COMMUN. STATIST.-THEOR. METH., A6(1), 15-32 (1977)

MULTIPLE COMPARISONS IN MODEL I ONE-WAY ANOVA WITH UNEQUAL VARIANCES

Ajit C. Tamhane

Northwestern University
Evanston, Illinois

Key Words & Phrases: fixed effects model; single-stage procedures; two-stage procedures; Behrens-Fisher problem.

ABSTRACT

A fixed effects one-way layout model of analysis of variance is considered where the variances are taken to be possibly unequal. Conservative single-stage procedures based on Banerjee's method for the solution of the Behrens-Fisher problem are proposed for the following multiple comparisons problems: 1) all pairwise comparisons with a control population mean, and 2) all pairwise comparisons and all linear contrasts among the means. Since these procedures are likely to be very conservative in practice, approximate procedures based on Welch's method for the solution of the Behrens-Fisher problem are suggested as alternatives. Monte Carlo studies indicate that the latter are much less conservative and hence may be better in practice. Both these sets of procedures need only the tables of the Student's t-distribution for their application and are very simple to use. Exact two-stage procedures are proposed for the following multiple comparisons problems: 1) all pairwise comparisons and all linear contrasts among the means, and 2) all linear combinations of the means.

1. INTRODUCTION

Consider the following one-way fixed effects model of analysis of variance,

$$X_{ij} = \mu_i + e_{ij}, \qquad (1.1)$$

where for $j = 1,2,\ldots,n_i$ and $i = 1,2,\ldots,k$, all the e_{ij} are independently distributed and $e_{ij} \sim N(0,\sigma_i^2)$. The main purpose of the present paper is to give certain multiple comparisons procedures concerning the populations means μ_i which, although somewhat conservative in nature, are very easy to apply in practice.

When the σ_i^2 are equal but the common value of the variance is unknown, single-stage procedures have been developed for the following multiple comparisons problems:

P1: Joint confidence intervals for all differences $\mu_i - \mu_k$ $(1 \le i \le k - 1)$ where we regard μ_k as the mean of the control population.

P2: Joint confidence intervals for all pairwise differences $\mu_i - \mu_j (1 \le i, j \le k, i \ne j)$ and all linear constrasts $\sum_{i=1}^k c_i\mu_i$ where $c_1,c_2,\ldots,c_k$ are set of arbitrary real constants satisfying $\sum_{i=1}^k c_i = 0$.

P3: Joint confidence intervals for all linear combinations $\sum_{i=1}^k a_i\mu_i$ where $a_1,a_2,\ldots,a_k$ are set of arbitrary real constants.

The original work in this area is due to Dunnett (for P1), Tukey and Scheffé (for P2 and P3); Miller (1966) is an excellent consolidated reference for all these procedures.

When the σ_i^2 are unequal and unknown the problem of multiple comparisons using single-stage procedures becomes relatively difficult. In the case of two populations the problem of comparison of μ_1 and μ_2 is well-known Behrens-Fisher problem. Various approximate methods have been suggested as solutions to this problem; one due to Banerjee (1961) strictly guarantees the specified confidence level for $\mu_1 - \mu_2$; another due to Welch (1938) only approximately guarantees the confidence level and involves Student's t with random number of degrees of freedom (d.f.). Both these methods are sketched briefly in the next section. In Section 3, we develop <u>conservative</u> single-stage procedures for P1 and P2 by making use of the method developed by Banerjee and Slepian (1962) type bounds. As alternatives to these, we propose <u>approximate</u> procedures based on Welch's method in Section 4. In Section 5 we give some Monte Carlo results which indicate that the procedures based on Banerjee's method are too conservative compared to the procedures based on Welch's method. We give some recommendations for the use of these procedures in practice. As a bibliographical note, we mention here that for P3 single-stage procedures have been given by Banerjee (op. cit.) (conservative), Spjøtvoll (1972) (exact Scheffé type) and Hochberg (1976) (exact Tukey type). Hochberg (op. cit.) has also given an approximate procedure for P2 which makes use of Bonferroni bounds and Welch's method.

It is well known that exact solutions can be obtained for the Behrens-Fisher problem using two-stage procedures in the spirit of Stein (1945); see, e.g., Chapman (1950) and Ghosh (1975). These procedures have the added advantage that the width of the confidence interval for $\mu_1 - \mu_2$ can be preassigned. A similar two-stage procedure for P1 has been given by Dudewicz and Ramberg (1972). In Section 6, we extend this work to provide two-stage procedures for P2 and P3. It may be noted that recently Hochberg (1975) has also given two-stage procedures for P2 and P3. Whereas his procedures are based on sample means, our procedures are based on "generalized" sample means. In this paper we make no attempt to compare the two approaches.

2. PRELIMINARIES AND NOTATION

Throughout $\bar{X}_i = \sum_{j=1}^{n_i} X_{ij}/n_i$ will denote the sample mean based on n_i observations from $N(\mu_i,\sigma_i^2)$ and S_i^2 will denote an unbiased estimate of σ_i^2 based on ν_i d.f. which is distributed independently of $\bar{X}_i$; usually one would use

$$S_i^2 = \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2/(n_i - 1)$$

with $\nu_i = n_i - 1$ d.f. $(1 \le i \le k)$. Also $t_{\nu,\beta}$ will denote the upper β point of the Student's t-distribution with ν d.f.

Let us first briefly sketch the two methods due to Banerjee and Welch for the solution of the Behrens-Fisher problem.

Banerjee's Method: The confidence coefficient of the statement

$$\mu_1 - \mu_2 \in [\bar{X}_1 - \bar{X}_2 \pm (\frac{t^2_{\nu_1,\alpha/2}S_1^2}{n_1} + \frac{t^2_{\nu_2,\alpha/2}S_2^2}{n_2})^{1/2}] \qquad (2.1)$$

is at least $1 - \alpha$. The exact value of $1 - \alpha$ is attained only when $\sigma_1^2/\sigma_2^2 = 0$ or ∞. This method is based on the following lemma due to Banerjee (op. cit.):

Lemma 2.1: Let U be a chi-square random variable with 1 d.f. and let V_i be chi-square random variables with ν_i d.f. $(1 \le i \le k)$, which are distributed mutually independently and also independently of U. Let $\lambda_i \geq 0$ $(1 \leq i \leq k)$ be a set of constants such that $\sum_{i=1}^k \lambda_i = 1$. Then

$$\Pr\{U \leq \sum_{i=1}^{k} t^2_{\nu_i,\alpha/2}(\lambda_i V_i/\nu_i)\} \geq 1 - \alpha.$$

Welch's Method: The confidence coefficient of the statement

$$\mu_1 - \mu_2 \in [\bar{X}_1 - \bar{X}_2 \pm t_{\hat{\nu}_{12},\alpha/2}(S_1^2/n_1 + S_2^2/n_2)^{1/2}] \qquad (2.2)$$

is approximately $1 - \alpha$. In (2.2)

$$\hat{\nu}_{12} = \frac{(S_1^2/n_1 + S_2^2/n_2)^2}{[S_1^4/n_1^2(n_1 - 1) + S_2^4/n_2^2(n_2 - 1)]}. \qquad (2.3)$$

Note that $\hat{\nu}_{12}$ is random and arbitrary (not necessarily an integer). The value of $t_{\hat{\nu}_{12},\alpha/2}$ can be found by interpolating in the t-tables. Wang (1971) has extensively studied Welch's intervals and found that the actual confidence levels are fairly close to the specified value $1 - \alpha$.

To develop the multiple comparisons procedures based on Banerjee's method we need the following additional lemmas:

Lemma 2.2 (Slepian (op. cit.)): Let $Y_1,\ldots,Y_k$ $(Z_1,\ldots,Z_k)$ be standard normal random variables with the correlation matrix $\{\rho_{ij}\}(\{\eta_{ij}\})$. If $\rho_{ij} \geq \eta_{ij}$ $1 \leq i < j \leq k$, then for any set of real constants $a_1,\ldots,a_k$, we have

$$\Pr\{Y_1 \leq a_1,\ldots,Y_k \leq a_k\} \geq \Pr\{Z_1 \leq a_1,\ldots,Z_k \leq a_k\}.$$

Lemma 2.3 (Šidák (1967)): Let $Y_1,\ldots,Y_k$ be standard normal random variables with an arbitrary correlation matrix $\{\rho_{ij}\}$. Then for any set of nonnegative constants $a_1,\ldots,a_k$, we have

$$\Pr\{|Y_1| \leq a_1,\ldots,|Y_k| \leq a_k\} \geq \prod_{i=1}^{k} \Pr\{|Y_i| \leq a_i\}.$$

For the next lemma, we need the following definition due to Esary, Proschan and Walkup (1967).

Definition: Random variables $Y_1, Y_2,\ldots,Y_k$ are said to be associated if for all real valued nondecreasing functions $\emptyset$ and Ψ of k arguments we have

$$\text{cov}(\emptyset(Y_1,\ldots Y_k),\ \Psi(Y_1,\ldots,Y_k)) \geq 0.$$

Now in the following lemma we have a generalization of a result obtained by Kimball (1951).

Lemma 2.4: Let $X_1,\ldots,X_k$ be independent real valued random variables and let $\Psi_j(x_1,\ldots,x_k)$ $(1 \leq j \leq p)$ be nonnegative real valued functions each of which is nondecreasing in each of its arguments x_i $(1 \leq i \leq k)$. Then denoting $Y_j = \Psi_j(X_1,\ldots,X_k)$ we have

$$E\{\prod_{j=1}^{p} Y_j\} \geq \prod_{j=1}^{p} E\{Y_j\}.$$

Proof: By Theorem 2.1 of Esary et al. (op. cit.), $X_1,\ldots,X_k$ are associated. Then by property P4 of Esary et al. (op. cit.) $Y_1,\ldots,Y_p$ are associated. Now let $\emptyset(Y_1,\ldots,Y_p) = Y_1$ and $\Psi(Y_1,\ldots,Y_p) = \prod_{j=2}^{p} Y_j$. Then $\emptyset$ and Ψ are nondecreasing functions of $Y_1,\ldots,Y_p$ and hence by the above mentioned property P4, they are associated. Therefore

$$\mathrm{cov}(\phi(Y_1,\ldots,Y_p),\ \Psi(Y_1,\ldots,Y_p)) \geq 0$$

$$\Rightarrow E\{\prod_{j=1}^{p} Y_j\} \geq E\{Y_1\}E\{\prod_{j=2}^{p} Y_j\}.$$

Repeated use of this argument gives the final result.

3. SINGLE-STAGE PROCEDURES BASED ON BANERJEE'S METHOD

3.1 Multiple Comparisons with a Control (Problem P1)

Consider the model in (1.1) and suppose that μ_k is to be regarded as the mean of the control population. Then the following theorem gives the upper/lower one-sided and two-sided joint confidence intervals for all differences $\mu_i - \mu_k (1 \leq i \leq k - 1)$.

Theorem 3.1: The joint confidence coefficient of each of the following families of confidence intervals is at least $1 - \alpha$ if $\beta \leq 1 - (1 - \alpha)^{1/(k-1)}$:

(i) Upper one-sided: For $1 \leq i \leq k - 1$,

$$\mu_i - \mu_k \leq \bar{X}_i - \bar{X}_k + \left(\frac{t^2_{\nu_i,\beta} S_i^2}{n_i} + \frac{t^2_{\nu_k,\beta} S_k^2}{n_k}\right)^{1/2}. \tag{3.1}$$

(ii) Lower one-sided: For $1 \leq i \leq k - 1$,

$$\mu_i - \mu_k \geq \bar{X}_i - \bar{X}_k - \left(\frac{t^2_{\nu_i,\beta} S_i^2}{n_i} + \frac{t^2_{\nu_k,\beta} S_k^2}{n_k}\right)^{1/2}. \tag{3.2}$$

(iii) Two-sided: For $1 \leq i \leq k - 1$,

$$\mu_i - \mu_k \in [\bar{X}_i - \bar{X}_k \pm \left(\frac{t^2_{\nu_i,\beta/2} S_i^2}{n_i} + \frac{t^2_{\nu_k,\beta/2} S_k^2}{n_k}\right)^{1/2}]. \tag{3.3}$$

Proof: We shall give the proof only for (3.1); the proofs of (3.2) and (3.3) are similar. In the proof of (3.3) we have to use Lemma 2.3 (Šidák inequality) instead of Lemma 2.2 (Slepian inequality); the latter is used in the proof of (3.1) below. Let P denote the actual confidence coefficient for (3.1). Then we have

$$P = \Pr\{\mu_i - \mu_k \leq \bar{X}_i - \bar{X}_k + \left(\frac{t^2_{\nu_i,\beta} S_i^2}{n_i} + \frac{t^2_{\nu_k,\beta} S_k^2}{n_k}\right)^{1/2} (1 \leq i \leq k-1)\}$$

$$= \Pr\{Z_i \leq [\left(\frac{t^2_{\nu_i,\beta} S_i^2}{n_i} + \frac{t^2_{\nu_k,\beta} S_k^2}{n_k}\right) / \left(\frac{\sigma_i^2}{n_i} + \frac{\sigma_k^2}{n_k}\right)]^{1/2} (1 \leq i \leq k-1)\}, \tag{3.4}$$

where for $1 \leq i \leq k - 1$,

$$Z_i = [\bar{X}_k - \bar{X}_i - (\mu_k - \mu_i)] / [\sigma_i^2/n_i + \sigma_k^2/n_k]^{1/2},$$

are standard normal random variables which are distributed independently of $\underset{\sim}{S}^2 = (S_1^2, S_2^2, \ldots, S_k^2)$. By conditioning on $\underset{\sim}{S}^2$ and noting that corr $(Z_i, Z_j) \geq 0$ for all $i \neq j$, we can use Lemma 2.1 to obtain from (3.4)

$$P \geq E\{\prod_{i=1}^{k-1} \Phi[\left(\frac{t^2_{\nu_i,\beta} S_i^2}{n_i} + \frac{t^2_{\nu_k,\beta} S_k^2}{n_k}\right)^{1/2} / \left(\frac{\sigma_i^2}{n_i} + \frac{\sigma_k^2}{n_k}\right)^{1/2}]\}$$

$$= E\{\prod_{i=1}^{k-1} \Psi_i(S_1^2, \ldots, S_k^2)\} \text{ (say)}, \tag{3.5}$$

where $\Phi(\cdot)$ denotes the standard normal cdf and the expectation in (3.5) is w.r.t. $\underset{\sim}{S}^2$. Now we note that $S_1^2, \ldots, S_k^2$ are independently distributed and $\Psi_1, \ldots, \Psi_{k-1}$ are nondecreasing functions of $S_1^2, \ldots, S_k^2$. Therefore by applying Lemma 2.4 we obtain

$$P \geq \prod_{i=1}^{k-1} \Pr\{Z_i \leq [\left(\frac{t^2_{\nu_i,\beta} S_i^2}{n_i} + \frac{t^2_{\nu_k,\beta} S_k^2}{n_k}\right) / \left(\frac{\sigma_i^2}{n_i} + \frac{\sigma_k^2}{n_k}\right)]^{1/2}\}$$

$$= \prod_{i=1}^{k-1} [1/2 + 1/2 \Pr\{Z_i^2 \leq t^2_{\nu_i,\beta} \lambda_i (S_i/\sigma_i)^2 + t^2_{\nu_k,\beta} (1 - \lambda_i)(S_k/\sigma_k)^2\}] \tag{3.6}$$

where $\lambda_i = (\sigma_i^2/n_i)/(\sigma_i^2/n_i + \sigma_k^2/n_k)$. We note that $(S_i/\sigma_i)^2$ are independently distributed as $\chi^2_{\nu_i}/\nu_i$ for $1 \leq i \leq k$. Applying Lemma 2.1

we obtain from (3.6)

$$P \geq \prod_{i=1}^{k-1} [1/2 + (1 - 2\beta)/2] = (1 - \beta)^{k-1} \geq 1 - \alpha.$$

Hence the theorem is proved.

3.2 Pairwise comparisons and Linear Contrasts (Problem P2)

Theorem 3.2: The joint confidence coefficient of all confidence statements for $1 \leq i < j \leq k$,

$$\mu_i - \mu_j \in [\bar{X}_i - \bar{X}_j \pm (\frac{t^2_{\nu_i,\beta/2} S_i^2}{n_i} + \frac{t^2_{\nu_j,\beta/2} S_j^2}{n_j})^{1/2}] \qquad (3.7)$$

is at least $1 - \alpha$ if $\beta \leq 1 - (1 - \alpha)^{2/k(k-1)}$.

Proof: The proof is similar to the proof of Theorem 3.1 and is hence omitted.

Corollary: The joint confidence coefficient of all confidence statements,

$$\sum_{i=1}^{k} c_i\mu_i \in [\sum_{i=1}^{k} c_i\bar{X}_i \pm \{\frac{2\sum_{i\in\theta_c}\sum_{j\in\eta_c} c_i(-c_j)(\frac{t^2_{\nu_i,\beta/2} S_i^2}{n_i} + \frac{t^2_{\nu_j,\beta/2} S_j^2}{n_j})^{1/2}}{\sum_{i=1}^{k}|c_i|}\}] \qquad (3.8)$$

for all contrasts $(c_1,\ldots,c_k;\ \sum_{i=1}^{k} c_i = 0)$ is at least $1 - \alpha$ if $\beta \leq 1 - (1 - \alpha)^{2/k(k-1)}$. In the above $\theta_c = \{i: c_i > 0\}$ and $\eta_c = \{j: c_j < 0\}$.

Proof: Follows from Lemma 3.1 of Hochberg (1974).

4. SINGLE-STAGE PROCEDURES BASED ON WELCH'S METHOD

In view of the conservative nature of the procedures based on Banerjee's method, we propose some approximate procedures based on Welch's method as alternatives to the previous procedures. The

formulae for the confidence intervals based on Welch's method are obtained by replacing $\{[(t^2_{\nu_i,\beta} S_i^2)/n_i] + [(t^2_{\nu_j,\beta} S_j^2)/n_j]\}^{1/2}$ terms in the appropriate previous formulae ((3.1) - (3.3), (3.7) and (3.8)) by $t_{\hat{\nu}_{ij},\beta}[(S_i^2/n_i) + (S_j^2/n_j)]^{1/2}$ where $\hat{\nu}_{ij}$ is obtained from (2.3) with obvious changes in the notation.

Thus these confidence intervals are in the same spirit as the previous ones (lower bounding a joint probability by the product of the individual probabilities) except that Welch's method is used for approximating the individual probabilities. Note that the joint confidence coefficient of these confidence intervals would be only close to the specified value of $(1 - \alpha)$ and in some situations may be less than $(1 - \alpha)$.

5. APPLICATIONS AND MONTE CARLO RESULTS

First we shall illustrate the use of the proposed procedures by means of an example.

Example

We consider the following data analyzed by Hochberg (1976). We have $k = 4$, $\underset{\sim}{n} = (6,6,6,6)$ and $\underset{\sim}{S}^2 = (178,60,98,68)$. It is desired to obtain 95% joint confidence intervals for all pairwise differences $\mu_i - \mu_j (1 \leq i < j \leq 4)$.

(i) Conservative intervals based on Banerjee's method: We obtain $\beta = 1 - (.05)^{1/6} = .0085$ and $t_{5,.00425} = 4.20$ (from Pearson and Hartley (1956), p. 132). Thus we obtain the following intervals for all six pairwise comparisons:

$$\mu_i - \mu_j \in [\bar{X}_i - \bar{X}_j \pm \frac{4.20}{\sqrt{6}}(S_i^2 + S_j^2)^{1/2}] \quad (1 \leq i < j \leq 4). \qquad (5.1)$$

(ii) Approximate intervals based on Welch's method: Using (2.3) we obtain $\hat{\nu}_{12} = 8.026$, $\hat{\nu}_{13} = 9.224$, $\hat{\nu}_{14} = 8.328$, $\hat{\nu}_{23} = 9.455$, $\hat{\nu}_{24} = 9.962$ and $\hat{\nu}_{34} = 9.692$. By doing linear interpolations in Table 9 of Pearson and Hartley (op. cit.) we find for $\beta/2 = .00425$, $t_{\hat{\nu}_{12},\beta/2} = 3.4642$, $t_{\hat{\nu}_{13},\beta/2} = 3.3344$, $t_{\hat{\nu}_{14},\beta/2} = 3.4300$,

$t_{\hat{\nu}_{23},\beta/2} = 3.3145$, $t_{\hat{\nu}_{24},\beta/2} = 3.2709$ and $t_{\hat{\nu}_{34},\beta/2} = 3.2941$. The confidence intervals are then given by the formula:

$$\mu_i - \mu_j \in [\bar{X}_i - \bar{X}_j \pm \frac{t_{\hat{\nu}_{ij},\beta/2}}{\sqrt{6}}(S_i^2 + S_j^2)^{1/2}] \quad (1 \le i < j \le 4). \quad (5.2)$$

For the same problem Hochberg (1976) has given the following two sets of approximate confidence intervals:

$$\text{(iii)}\ \mu_i - \mu_j \in [\bar{X}_i - \bar{X}_j \pm \frac{5.05}{\sqrt{6}} \max(S_i, S_j)] \quad (1 \le i < j \le 4). \quad (5.3)$$

This is obtained by applying the result of Theorem 2.1 of Hochberg (1976) regarding the confidence intervals for all linear combinations of the μ's and approximating the augmented range distribution. (See Miller (op. cit.) for the definitions of these distributions.)

$$\text{(iv)}\ \mu_i - \mu_j \in [\bar{X}_i - \bar{X}_j \pm \frac{3.36}{\sqrt{6}}(S_i^2 + S_j^2)^{1/2}] \quad (1 \le i < j \le 4). \quad (5.4)$$

This is obtained by using Bonferroni bounds and the Welch approximation.

In this specific example we find that (5.1) gives the widest confidence intervals, (5.2) gives shorter confidence intervals than (5.3) in all the comparisons and than (5.4) in 4 out of 6 comparisons. Also (5.4) gives shorter confidence intervals than (5.3) in 4 out of 6 comparisons. Thus it would appear that the intervals based on Banerjee's method are most conservative. However one must bear in mind that only for these intervals it can be rigorously proved that the specified confidence level is guaranteed. They are also easiest to compute and need only readily available t-tables for their application.

Monte Carlo experiments were performed to study the actual confidence levels attained by the multiple comparisons procedures based on Banerjee's method and Welch's method. The problem of constructing joint confidence intervals for all pairwise differences

of the population means was considered for k = 4, 6, and 8, and $1 - \alpha = 0.90$, 0.95, and 0.99. For each combination of values of k and $1 - \alpha$, various configurations of values of σ_i^2 and n_i were studied. For each case, for $1 - \alpha$ - 0.90 and 0.95, N = 1000 experiments were carried out and for $1 - \alpha = 0.99$, N = 2000 experiments were carried out. In each experiment, k independent pairs of values of $\bar{X}_i \sim N(0, \sigma_i^2/n_i)$ and $S_i^2 \sim \sigma_i^2 \chi^2_{n_i - 1}/(n_i - 1)$ were generated and the confidence intervals were computed for both the procedures for all pairwise differences $\mu_i - \mu_j$ $(1 \le i < j \le k)$. The necessary values of the upper points of the t-variables were computed by linear interpolation in Table 9 of Pearson and Hartley (op. cit.). For each procedure an estimate of the actual confidence level was found by calculating the fraction of the total number of experiments in which all the confidence intervals covered the corresponding pairwise differences $\mu_i - \mu_j$ $(1 \le i < j \le k)$. The results of the experiments are given in Table I.

We find that both the procedures guarantee the specified confidence levels for $1 - \alpha = 0.90$ and 0.95. But the procedure based on Welch's method fails to guarantee the specified confidence level of 0.99 in some cases. (The estimated confidence level is less than 0.99 by a statistically significant amount.) Apparently this occurs when the configurations of σ_i^2 and n_i are unbalanced, i.e., more observations are taken on the populations having smaller variances and vice-versa. We also note that the procedure based on Banerjee's method is highly conservative.

Based on this Monte Carlo study and previous theoretical work we conclude that the procedures based on Welch's method may be better in practice (give shorter confidence intervals having approximately the specified confidence level). The procedures based on Banerjee's method may still be useful in practice for short cut and quick computations and for severely unbalanced configurations of σ^2-values. Hochberg's procedure used in (5.4) would be unattractive in practice because of its trial and error nature. To implement his

TABLE I

Results of Monte Carlo Experiments[1,2]

k	$\sigma_1^2,\ldots,\sigma_k^2$	$n_1,\ldots,n_k$	Estimates of confidence levels $1-\alpha = .90$ B	W	$1-\alpha = .95$ B	W	$1-\alpha = .99$ B	W
4	1,1,1,1	5,5,5,5	.978	.930	.992	.966	.9990	.9935
	1,1,1,1	5,7,9,11	.962	.925	.986	.959	.9995	.9915
	1,2,3,4	5,5,5,5	.971	.926	.992	.964	1.0000	.9960
	1,2,3,4	5,7,9,11	.970	.938	.988	.970	1.0000	.9930
	1,2,3,4	11,9,7,5	.950	.915	.984	.954	.9985	.9880
	1,4,7,10	5,5,5,5	.971	.927	.991	.960	.9995	.9895
	1,4,7,10	5,7,9,11	.968	.930	.984	.968	.9995	.9975
	1,4,7,10	11,9,7,5	.959	.925	.980	.957	.9955	.9880
6	1,1,1,1,1,1	7,7,7,7,7,7	.981	.927	.993	.965	.9995	.9950
	1,1,1,1,1,1	7,9,9,11,11,13	.961	.927	.987	.955	.9990	.9950
	1,2,2,3,3,4	7,7,7,7,7,7	.978	.933	.990	.967	.9995	.9915
	1,2,2,3,3,4	7,9,9,11,11,13	.973	.923	.987	.966	.9990	.9935
	1,2,2,3,3,4	13,11,11,9,9,7	.961	.928	.987	.958	.9975	.9930
	1,4,4,7,7,10	7,7,7,7,7,7	.969	.945	.988	.966	1.0000	.9940
	1,4,4,7,7,10	7,9,9,11,11,13	.973	.936	.987	.964	.9985	.9915
	1,4,4,7,7,10	13,11,11,9,9,7	.961	.924	.983	.958	.9960	.9870
8	1,1,1,1,1,1,1,1	9,9,9,9,9,9,9,9	.977	.935	.997	.969	.9995	.9930
	1,1,1,1,1,1,1,1	9,9,11,11,13,13,15,15	.979	.944	.990	.972	.9990	.9925
	1,1,2,2,3,3,4,4	9,9,9,9,9,9,9,9	.980	.940	.994	.974	.9980	.9935
	1,1,2,2,3,3,4,4	9,9,11,11,13,13,15,15	.971	.939	.989	.962	1.0000	.9935
	1,1,2,2,3,3,4,4	15,15,13,13,11,11,9,9	.965	.925	.984	.962	.9975	.9915
	1,1,4,4,7,7,10,10	9,9,9,9,9,9,9,9	.969	.929	.987	.962	.9990	.9895
	1,1,4,4,7,7,10,10	9,9,11,11,13,13,15,15	.978	.949	.988	.974	.9985	.9925
	1,1,4,4,7,7,10,10	15,15,13,13,11,11,9,9	.966	.943	.984	.965	.9985	.9920

1. B = Multiple comparisons procedure based on Banerjee's method.
2. W = Multiple comparisons procedure based on Welch's method.

procedure used in (5.3) in practice, would require tables of the augmented range distribution (or for approximate purposes, the range distribution) of t-variables for all possible combinations of $\nu_1,\ldots\nu_k$ which are commonly encountered. The difficulty of constructing such tables would inhibit the application of this latter procedure.

6. TWO-STAGE PROCEDURES

In deriving two-stage procedures we use the following basic result.

Lemma 6.1 (Stein (op. cit.)): Let X_i $(i = 1,2,\ldots)$ be i.i.d. $N(\mu,\sigma^2)$ random variables. Let n be a fixed nonnegative integer and let S^2 be an unbiased estimate of σ^2, based on ν d.f., distributed independently of $\sum_{i=1}^n X_i$ and $X_{n+1}, X_{n+2},\ldots$. If a positive integer N satisfies

$$N \geq \max\{n + 1, [(S/d)^2]\}$$

where $d > 0$ is an arbitrary constant and $[x]$ denotes the smallest integer $\geq x$ then there exist real numbers $A_1, A_2, \ldots, A_N$ such that

$$A_1 = \ldots = A_n, \ \sum_{i=1}^N A_i = 1 \ \text{and} \ S^2 \sum_{i=1}^N A_i^2 = d^2.$$

Further $(\sum_{i=1}^N A_i X_i - \mu)/d$ has a Student's t-distribution with ν d.f.

6.1 Pairwise Comparisons and Linear Contrasts (Problem P2)

In this case the experimenter may have one of the following goals for specified values of constants $d > 0$ and $0 < \alpha < 1$.

Goal I: Establish joint two-sided confidence intervals for all pairwise differences $\mu_i - \mu_j (1 \leq i < j \leq k)$ of width (for each pairwise difference) = 2d and overall confidence coefficient = $1 - \alpha$.

Goal II: Establish joint two-sided confidence intervals for all contrasts $\sum_{i=1}^k c_i\mu_i$ where $\underset{\sim}{c} = (c_1,\ldots,c_k) \in \mathcal{C} = \{\underset{\sim}{c} \in \mathbb{R}^k; \sum_{i=1}^k c_i = 0, \sum_{i=1}^k |c_i| = 2\}$ with width (for each such contrast) = 2d and overall confidence coefficient = $1 - \alpha$.

Let $f_\alpha(\nu_1,\ldots,\nu_k)$ denote the upper α point of the range $R(t_{\nu_1},\ldots,t_{\nu_k}) = \max_{1\leq i\leq k} t_{\nu_i} - \min_{1\leq i\leq k} t_{\nu_i}$ of k independent Student t-variables with $\nu_1,\ldots,\nu_k$ d.f. which are denoted by $t_{\nu_1},\ldots,t_{\nu_k}$, respectively. Hochberg (1976) has tabulated the values of $f_\alpha(\nu_1,\ldots,\nu_k)$ when $\nu_1 = \ldots = \nu_k = \nu$ (say), $\alpha = .05$, .10 and for selected values of ν. If the experiment is a designed one--which is usually the case when using two-stage procedures, it seems reasonable to demand that all ν_i be chosen equal. In such situations the difficulty of tabulating the values of $f_\alpha(\nu_1,\ldots,\nu_k)$ for all practically encountered combinations of $\nu_1,\ldots,\nu_k$ would not be a major obstacle in applying our two-stage procedure R_1 which we propose below. In Theorem 6.1 we show that R_1 guarantees the fulfillment of Goals I and II.

Procedure R_1: 1) In the first stage take independent observations $X_{ij}(1 \le j \le n_i)$ from $N(\mu_i,\sigma_i^2)$. Let S_i^2 denote an unbiased estimate of σ_i^2 based on ν_i d.f. $(1 \le i \le k)$ which is distributed independently of $\sum_{j=1}^{n_i} X_{ij}$ and $X_{i,n_i+1}, X_{i,n_i+2},\ldots$ etc.

2) Let $N_i = \max\{n_i + 1, [S_i^2 f_\alpha^2(\nu_1,\ldots,\nu_k)/d^2]\}$. Take additional independent observations $X_{ij}(n_i + 1 \le j \le N_i)$ from $N(\mu_i,\sigma_i^2)(1 \le i \le k)$. Choose real numbers $A_{ij}(1 \le j \le N_i)$ satisfying

$$A_{i1} = \ldots = A_{in_i}, \quad \sum_{j=1}^{N_i} A_{ij} = 1, \text{ and } S_i^2 \sum_{j=1}^{N_i} A_{ij}^2 = d^2/f_\alpha^2(\nu_1,\ldots,\nu_k),$$

and compute generalized sample means $\tilde{\bar{X}}_i = \sum_{j=1}^{N_i} A_{ij}X_{ij}$ for $1 \le i \le k$.

3) (i) Assert that Goal I is fulfilled by the set of joint confidence intervals

$$\mu_i - \mu_j \in [\tilde{\bar{X}}_i - \tilde{\bar{X}}_j \pm d] \quad (1 \le i < j \le k).$$

(ii) Assert that Goal II is fulfilled by the set of joint confidence intervals

$$\sum_{i=1}^{k} c_i\mu_i \in [\sum_{i=1}^{k} c_i\tilde{\bar{X}}_i \pm d]$$

for $\underset{\sim}{c} \in \mathcal{C}$.

Theorem 6.1: Procedure R_1 guarantees the fulfillment of Goals I and II.

Proof:

$$\Pr\{\mu_i - \mu_j \in [\tilde{\bar{X}}_i - \tilde{\bar{X}}_j \pm d] \ \forall\ i < j\}$$

$$= \Pr\{|[(\tilde{\bar{X}}_i - \mu_i)/\{S_i(\sum_{\ell=1}^{N_i} A_{i\ell}^2)^{1/2}\}] - [(\tilde{\bar{X}}_j - \mu_j)/\{S_j(\sum_{\ell=1}^{N_j} A_{j\ell}^2)^{1/2}\}]|$$

$$\le f_\alpha(\nu_1,\ldots,\nu_k) \ \forall\ i < j\}$$

$$= \Pr\{\max_{1\le i\le k} t_{\nu_i} - \min_{1\le i\le k} t_{\nu_i} \le f_\alpha(\nu_1,\ldots,\nu_k)\}$$

$$= 1 - \alpha.$$

Hence R_1 fulfills Goal I. We have made use of Lemma 6.1 in concluding that $(\tilde{\bar{X}}_i - \mu_i)/[S_i(\sum_{\ell=1}^{N_i} A_{i\ell}^2)^{1/2}]$ are distributed independently as t_{ν_i} for $1 \le i \le k$ in the above proof. Fulfillment of Goal II now follows from Lemma 1, p. 44 of Miller (op. cit.).

6.2 Linear Combinations (Problem P3)

In this case the experimenter may have the following goal:

Goal III: For specified constants $d > 0$, $0 < \alpha < 1$, establish joint two-sided confidence intervals for all linear combinations $\sum_{i=1}^k a_i\mu_i$ where $\underset{\sim}{a} = (a_1,a_2,\ldots,a_k) \in \mathcal{A} = \{\underset{\sim}{a} \in \mathbb{R}^k; \sum_{i=1}^k a_i^2 = 1\}$, with width (for each such linear combination) = 2d and overall confidence coefficient = $1 - \alpha$.

Let $g_\alpha(\nu_1,\ldots,\nu_k)$ denote the upper α point of the distribution of $\sum_{i=1}^k t_{\nu_i}^2$ where $t_{\nu_1},\ldots,t_{\nu_k}$ are independent Student t-variables with $\nu_1,\ldots,\nu_k$ d.f. respectively. This distribution is not yet tabulated but Spjøtvoll (op. cit.) has given the following approximation to it:

$$g_\alpha(\nu_1,\ldots,\nu_k) \cong m_1F_\alpha(k,m_2). \qquad (6.1)$$

In (6.1) $F_\alpha(k,m_2)$ denotes the upper α point of the F-distribution with k and m_2 d.f.,

$$m_2 = \frac{(k-2)[\sum_{i=1}^{k}\{\nu_i/(\nu_i-2)\}]^2 + 4k\sum_{i=1}^{k}\{\nu_i^2(\nu_i-1)/(\nu_i-2)^2(\nu_i-4)\}}{k\sum_{i=1}^{k}\{\nu_i^2(\nu_i-1)/(\nu_i-2)^2(\nu_i-4)\} - [\sum_{i=1}^{k}\{\nu_i/(\nu_i-2)\}]^2}, \quad (6.2)$$

and

$$m_1 = (1 - 2/m_2)\sum_{i=1}^{k}\{\nu_i/(\nu_i - 2)\}. \quad (6.3)$$

Now we propose the following Scheffé-type procedure which guarantees the fulfillment of Goal III as shown in Theorem 6.2 below.

Procedure R_2: 1) Same as in R_1.
2) Same as in R_1 with $f_\alpha(\nu_1,\ldots,\nu_k)$ replaced by $(g_\alpha(\nu_1,\ldots,\nu_k))^{1/2}$.
3) Assert that Goal III is fulfilled by the set of joint confidence intervals

$$\sum_{i=1}^{k} a_i\mu_i \in [\sum_{i=1}^{k} a_i\tilde{\bar{X}}_i \pm d]$$

for all $\underset{\sim}{a} \in \mathcal{A}$.

Theorem 6.2: Procedure R_2 guarantees the fulfillment of Goal III.

Proof:

$$\Pr\{\sum_{i=1}^{k} a_i\mu_i \in [\sum_{i=1}^{k} a_i\tilde{\bar{X}}_i \pm d]\ \forall \underset{\sim}{a} \in \mathcal{A}\}$$

$$= \Pr\{|\sum_{i=1}^{k}\{a_i(\tilde{\bar{X}}_i - \mu_i)/S_i(\sum_{\ell=1}^{N_i} A_{i\ell}^2)^{1/2}\}| \le (g_\alpha(\nu_1,\ldots,\nu_k))^{1/2}\ \forall \underset{\sim}{a} \in \mathcal{A}\}$$

$$= \Pr\{|\sum_{i=1}^{k} a_i t_{\nu_i}| \le [g_\alpha(\nu_1,\ldots,\nu_k)(\sum_{i=1}^{k} a_i^2)]^{1/2}\ \forall \underset{\sim}{a} \in \mathcal{A}\}$$

$$= \Pr\{\sum_{i=1}^{k} t_{\nu_i}^2 \le g_\alpha(\nu_1,\ldots,\nu_k)\}$$

$$= 1 - \alpha.$$

In the second to last step above we have made use of Lemma 2, p. 63 of Miller (op. cit.).

A two-stage Tukey type procedure can be derived based on the distribution of the augmented range

$$\bar{R}(t_{\nu_1},\ldots,t_{\nu_k}) = \max\{\max_{1\le i\le k}|t_{\nu_i}|, R(t_{\nu_1},\ldots,t_{\nu_k})\}$$

of k independent Student t-variables $t_{\nu_1},\ldots,t_{\nu_k}$. For this procedure the width of the confidence interval for each linear combination $\sum_{i=1}^{k} a_i\mu_i$ would be $2dM(a_1,\ldots,a_k)$ where $M(a_1,\ldots,a_k) = \{\sum_{i\in\mathcal{P}_a} a_i, -\sum_{j\in\mathcal{N}_a} a_j\}$. The details of this procedure are omitted for brevity.

ACKNOWLEDGEMENTS

This work was partially supported by the Office of Research and Sponsored Programs and the Technological Institute, Northwestern University.

BIBLIOGRAPY

1. Banerjee, S. K. (1961). On confidence interval for two-means problem based on separate estimates of variances and tabulated values of t-variable. Sankhyā A 23, 359-78.

2. Chapman, D. G. (1950). Some two-sample tests. Ann. Math. Statist. 21, 601-6.

3. Dudewicz, E. J. and Ramberg, J. S. (1972). Multiple comparisons with a control: unknown variances. Annual Tech. Conf. Trans. Amer. Soc. Quality Control 26, 483-8.

4. Esary, J. D., Proschan, F., and Walkup, D. W. (1967). Association of random variables with applications. Ann. Math. Statist. 38, 1466-74.

5. Ghosh, B. K. (1975). A two-stage procedure for the Behrens-Fisher problem. J. Amer. Statist. Assoc. 70, 457-62.

6. Hochberg, Y. (1974). Some generalizations of the T-method in simultaneous inference. J. Multivariate Anal. 4, 224-34.

7. Hochberg, Y. (1975). Simultaneous inference under Behrens-Fisher conditions--a two sample approach. Commun. Statist. 4, 109-19.

8. Hochberg, Y. (1976). A modification of the T-method of multiple comparisons for a one-way layout with unequal variances. J. Amer. Statist. Assoc. 71, 200-3.

9. Kimball, A. W. (1951). On dependent tests of significance in the analysis of variance. Ann. Math. Statist. 22, 600-2.

10. Miller, R. G. (1966). Simultaneous Statistical Inference. New York: McGraw-Hill Book Co.

11. Pearson, E. S. and Hartley, H. O. (1956). Biometrika Tables for Statisticians, Vol. I. Cambridge: Cambridge University Press.

12. Šidák, Z. (1967). Rectangular confidence regions for the means of multivariate normal distributions. J. Amer. Statist. Assoc. 62, 626-33.

13. Slepian, D. (1962). The one-sided barrier problem for gaussian noise. Bell System Tech. J. 41, 463-501.

14. Spjøtvoll, E. (1972). Joint confidence intervals for all linear functions in ANOVA with unknown variances. Biometrika 59, 684-5.

15. Stein, C. (1945). A two-sample test for a linear hypothesis whose power is independent of the variance. Ann. Math. Statist. 16, 243-58.

16. Wang, Y. Y. (1971). Probabilities of Type I errors of Welch tests for the Behrens-Fisher problem. J. Amer. Statist. Assoc. 66, 605-8.

17. Welch, B. L. (1938). The significance of the difference between two means when the population variances are unequal. Biometrika 25, 350-62.

Received May 1976; revised June 1976; retyped September 1976.

Recommended by Meyer Dwass, Northwestern University, Evanston, IL

Refereed by Rupert G. Miller, Stanford University, Stanford, CA