Recent Developments in Multiple Comparison Procedures
Institute of Mathematical Statistics
Lecture Notes – Monograph Series
Vol. 47 (2004) 119–128

Multiple comparisons with more than one control

John D. Spurrier[1] and Eleanne Solorzano[2]

University of South Carolina and University of New Hampshire

Abstract: This article provides a survey of recent work on comparing more than one experimental treatment to more than one control based on differences in the location parameters. The primary focus is on simultaneous confidence intervals for the amount of improvement that the experimental treatments offer relative to each control. Some discussion of single-step and step-down tests is also included. Both the parametric and nonparametric cases are considered. Data from a medical example are used to illustrate the techniques.

1. Introduction

Dunnett (1955) proposed the problem of comparing the means of k_t experimental treatments to the mean of a control or standard treatment in the one-way layout under the assumption of i.i.d. normal errors. Numerous papers on comparing experimental treatments to one control followed Dunnett's seminal paper.

In many settings, one wishes to compare the k_t experimental treatments to $k_c > 1$ controls or standard treatments. Hoover (1991) discusses the example of comparing experimental pain relievers to $k_c = 2$ standard treatments, aspirin and acetaminophen. Blake and Boockfor (1997) studied the effects of four compounds injected into rats via a corn oil vehicle. They used $k_c = 2$ controls, an active control in which rats received an injection of only corn oil and a passive control in which rats received no injection. Shaffer (1977) discusses comparing each mean in one group of treatments to each mean in another group of treatments. The second group could be a set of controls or standard treatments but does not have to be. Finally, Hilden (2000) suggested comparing an experimental treatment to a highly effective, but not practical, "ideal" treatment and to a control. His interest is in what fraction of the improvement offered by the "ideal" treatment relative to the control is obtained by the experimental treatment.

This article provides a survey of the literature on comparisons with more than one control. Although the focus will primarily be on $100(1–\alpha)\%$ simultaneous confidence bounds, both single-step and step-down testing procedures will be discussed. We will index the k_t experimental treatments by $i = 1, \ldots, k_t$ and the k_c controls or standard treatments by $j = 1, \ldots, k_c$. Section 2 considers procedures for the one-way layout under the assumption of i.i.d. normal errors. Section 3 deals with a nonparametric analysis of the one-way layout under the assumption of i.i.d. continuous errors. Section 4 discusses a class of incomplete block designs for comparing more than one experimental treatment with more than one control. In Section 5,

[1]Department of Statistics, University of South Carolina, Columbia, SC 29208, USA. e-mail: spurrier@stat.sc.edu

[2]Department of Decision Sciences, University of New Hampshire, Durham, NH 03824, USA. e-mail: eleannes@cisunix.unh.edu

Keywords and phrases: fractional improvement, incomplete block, nonparametric, one-way layout, simultaneous confidence intervals, simultaneous tests.

AMS 2000 subject classifications: primary 62F25, 62G15; secondary 62F03, 62G10.

Hilden's (2000) fractional improvement problem is discussed. An example analyzing some of Blake and Boockfor's (1997) data is given in Section 6.

2. One-way layout, normal errors

Consider a one-way layout with i.i.d. $N(0, \sigma^2)$ errors having n_t observations from each experimental treatment and n_c observations from each control (or standard treatment). Denote the expected response and the sample mean of experimental treatment i by ξ_i and $\overline{X}_i$, $i = 1, \ldots, k_t$. Denote the expected response and the sample mean of control j by η_j and $\overline{Y}_j$, $j = 1, \ldots, k_c$. Let S^2 denote the pooled estimator of σ^2. This estimator has $\nu = k_t(n_t - 1) + k_c(n_c - 1)$ degrees of freedom.

2.1. Simultaneous intervals

Shaffer (1977) constructs $100(1 - \alpha)\%$ simultaneous two-sided confidence bounds for all $\xi_i - \eta_j$ by using the pivotal variables

$$T_{ij} = \frac{(\overline{X}_i - \overline{Y}_j) - (\xi_i - \eta_j)}{S(1/n_t + 1/n_c)^{1/2}}, \quad i = 1, \ldots, k_t;\ j = 1, \ldots, k_c. \tag{1}$$

The simultaneous bounds

$$\xi_i - \eta_j \in (\overline{X}_i - \overline{Y}_j) \pm a_2 S(1/n_t + 1/n_c)^{1/2}, \quad i = 1, \ldots, k_t;\ j = 1, \ldots, k_c \tag{2}$$

follow from inverting the probability statement

$$P(|T_{ij}| \le a_2, \quad i = 1, \ldots, k_t;\ j = 1, \ldots, k_c) = 1 - \alpha. \tag{3}$$

Hoover (1991) uses analogous methods to get the one-sided simultaneous lower bounds

$$\xi_i - \eta_j \ge (\overline{X}_i - \overline{Y}_j) - a_1 S(1/n_t + 1/n_c)^{1/2}, \quad i = 1, \ldots, k_t;\ j = 1, \ldots, k_c \tag{4}$$

by inverting the probability statement

$$P(T_{ij} \le a_1, \quad i = 1, \ldots, k_t;\ j = 1, \ldots, k_c) = 1 - \alpha. \tag{5}$$

We must evaluate a_1 and a_2 to use these methods.

Following Hoover (1991), we can write

$$T_{ij} = \frac{\left(\frac{n_c}{n_c + n_t}\right)^{1/2} Z_{ti} - \left(\frac{n_t}{n_c + n_t}\right)^{1/2} Z_{cj}}{S/\sigma}, \tag{6}$$

where $Z_{ti} = \frac{\overline{X}_i - \xi_i}{\sigma/(n_t)^{1/2}}$ and $Z_{cj} = \frac{\overline{Y}_j - \eta_j}{\sigma/(n_c)^{1/2}}$ for $i = 1, \ldots, k_t$; $j = 1, \ldots, k_c$. Let $Z_{c(1)}$ and $Z_{c(k_c)}$ denote the smallest and largest order statistics of the Z_{cj}'s, respectively. Now, equation (5) is equivalent to

$$P\left[\left(\frac{n_c}{n_c + n_t}\right)^{1/2} Z_{ti} \le \left(\frac{n_t}{n_c + n_t}\right)^{1/2} Z_{c(1)} + \frac{a_1 S}{\sigma},\ i = 1, \ldots, k_t\right]$$
$$= \int_0^\infty \int_{-\infty}^\infty \left\{\Phi\left[\left(\frac{n_t}{n_c}\right)^{1/2} z + \left(\frac{n_c + n_t}{n_c}\right)^{1/2} a_1 u\right]\right\}^{k_t} f(z) g(u)\, dz\, du = 1 - \alpha, \tag{7}$$

where $f(z)$ and $g(u)$ are the density functions of $Z_{c(1)}$ and S/σ, respectively. This double integral can be evaluated numerically, and a root finding technique can be used to solve equation (7) for a_1. Equation (3) is equivalent to

$$P\left[\left(\frac{n_t}{n_c+n_t}\right)^{1/2} Z_{c(k_c)} - \frac{a_2 S}{\sigma} \le \left(\frac{n_c}{n_c+n_t}\right)^{1/2} Z_{ti} \le \left(\frac{n_t}{n_c+n_t}\right)^{1/2} Z_{c(1)} + \frac{a_2 S}{\sigma},\ i=1,\ldots,k_t\right] \tag{8}$$

which can be rewritten involving a triple integral by conditioning on the values of $Z_{c(1)}$, $Z_{c(k_c)}$, and S/σ. Again, numerical methods can be used to solve equation (3) for a_2. Hoover (1991) and Solorzano and Spurrier (1999) provide tables of the probability points.

Hoover (1991) shows one minimizes the standard errors of the treatment versus control contrast estimators for a fixed total sample size by taking

$$n_c \approx (k_t/k_c)^{1/2} n_t. \tag{9}$$

Thus, one should only use equal sample sizes only if $k_c = k_t$.

2.2. Simultaneous hypothesis tests

Let us now turn our attention to simultaneously testing the null hypotheses

$$\begin{aligned} &H_0\text{: } \xi_i = \eta_j \quad \text{for } i=1,\ldots,k_t;\ j=1,\ldots,k_c \text{ versus} \\ &H_a\text{: } \xi_i \ne \eta_j \text{ for at least one } (i,j) \end{aligned} \tag{10}$$

with an experimentwise type I error rate of α. The simultaneous tests are based on the statistics

$$T_{ij}^* = \frac{(\overline{X}_i - \overline{Y}_j)}{S(1/n_t + 1/n_c)^{1/2}}, \quad i=1,\ldots,k_t;\ j=1,\ldots,k_c. \tag{11}$$

The single step testing procedure declares $\xi_i \ne \eta_j$ if $|T_{ij}^*| \ge a_2$.

For the step-down testing procedure, one orders the T_{ij}^* statistics from largest to smallest with respect to absolute value. Let $|T_q^*|$ denote the qth largest $|T_{ij}^*|$, $q=1,\ldots,k_tk_c$. Let $A_1 = \{(i,j),\ i=1,\ldots,k_t;\ j=1,\ldots,k_c\}$. Let $a_{21} = a_2$.

In the first step of the step-down procedure, we stop if $|T_1^*| < a_{21}$. Otherwise, we declare $\xi_i \ne \eta_j$ for the (i,j) pair corresponding to the statistic $|T_1^*|$ and continue to step 2.

In the qth step for $q=2,\ldots,k_tk_c$, we let A_q be the set of (i,j) pairs in A_{q-1} minus the (i,j) pair corresponding to the statistic $|T_{q-1}^*|$. Let a_{2q} be the probability point such that

$$P\left[\left|T_{ij}^*\right| \le a_{2q} \quad \text{for all } (i,j) \in A_q \mid H_0\right] = 1-\alpha. \tag{12}$$

We stop if $|T_q^*| < a_{2q}$. Otherwise, we declare $\xi_i \ne \eta_j$ for the (i,j) pair corresponding to the statistic $|T_q^*|$, and continue to step $q+1$.

The value of a_{2q}, which depends on which (i,j) pairs are in A_q, is the solution of

$$\begin{aligned} 1-\alpha &= P\left[\left|T_{ij}^*\right| \le a_{2q} \quad \text{for all } (i,j) \in A_q \mid H_0\right] \\ &= P\left[-a_{2q} \le \frac{(1/2)^{1/2}(Z_{ti} - Z_{cj})}{S/\sigma} \le a_{2q} \quad \text{for all } (i,j) \in A_q \mid H_0\right] \\ &= P\left[Z_{cj} - 2^{1/2} a_{2q} S/\sigma \le Z_{ti} \le Z_{cj} + 2^{1/2} a_{2q} S/\sigma \quad \text{for all } (i,j) \in A_q \mid H_0\right] \end{aligned} \tag{13}$$

Note that a_{2q} is decreasing in q.

For the case of $k_c = 2$, let b_{1q} denote the number of i's such that $(i,1)$ and $(i,2)$ are both in A_q. Let b_{2q} denote the number of i's such that $(i,1)$ is in A_q but $(i,2)$ is not. Let b_{3q} denote the number of i's such that $(i,2)$ is in A_q but $(i,1)$ is not. If $b_{2q} = b_{3q} = 0$, the value of a_{2q} can be found in Hoover's (1991) Table 2. Otherwise, one may need to find the value.

Conditioning on $Z_{c1} = z_1, Z_{c2} = z_2$, and $S/\sigma = u$, equation (13) becomes

$$\begin{aligned} 1-\alpha \;=\; & \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty I\big[\max(z_1,z_2) - 2^{1/2}a_{2q}u < \min(z_1,z_2) + 2^{1/2}a_{2q}u\big] \\ & \times \big\{\Phi\big[\min(z_1,z_2) + 2^{1/2}a_{2q}u\big] - \Phi\big[\max(z_1,z_2) - 2^{1/2}a_{2q}u\big]\big\}^{b_{1q}} \\ & \times \big\{\Phi\big[z_1 + 2^{1/2}a_{2q}u\big] - \Phi\big[z_1 - 2^{1/2}a_{2q}u\big]\big\}^{b_{2q}} \\ & \times \big\{\Phi\big[z_2 + 2^{1/2}a_{2q}u\big] - \Phi\big[z_2 - 2^{1/2}a_{2q}u\big]\big\}^{b_{3q}} \phi(z_1)\phi(z_2)g(u)\,dz_1\,dz_2\,du \end{aligned} \tag{14}$$

provided that $b_{1q} > 0$. If $b_{1q} = 0$, the first two factors are removed from the integrand. The integral in (14) can be evaluated numerically and an iterative method can be used to solve for a_{2q}.

The step-down testing procedures for one-sided alternatives are analogous.

3. One-way layout, nonparametric

Solorzano and Spurrier (2001a) investigate nonparametric comparisons with more than one control under the setting of a one-way layout with i.i.d. continuous random errors. In this setting, ξ_i and η_j are medians rather than means and two-sample rank statistics are used in placed of the pivotals in (1). Under this model, subtracting ξ_i from each treatment i observation and subtracting η_j from each control j observation produces i.i.d. random variables.

Let U_{ij} be the Mann–Whitney (1947) statistic for comparing experimental treatment i to control j. That is, U_{ij} is the number of times that an observation from experimental treatment i is less than an observation from control j. Small values of U_{ij} suggest that $\xi_i > \eta_j$. To form simultaneous confidence intervals for all $\xi_i - \eta_j$, we must consider the joint behavior of the U_{ij}'s under the null hypothesis in (10).

We will consider two types of rankings of observations. A separate ranking will refer to the ranking of the $n_t + n_c$ observations from experimental treatment i and control j. If $\xi_i = \eta_j$, then all separate rankings are equally likely. We can represent a separate ranking by the n_c-tuple with mth element equal to the number of experimental treatment i observations smaller than mth order statistic from control j. The n_c-tuple is known as a partition. For example, the separate ranking C T C C T C T T yields the partition $(0,1,1,2)$. There is a 1-1 correspondence between partitions and separate rankings. The statistic U_{ij} equals the sum of the elements of the partition for experimental treatment i and control j.

A joint ranking will refer to the ranking of observations from all experimental treatments and controls under consideration at a given time. There are $\frac{(n_t+2n_c)!}{n_c!n_c!n_t!}$ joint rankings of the observations from one experimental treatment and two controls. All joint rankings are equally likely under H_0.

To form one-sided confidence bounds for all $\xi_i - \eta_j$, we must find the largest integer a such that

$$P(\min U_{ij} \le a \mid \mathrm{H}_0) \le \alpha. \tag{15}$$

Let A_{ij} be the event $\{U_{ij} \le a\}$. We can write the event $\{\min U_{ij} \le a\}$ in terms of unions and intersections of the A_{ij}'s. For $k_t = k_c = 2$,

$$\begin{aligned}
&P(\min U_{ij} \le a \mid \mathrm{H}_0) \\
&\quad = P(A_{11}) + P(A_{12}) + P(A_{21}) + P(A_{22}) - P(A_{11} \cap A_{12}) - P(A_{11} \cap A_{21}) \\
&\qquad - P(A_{11} \cap A_{22}) - P(A_{12} \cap A_{21}) - P(A_{12} \cap A_{22}) - P(A_{21} \cap A_{22}) \\
&\qquad + P(A_{11} \cap A_{12} \cap A_{21}) + P(A_{11} \cap A_{12} \cap A_{22}) + P(A_{11} \cap A_{21} \cap A_{22}) \\
&\qquad + P(A_{12} \cap A_{21} \cap A_{22}) - P(A_{11} \cap A_{12} \cap A_{21} \cap A_{22}).
\end{aligned}$$

Using equal in distribution arguments, we get

$$\begin{aligned}
&P(\min U_{ij} \le a \mid \mathrm{H}_0) \\
&\quad = 4P(A_{11}) - 2P(A_{11} \cap A_{12}) - 2P(A_{11} \cap A_{21}) - 2\big[P(A_{11})\big]^2 \\
&\qquad + 4P(A_{11} \cap A_{12} \cap A_{21}) - P(A_{11} \cap A_{12} \cap A_{21} \cap A_{22}).
\end{aligned} \tag{16}$$

Let Λ be the set of partitions whose elements sum to $\le a$. Under H_0,

$$P(A_{11}) = \frac{\text{number of elements in } \Lambda}{\binom{n_t+n_c}{n_c}}, \tag{17}$$

and

$$P(A_{11} \cap A_{12}) = \frac{\text{number of joint rankings with } U_{11} \le a \text{ and } U_{12} \le a}{\frac{(n_t+2n_c)!}{n_c!n_c!n_t!}}. \tag{18}$$

To calculate the number of joint rankings in equation (18), one can sum the number of joint rankings consistent with each element of $\Lambda \times \Lambda$. Similar methods are used for computing the other terms in equation (16). Solorzano and Spurrier (2001a) give algorithms for computing the number of such joint rankings. They also provide tables of probability points, some probability inequalities, and asymptotic results for the broader class of Chernoff-Savage (1958) linear rank statistics.

For two-sided confidence bounds, we must find the largest integer a such that

$$P\big(\min U_{ij} \le a \text{ or } \max U_{ij} \ge n_t n_c - a \mid \mathrm{H}_0\big) \le \alpha. \tag{19}$$

The computational techniques are similar to those used in the one-sided case.

One forms simultaneous distribution-free confidence bounds for $\xi_i - \eta_j$ by inverting the two-sample Mann–Whitney test. The probability point from inequality (15) or (19) is used in place of the two-sample probability point. The optimal design for comparing k_t experimental treatments with k_c controls is identical to that in Section 2.

The single-step and step-down simultaneous testing procedures are analogous to those in Section 2 with U_{ij} replacing T^*_{ij}.

4. Incomplete blocks

Solorzano and Spurrier (2001b) investigate the comparison of k_t experimental treatments with k_c controls with incomplete block designs with b blocks of size $p < k_t + k_c$.

In this setting, it is natural to seek balance among the experimental treatments and balance among the controls. Thus, one desires an incomplete block design with

$$\begin{aligned}
&\mathrm{Var}\big(\hat{\xi}_i - \hat{\eta}_j\big) \text{ equal for all } i \text{ and } j,\\
&\mathrm{Cov}\big(\hat{\xi}_i - \hat{\eta}_j, \hat{\xi}_{i'} - \hat{\eta}_j\big) \text{ equal for all } ti \neq i' \text{ and } j,\\
&\mathrm{Cov}\big(\hat{\xi}_i - \hat{\eta}_j, \hat{\xi}_i - \hat{\eta}_{j'}\big) \text{ equal for all } i \text{ and } j \neq j',\\
&\mathrm{Cov}\big(\hat{\xi}_i - \hat{\eta}_j, \hat{\xi}_{i'} - \hat{\eta}_{j'}\big) \text{ equal for all } i \neq i' \text{ and } j \neq j'.
\end{aligned} \tag{20}$$

A class of incomplete block designs satisfying condition (20) has been studied by Majumdar (1986) and by Jaggi, Gupta, and Parsad (1996). This class is known as balanced bipartite block (BBPB) designs. If $k_c = 1$, the class becomes Bechhofer and Tamhane's (1981) class of balanced treatment incomplete block designs. Let n_{tih} denote the number of times experimental treatment i appears in block h, and n_{cjh} denote the number of times control j appears in block h. A design is BBPB if there exists integers λ_1, λ_2, and λ_3 such that

$$\begin{aligned}
&\sum_{h=1}^{b} n_{tih} n_{ti'h} = \lambda_1 \quad \text{for all } i \neq i',\\
&\sum_{h=1}^{b} n_{cjh} n_{cj'h} = \lambda_2 \quad \text{for all } j \neq j',\\
&\sum_{h=1}^{b} n_{tih} n_{cjh} = \lambda_3 > 0 \quad \text{for all } i \neq j.
\end{aligned} \tag{21}$$

An example of a BBPB design with $k_t = 3$ experimental treatments (treatments 1, 2, and 3) and $k_c = 2$ controls (treatments 4 and 5) using $b = 9$ blocks of size $p = 3$ is given in Table 1. This design has $\lambda_1 = 3, \lambda_2 = 4$, and $\lambda_3 = 2$. Solorzano (1999) shows that BBPB designs are the only designs satisfying conditions (20).

Solorzano and Spurrier (2001b) give some A-optimality results for BBPB designs. For example, when comparing two experimental treatments to two controls using a large number of blocks of size 2, approximately 82.8% of the blocks should contain one treatment and one control, 8.6% of the blocks should contain both experimental treatments, and 8.6% of the blocks should contain both controls. When

Table 1: An example BBPB design with 3 experimental treatments (1, 2, and 3) and 2 controls (4 and 5) in 9 blocks of size 3

Block	Treatments
1	1 2 3
2	4 4 5
3	4 5 5
4	1 2 4
5	1 3 4
6	1 2 5
7	1 3 5
8	2 3 4
9	2 3 5

comparing two experimental treatments to two controls with blocks of size 3, the A-optimal BBPB design is a balanced incomplete block design if b is divisible by 4.

Solorzano and Spurrier (2001b) also give algorithms for simultaneous confidence bounds for all $\xi_i - \eta_j$ under i.i.d. normal errors. The probability points for these bounds differ from those in the one-way layout.

5. Fractional improvement

Hilden (2000) suggested a variation of the comparison with control problem in which one wishes to compare an experimental treatment having mean ξ with $k_c = 2$ other treatments, a control having mean η_1 and a highly effective "ideal" treatment having mean η_2. It is assumed that the "ideal" treatment is not practical due to high costs or other constraints. For this problem, one could reverse the roles of the experimental treatment and the controls and get simultaneous confidence intervals for $\xi - \eta_1$ and $\xi - \eta_2$ using Dunnett's (1955, 1964) results. Hilden (2000) was interested in the fraction of the improvement of the "ideal" treatment to the control that is achieved by the experimental treatment. Denote this fractional improvement by

$$\gamma = \frac{\xi - \eta_1}{\eta_2 - \eta_1} \tag{22}$$

It is assumed that $\eta_2 \neq \eta_1$.

Zerbe (1978) presented a confidence region for the ratio of two linear combinations of general linear model parameters. His results make use of Fieller's (1944) theorem. The fractional improvement parameter is an important special case of Zerbe's ratio.

The $100(1-\alpha)\%$ confidence region for γ is found by inverting the size α hypothesis test of

$$\begin{aligned} &\mathrm{H}_0\colon \gamma = \gamma_0 \leftrightarrow (\xi - \eta_1) - \gamma_0(\eta_2 - \eta_1) = 0 \text{ versus} \\ &H_a\colon \gamma \neq \gamma_0 \leftrightarrow (\xi - \eta_1) - \gamma_0(\eta_2 - \eta_1) \neq 0, \end{aligned} \tag{23}$$

where γ_0 is an arbitrary constant. These hypotheses can be tested using the t-test for testing that a linear combination of general linear model parameters equals zero. Alternatively, one might be interested in testing that γ equals a specific constant.

6. A normal theory example with two-sided inference

Blake and Boockfor (1997) studied the effects of $k_t = 4$ compounds injected into rats via a corn oil vehicle. They used $k_c = 2$ controls, an active control in which rats received an injection of only corn oil and a passive control in which rats received no injection. They used a one-way layout with $n_t = n_c = 6$ rats. Part of their data involves the weights of the rats' pituitary glands one month after treatment. This data are summarized in Table 2.

For illustration purposes, we will assume i.i.d. normal errors and make two-sided inferences with $\alpha = 0.05$. We have $\nu = 4(6-1) + 2(6-1) = 30$ degrees of freedom. From Hoover's (1991) Table 2, we have $a_2 = 2.848$. The two-sided simultaneous 95% confidence intervals are

$$\xi_i - \eta_j \in (\overline{X}_i - \overline{Y}_j) \pm 2.848\big[1.0626(1/6 + 1/6)\big]^{1/2} \tag{24}$$

which simplifies to $(\overline{X}_i - \overline{Y}_j) \pm 1.69$. The 95% simultaneous confidence bounds are given in Table 3. The $\pm$ terms for the Bonferroni and Tukey (1953) methods are ± 1.75 and ± 1.81, respectively.

Table 2: Descriptive statistics for Blake and Boockfor's (1997) weights (mg) of pituitary glands for 4 experimental treatments and two controls

Treatment	Mean weight	Mean square error
Compound 1 ($i = 1$)	6.96	1.0626
Compound 2 ($i = 2$)	7.83	
Compound 3 ($i = 3$)	9.02	
Compound 4 ($i = 4$)	14.12	
Passive control ($j = 1$)	7.19	
Active control ($j = 2$)	6.70	

Table 3: 95% simultaneous two-sided confidence intervals for difference in Blake and Boockfor's (1997) mean weight (mg) of pituitary glands, compound – control

Contrast	Confidence interval
Compound 1 – Passive Control	$(-1.92, 1.46)$
Compound 2 – Passive Control	$(-1.05, 2.33)$
Compound 3 – Passive Control	$(0.14, 3.52)$
Compound 4 – Passive Control	$(5.24, 8.62)$
Compound 1 – Active Control	$(-1.43, 1.95)$
Compound 2 – Active Control	$(-0.56, 2.82)$
Compound 3 – Active Control	$(0.63, 4.01)$
Compound 4 – Active Control	$(5.73, 9.11)$

The single-step and step-down simultaneous tests are based on the T_{ij}^* and $|T_q^*|$ statistics, respectively. The value of these statistics are given in Table 4. The single-step test declares $\xi_i \neq \eta_j$ if $|T_{ij}^*| \geq a_2 = 2.848$. From Table 4, we see that the means for compounds 3 and 4 are significantly different from the means for both controls. This can also be seen from Table 3, where the corresponding confidence intervals do not contain zero.

To do the more powerful step-down test with this data, we begin by comparing $|T_1^*| = 12.468$ to $a_{21} = a_2 = 2.848$. As $12.468 \geq 2.848$, we conclude that the mean for compound 4 ($i = 4$) differs from the mean for the active control ($j = 2$) and move to step 2.

For step 2, the set

$$A_2 = \big\{(1,1),(1,2),(2,1),(2,2),(3,1),(3,2),(4,1)\big\}. \tag{25}$$

As $|T_2^*| = 11.645 \geq a_{21}$ which is larger than a_{22}, we can conclude that the mean for compound 4 ($i = 4$) is different from the mean for the passive control ($j = 1$) and move to step 3 without computing a_{22}.

If necessary, we could find the value of a_{22} by solving the equation

$$\begin{aligned} 1-\alpha \;=\; & \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty I\big[\max(z_1,z_2) - 2^{1/2}a_{22}u < \min(z_1,z_2) + 2^{1/2}a_{22}u\big] \\ & \times \big\{\Phi\big[\min(z_1,z_2) + 2^{1/2}a_{22}u\big] - \Phi\big[\max(z_1,z_2) - 2^{1/2}a_{22}u\big]\big\}^3 \\ & \times \big\{\Phi\big[z_1 + 2^{1/2}a_{22}u\big] - \Phi\big[z_1 - 2^{1/2}a_{22}u\big]\big\}\phi(z_1)\phi(z_2)g(u)\,dz_1\,dz_2\,du, \end{aligned} \tag{26}$$

Table 4: Values of T_{ij}^* and $|T_q^*|$ statistics for Blake and Boockfor's (1997) pituitary gland data

Contrast	i	j	T_{ij}^*	q	$\lvert T_q^*\rvert$
Compound 1 – Passive Control	1	1	−0.386	8	0.386
Compound 2 – Passive Control	2	1	1.075	6	1.075
Compound 3 – Passive Control	3	1	3.075	4	3.075
Compound 4 – Passive Control	4	1	11.645	2	11.645
Compound 1 – Active Control	1	2	0.437	7	0.437
Compound 2 – Active Control	2	2	1.899	5	1.899
Compound 3 – Active Control	3	2	3.899	3	3.899
Compound 4 – Active Control	4	2	12.468	1	12.468

where I is the indicator function, $g(u)$ is the density of S/σ, and ϕ and Φ are the density and c.d.f. of the standard normal distribution. Using numerical integration and the secant method, we find $a_{22} = 2.800$.

For step 3, the set

$$A_3 = \big\{(1,1),(1,2),(2,1),(2,2),(3,1),(3,2)\big\}. \tag{27}$$

This is the complete set of (i,j) pairs for comparing 3 experimental treatments and 2 controls. With $\nu = 30$, we find $a_{23} = 2.745$ from Hoover's (1991) Table 2. As $|T_3^*| = 3.899 \geq 2.745$, we declare the mean for compound 3 ($i = 3$) to be different from the mean for the active control ($j = 2$) and move to step 4.

For step 4, the set

$$A_4 = \big\{(1,1),(1,2),(2,1),(2,2),(3,1)\big\}. \tag{28}$$

As $|T_4^*| = 3.075 > a_{23}$ which is greater than a_{24}, we declare the mean for compound 3 ($i = 3$) to be different from the mean for the passive control ($j = 1$) and move to step 5. If necessary, we could have found the value of a_{24}, by solving

$$\begin{aligned} 1-\alpha \;=\; & \int_0^\infty\!\!\int_{-\infty}^\infty\!\!\int_{-\infty}^\infty I\big[\max(z_1,z_2) - 2^{1/2}a_{24}u < \min(z_1,z_2) + 2^{1/2}a_{24}u\big] \\ & \times \big\{\Phi\big[\min(z_1,z_2) + 2^{1/2}a_{24}u\big] - \Phi\big[\max(z_1,z_2) - 2^{1/2}a_{24}u\big]\big\}^2 \\ & \times \big\{\Phi\big[z_1 + 2^{1/2}a_{24}u\big] - \Phi\big[z_1 - 2^{1/2}a_{24}u\big]\big\}\phi(z_1)\phi(z_2)g(u)\,dz_1\,dz_2\,du, \end{aligned} \tag{29}$$

The solution is $a_{24} = 2.679$.

For step 5, the set

$$A_5 = \big\{(1,1),(1,2),(2,1),(2,2)\big\}. \tag{30}$$

This is the complete set of (i,j) pairs for comparing 2 experimental treatments and 2 controls. With $\nu = 30$, we find $a_{25} = 2.594$ from Hoover's (1991) Table 2. As $|T_5^*| = 1.899 < 2.594$, we stop the test and declare no more differences.

References

[1] Bechhofer, R. E. and Tamhane, A. C. (1981). Incomplete block designs for comparing treatments with a control: General theory. *Technometrics* **23**, 45–57. MR604909

[2] Blake, C. A. and Boockfor, F. R. (1997). Chronic administration of the environmental pollutant 4-tert-octylphenol to adult male rats interferes with the secretion of luteinizing hormone, follicle-stimulating hormone, prolactin, and testosterone. *Biology of Reproduction* **57**, 255–266.

[3] Chernoff, H. and Savage, I. R. (1958). Asymptotic normality and efficiency of certain nonparametric test statistics. *Ann. Math. Statist.* **29**, 972–994. MR100322

[4] Dunnett, C. W. (1955). A multiple comparison procedure for comparing several treatments with a control. *J. Amer. Statist. Assoc.* **50**, 1096–1121.

[5] Dunnett, C. W. (1964). New tables for multiple comparisons with a control. *Biometrics* **20**, 482–491.

[6] Fieller, E. C. (1944). A fundamental formula in the statistics of biological assay, and some applications. *Quarterly J. of Pharmacy and Pharmacology* **17**, 117–123.

[7] Hilden, J. (2000). personal communication.

[8] Hoover, D. R. (1991). Simultaneous comparisons of multiple treatments to two (or more) controls. *Biom. J.* **33**, 913–921.

[9] Jaggi, S., Gupta, V. K., and Parsad, R. (1996). A-efficient block designs for comparing two disjoint sets of treatments. *Comm. Statist. Theory Methods* **25**, 967–983. MR1399205

[10] Mann, H. B. and Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Statist.* **18**, 50–60. MR22058

[11] Majumdar, D. (1986). Optimal designs for comparisons between two sets of treatments. *J. Statist. Plann. Inference* **14**, 359–372. MR859982

[12] Shaffer, J. P. (1977). Multiple comparisons emphasizing selected contrasts: An extension and generalization of Dunnett's procedure. *Biometrics* **33**, 293–303. MR448737

[13] Solorzano, E. (1999). *Simultaneous comparisons with more than one control*, Ph.D. Dissertation, Dept. Statistics, University of South Carolina.

[14] Solorzano, E. and Spurrier, J. D. (1999). One-sided simultaneous comparisons with more than one control. *J. Statist. Comput. Simulation* **63**, 37–46. MR1703043

[15] Solorzano, E. and Spurrier, J. D. (2001a). Nonparametric multiple comparisons with more than one control. *J. Nonparametr. Statist.* **13**, 397–412. MR1932049

[16] Solorzano, E. and Spurrier, J. D. (2001b). Comparing more than one treatment to more than one control in incomplete blocks. *J. Statist. Plann. Inference* **97**, 385–398. MR1861161

[17] Tukey, J. W. (1953). The problem of multiple comparisons. unpublished manuscript.

[18] Zerbe, G. O. (1978). On Fieller's theorem and the general linear model. *Amer. Statist.* **32**, 103–105. MR514336