DF, CHI)

      Q = CHI/DF

      DO 50 I = 1, K
              AV= 0.
              J = I
              SN(I,I-1) = 0.
40            AV = AV + Z(J)*S(J)**0.5
              SN(I,J) = SN(I,J-1) + S(J)
              A(I,J) = AV/SN(I,J)

              IF (J .GE. K) GO TO 50
                     J = J+1
              GO TO 40

50    CONTINUE
```

```
      T = 0.
      DO 90 IP = 1, K-1
              DO 80 IQ = IP, K-1
              S1 = A(IP,IQ)
              WS(IP,IQ) = 1./SN(IP,IQ)

                      DO 70 IR = IQ+1, K
                              DO 60 IS = IR, K
                                 WS(IR,IS) = 1./SN(IR,IS)
                                 F = WS(IP,IQ) + WS(IR,IS)
                                 F = SQRT(F)
                                 U = (A(IR, IS) - S1)/F
                                 U = U/SQRT(Q)
                                 IF (U .GT. T)  THEN
                                     T = U
                                 END IF
60                            CONTINUE
70                    CONTINUE

80            CONTINUE
90    CONTINUE


      I = 0
100   I = I + 1

      IF ( T .LT. C(I))  GO TO 110
              P(I) = P(I) + 1
              IF (I .GE. 30)  GO TO 110
              GO TO 100

110   CONTINUE

      DO 120  I = 1, 30
              P(I) = P(I)/NIT
120   CONTINUE

130   FORMAT(4I10/)

      WRITE(6, 140) (C(I), I = 1, 10)
```

```
      WRITE(6, 150) (P(I), I = 1, 10)
      WRITE(6, 140) (C(I), I = 11, 20)
      WRITE(6, 150) (P(I), I = 11, 20)
      WRITE(6, 140) (C(I), I = 21, 30)
      WRITE(6, 150) (P(I), I = 21, 30)

140   FORMAT(10F8.3/)
150   FORMAT(10F8.5//)
      CLOSE(6)

      STOP
      END

C ******************************************************************
C * SUBROUTINE   NORMO1                                            *
C *                                                                *
C * Purpose  :                                                     *
C *       Generate a sample from a standard normal distribution.*
C *                                                                *
C * Variables:                                                     *
C *       ISEED  -  Seed of the intrinsic uniform random           *
C *                  generator, usually a very large integer       *
C *       N      -  Total sample size                              *
C *       DZ1    -  One dimensional array                          *
C *                 array size (N + 1)                             *
C *                                                                *
C * Subroutines:  none                                             *
C ******************************************************************
C

        SUBROUTINE     NORMO1(ISEED,K,Z)
        INTEGER        ISEED , K
        REAL           Z(20), U(20)
        REAL           WA, WB, WC, WPIE
C
C ******generate K+1 pseudo-ran numbers from U(0,1)
C
        WA = RAN(ISEED)
        DO 200 I = 1, K+1
                WA = RAN(ISEED)
                DO WHILE (WA .LE. 1.E-5 .OR. WA .GE. 1.-1.E-5)
```

```
                     WA = RAN(ISEED)
                  END DO
                  U(I) = WA
200         CONTINUE

C
C****** transform U(0,1) to standard normal (Box-muller)
C
         WPIE = ACOS(-1.)
         DO  300  I = 1, K, 2
             WA = SQRT(-2.*LOG(U(I)))
             WB = COS(2.*WPIE*U(I+1))
             WC = SIN(2.*WPIE*U(I+1))
             Z(I) = WA*WB
             Z(I+1) = WA*WC
300      CONTINUE

C
C****** END
C
        RETURN
        END

C *******************************************************************
C *  SUBROUTINE CHISQ            1998/07/20/                        *
C *                                                                 *
C *  Purpose: To generate the Chi-square random variables           *
C *                                                                 *
C *                                                                 *
C *******************************************************************


          SUBROUTINE      CHISQ(ISEED,  DF, CHI)
          INTEGER         ISEED , DF
          REAL            Z, CHI, U(300), WCHI(300)
          REAL            WD, WE, WPIE
          CHI = 0.

C
C ******generate K+1 pseudo-ran numbers from U(0,1)
C
```

```
          M = INT(DF/2)

          WD = RAN(ISEED)
          DO 1000 I = 1, M+2
                  WD = RAN(ISEED)
                  DO WHILE (WD .LE. 1.E-5 .OR. WD .GE. 1.-1.E-5)
                       WD = RAN(ISEED)
                  END DO
                  U(I) = WD
1000      CONTINUE

C
C******tranform U(0,1) to Chi-square with 2df
C
         DO  2000  I = 1, M
                 CHI = CHI  -2.*LOG(U(I))
2000     CONTINUE

         IF (MOD(DF,2)  .EQ. 1) THEN
             WPIE = ACOS(-1.)
             WD = SQRT(-2.*LOG(U(M+1)))
             WE = COS(2.*WPIE*U(M+2))
             Z  = WD*WE
             CHI = CHI + Z**2
         ENDIF

         RETURN
         END
```

3. Program for Computing the OMCT Simultaneous Lower Bounds:

```
C ***************************************************************
C *    Program   LBOMCT.f      1998/08/27                        *
C *                                                             *
C *    Purpose : Construct the  OMCT simultaneous lower bounds *
C *                                                             *
C *    Varables:                                                *
C *                                                             *
C *    B     - Constant                                         *
C *    STD   - Pooled standard  deviation                       *
C *    CVL   - The OMCT critical value                          *
C *    Y     - One dimension array                              *
C *            Y(I) is the  sample mean of the Ith population *
C *    S     - One dimension array                              *
C *            S(I) is the sample size of the Ith population  *
C *    YS    - Two dimension array                              *
C *          - YS(I,J) is the  sum of observations from Ith    *
C *            to Jth population                                *
C *    YB    - Two dimension array                              *
C *          - YB(I,J) is the  mean of the obsevation from     *
C *            Ith to Jth population                            *
C *    WS    - Two dimension array                              *
C *          - WS(I,J) is the sum of the sample size from      *
C *            Ith to Jth population                            *
C *    LL    - Two dimension array                              *
C *          - The OMCT lower bound                             *
C ***************************************************************

      INTEGER     B
      PARAMETER   ( B=20 )

      INTEGER     K
      REAL        Y(B), YS(0:B,B), YB(B,B), S(B)
      REAL        L(B,B),  LL(B,B), LLT(B,B), WS(0:B,B), WK(B)
      REAL        D, CVL, STD

      OPEN(2, FILE='par.dat',  STATUS='OLD')
      READ(2,*) K,  CVL, STD
      READ(2,*) (Y(I), I=1, K)
      READ(2,*) (S(I), I=1, K)
      CLOSE(2)
```

```
      OPEN(3, FILE='test.dat' , STATUS='UNKNOWN')
      WRITE(3,*)   K, CVL
      WRITE(3,*) (Y(I), I=1,K)
      WRITE(3,*) (S(I), I=1,K)


      DO 10 IP = 1, K
                  YS(IP,IP) = Y(IP)*S(IP)
                  WS(IP,IP) = S(IP)
                  YB(IP,IP) = Y(IP)
10      CONTINUE


       DO 12  IP = 1, K-1
             DO 13  IQ = IP+1 ,K
                   YS(IQ,IP) = YS(IQ-1,IP) + Y(IQ)*S(IQ)
                   WS(IQ,IP) = WS(IQ-1,IP) + S(IQ)
                   YB(IQ,IP) = YS(IQ,IP)/WS(IQ,IP)
13    CONTINUE
12    CONTINUE


      DO 20  IS= 2 , K
         DO 30  IP = 1, IS-1

            LL(IS, IP) = -1.0E5

             DO 40 IQ = IP, IS-1
                DO 50  IR = IQ+1, IS

                   D = STD*CVL*(1./WS(IQ, IP)+ 1./WS(IS, IR))**0.5
                   L(IR, IQ) = YB(IS, IR) - YB(IQ, IP) - D

                       IF (L(IR,IQ) .GT. LL(IS, IP)) THEN
                          LL(IS, IP) = L(IR,IQ)
                          LLT(IS,IP) = LL(IS,IP)
                       ENDIF

50                CONTINUE
40             CONTINUE
```

```
30        CONTINUE
20     CONTINUE


       DO 300  IS = 2, K
           DO 400  IP = 1, IS-1

              DO 500 IQ = IP, IS-1
                 DO 600  IR = IQ+1, IS

                        IF ( LLT(IR, IQ) .GT.  LL(IS, IP) )     THEN
                                LL(IS, IP) = LLT(IR, IQ)
                        ENDIF

600                 CONTINUE
500            CONTINUE

400       CONTINUE
300    CONTINUE


         WRITE(3,*)    'OMCT lower bound'
         WRITE(3,90)  ( (LL(I,J), J= 1, K), I= 1, K)

90       FORMAT(5X, 9F8.2)
         CLOSE(3)

      STOP
      END
```

References

Abelson, R. P and Tukey, J. W. (1963), "Efficient Utilization of Non-Numerical Information in Quantitative Analysis: General Theory and the Case of the Simple Order," *The Annals of Mathematical Statistics*, **34**, 1347-1369.

Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (1992), *A First Course in Order Statistics*. London: John Wiley.

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

Barlow, R. E. and Ubhaya, V. A. (1971), "Isotonic Approximation," in J. S. Rustagi (ed.), *Optimizing Methods in Statistics*, Academic Press, New York.

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 of Ordered Alternatives," *Biometrika*, 46, 36-48.

Bartholomew, D.J. (1959b), "A Test of Homogeneity of 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.

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.

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," *The Annals of Mathematical Statistics*, **43**, 1541-1548.

Chase, G. R. (1974), "On Testing for Ordered Alternatives with Increased Sample Size for a Control," *Biometrika*, **61**, 569-578.

Dunnett, C. W. (1955), "A Multiple Comparisons Procedure for Comparing Several Treatment with a Control," *Journal of the American Statistical Association*, **50**, 1096-1121.

Eden, C. Van (1958), "Testing and Estimating Ordered Parameters of Probability Distributions," Doctoral Dissertation, University of Amsterdam, Studentendrukkerij Poortpers, Amsterdam.

Genz, A.(1992), "Numerical Computation of Multivariate Normal Probabilities," *Journal of Computational and Graphical Statistics*, 1, 141-149.

Grove, D. M. (1980), "A Test of Independence Against a Class of Ordered Alternatives in a $2 \times C$ Contingency Table", *Journal of American Statistical Association*, **75**, 454-459.

Hayter, A. J. (1984), "A Proof of the Conjecture that the Tukey-Kramer Multiple Comparisons Procedure is Conservative," *The Annals of Statistics*, **12**, 61-75.

Hayter, A. J. (1990), " A One-Sided Studentized Range Test for Testing Against a Simple Ordered Alternative," *Journal of the American Statistical Association*, **85**, 778-785.

Hayter, A. J. (1992), "Multiple Comparisons of Three Ordered Normal Means for Unbalanced Models," *Computational Statistics and Data Analysis*, 13, 153-162.

Hirotsu , C., Kuriki, S., and Hayter, A. J. (1992), "Multiple Comparison Procedures Based On the Maximal Component of Cumulative Chi-Squared Statistic," *Biometrika*, 79, 381-392.

Hochberg, Y. and Tamhane, A. C. (1987), *Multiple Comparison Procedures*. Wiley, New York.

Hogg, R. V. (1965), "On Models and Hypotheses with Restricted Alternatives," *Journal of the American Statistical Association*, **60**, 1153-1162.

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.

Kanowith-Klein, S., Vitetta, E. S. , Korn, E. L. and Ashman R. F. (1979), "Antigen-induced Changes in the Proportion of Antigen-binding Cells Expressing IgM, IgG, and IgD receptors," *Journal of Immunology*, **122**, 2349-2355.

Kelly, R. E. (1989), "Stochastic Reduction of Loss in Estimating Normal Means by Isotonic Regression," *The Annals of Statistics*, **17**, 937-940.

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.

Lee, C. C. (1981), "The Quadratic Loss of Isotonic Regression under Normality," *The Annals of Statistics*, **9**, 686-668.

Lee, C. C. (1988), "Quadratic Loss of Order Restricted Estimators for Treatment Means with a Control," *The Annals of Statistics*, **16**, 751-758.

Lee, C. C., Robertson, T. and Wright, F. T. (1993), "Bounds on Distributions Arising in Order Restricted Inferences with Restricted Weights," *Biometrika*, **80**, 405-416.

Lee, C. C. (1996), "On Estimation for Monotone Dose-Response Curves", *Journal of the American Statistical Association*, **91**, 1110-1119.

Lee, R. E. and Spurrier, J. (1995), "Successive Comparisons between ordered Treatments," *Journal of Statistical Planning and Inference*, **43**, 323-330.

Liu, W., Miwa, T. and Hayter, A. J. (2000), "Simultaneous Confidence Interval Estimation for Successive Comparisons of Ordered Treatment Effects," *Journal of Statistical Planning and Inference*, **88**, 75-86.

Marcus, R. (1976), "The Powers of Some Tests of the Equality of Normal Means Against an Ordered Alternative," *Biometrika*, **63**, 177-183.

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. (1978), " Further Results on Simultaneous Confidence Bounds in Normal Models With Restricted Alternatives," *Communications in Statistics, Part A - Theory and Methods*, **A7**, 573-590.

Marcus, R. (1982), "Some Results on Simultaneous Confidence Intervals for Monotone Contrasts in One-way ANOVA model," *Communications in Statistics, Part A - Theory and Methods*, **11**, 615-622.

Miller, R. G., Jr. (1981), *Simultaneous Statistical Inference*, 2nd edn. New York: Springer-Verlag.

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.

McDermott, M. P. (1999), "Generalized Orthogonal Contrast Tests for homogeneity of Ordered Means," *The Canadian Journal of Statistics*, **27**, 457-470.

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. and Wright, F. T. (1974), "A Norm Reducing Property for Isotonized Cauchy Mean Values Functions," *The Annals of Statistics,* **2**, 1302-1307.

Robertson, T. and Wright, F. T. (1982), "Bounds on Mixtures of Distributions Arising in Order Restricted Inference," *The Annals of statistics,* **10**, 302-306.

Robertson, T. and Wright, F. T. (1983), "On Approximation of the level Probabilities and Associated Distributions in Order Restricted Inference," *Biometrika*, **70**, 597-606.

Robertson, T., Wright, F. T. and Dykstra, R. L. (1988) *Order Restricted Statistical Inference.* Chichester, U.K.: John Wiley.

Rockafellar, R. T. (1972) *Convex Analysis.* Princeton, New Jersey.

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), "Contrasts for Identifying the Minimum 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**(1), 15-42.

Schaafsma, W. and Smid, L. J. (1966), "Most Stringent Somewhere Most Powerful Test Against Alternatives Restricted by a Number of Linear Inequalities," *The Annals of Mathematical Statistics*, **37**, 1161-1172.

Scheffé, H. (1959), "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 Experiments, and Low-Dose Extrapolation," *Journal of the American Statistical Association*, **81**, 186-195.

Siskind (1976), "Approximate Probability Integrals and Critical Values for Bartholomew's Test for Ordered Means," *Biometrika*, **63**, 647-654.

Snidjers, T. (1979), *Asymptotic Optimality Theory for Testing Problems with Restricted Alternatives*, Mathmematisch Centrum, Amsterdam.

Tamhane, A.C., Hochberg, Y., and Dunnett, C.W. (1996), "Multiple Testing Procedures for Dose Finding," *Biometrics*, **52**, 21-37.

Tang, D. and Lin, S. P. (1997), "An Approximate Likelihood Ratio Test for Comparing Several Treatments to a Control", *Journal of the American Statistical Association.* **92** 1155-1162.

Thompson, W. A. Jr. (1962), "The Problem of Negative Estimates of Variance Components", *The Annals of Mathematical Statistics*, **33**, 273-289.

Tsu, J. C. (1996), *Multiple Comparisons: Theory and Methods*, London: Chapman and Hall.

Williams, D. A. (1971), "A Test for Difference Between Treatment means When Several Dose Levels are Compared with a Zero Dose Control," *Biometrics*, **27**, 103-117.

Williams, D. A. (1972), "The Comparison of Several Dose Levels with a Zero Dose Control," *Biometrics*, **28**, 519-531.

Williams, D. A. (1977), "Some Inference Procedures for Monotonically Ordered Normal Means," *Biometrika*, **64**, 9-14.

Wright, F. T. (1982), "Monotone Regression Estimates for Grouped Observations," *The Annals of Statistics*, **10**, 278-286.

Wright, F. T. and Tran, T. (1985), " Approximating the Level Probabilities in Order Restricted Inference: The Simple Tree Ordering," *Biometrika*, **72**, 429-439.

Wright, F. T. (1988), "The One-way Analysis of Variance with Ordered Alternatives: A Modification of Bartholomew's $\bar{E}^2$ Test", *The Canadian Journal of Statistics*, **16**, 75-85.