BU-825-M

MULTIPLE COMPARISONS PROCEDURES FOR SOME SPLIT PLOT AND SPLIT BLOCK DESIGNS

Walter T. Federer
Biometrics Unit
Cornell University
Ithaca, New York

Charles E. McCulloch
Biometrics Unit
Cornell University
Ithaca, New York

1. INTRODUCTION

The subject of multiple comparisons for a set of v means has received considerable attention in scientific and textbook literature. Some of the earlier references are Duncan (1947, 1955), Keuls (1952), Tukey (1953), Federer (1955), Dunnett (1955), and Miller (1966). More recent discussions are given by Miller (1976), and Chew (1977). These latter two references contain bibliographies.

A closely related area is that of ranking and selection. Historically, the area has been divided into two branches, the indifference zone approach and the subset selection approach (Gibbons, Olkin and Sobel, 1977). Some early references are Bechhofer (1954), Gupta and Sobel (1958), Paulson (1964), Gupta (1965) and Bechhofer, Kiefer and Sobel (1968). Recently, however, Hsu (1981, 1982) has demonstrated a relationship between the two branches and also between them and a special kind of multiple comparisons, called simultaneous confidence intervals with the best.

All of these procedures relate to means of treatments for which there is no structure among the treatments. When there is structure, it should be taken into account. Failure to do this has resulted in misuses of multiple comparisons procedures. Examples of structure that should be utilized include sets of orthogonal contrasts or quantitative levels of a factor. Some of these are discussed by Chew (1977), Peterson (1977), Little (1978), Brian-Jones and Finney (1983) and Carmer and Walker (1982). Also, thought should be given to the goal of setting simultaneous confidence intervals and the goal of significance testing. In many experimental situations, it is known (almost certainly) that the overall null hypothesis of equality for a set of v treatments is false. Hence there is no point in significance testing since finding a significant difference among the v means is mainly a function of sample size. In this paper, we will concentrate on simultaneous confidence intervals for the set of contrasts under consideration.

Some of the more widely known and used multiple comparisons procedures are the following:

(i) lsd: The comparisonwise confidence interval on pairs of means is computed as $\bar{Y}_{i\cdot} - \bar{Y}_{i'\cdot} \pm t_{\alpha,f}\sqrt{2s^2/r}$, where s^2 is an estimate of the experimental error with f degrees of freedom, $\bar{Y}_{i\cdot}$ is the sample mean of treatment $i \neq i' = 1,2,\ldots,v$, r is the number of replicates for the ith treatment mean, and $t_{\alpha,f}$ is the tabular value of the two-tailed Students' t at the α level and for f degrees of freedom. This is the least significant difference procedure.

(ii) hsd: The experimentwise confidence interval on pairs of means is computed as $\bar{Y}_{i\cdot} - \bar{Y}_{i'\cdot} \pm q_{\alpha,v,f}\sqrt{s^2/r}$, where $q_{\alpha,v,f}$ is the tabulated value of the Studentized-range statistic at the α level for v treatments and for f degrees of freedom, and the other values are as defined in (i). This is commonly called the honestly significant difference or Tukey's range procedure.

(iii) esd: The per experiment confidence interval on m-pairs of means is computed as $\bar{Y}_{i\cdot} - \bar{Y}_{i'\cdot} \pm t_{\alpha/m,f}\sqrt{2s^2/r}$. This is sometimes called the Bonferroni procedure.

(iv) ssd: The experimentwise confidence interval for contrasts of the form $\sum_{i=1}^{v} c_i\mu_i$, where $\sum_{i=1}^{v} c_i = 0$, is computed as

$$\sum_{i=1}^{v} c_i \bar{Y}_{i\cdot} \pm \sqrt{(v-1)F_\alpha(v-1,f)s^2 \sum_{i=1}^{v} c_i^2/r}$$

, where $F_\alpha(v-1,f)$ is the tabulated value of F at the α level for v-1 degrees of freedom in the numerator and f in the denominator. This is known as Scheffe's procedure.

(v) csd: The experimentwise confidence interval for comparing a control or standard treatment mean with p other treatment means is computed as $\bar{Y}_{c\cdot} - \bar{Y}_{i\cdot} \pm d_{v,f,\alpha/2}\sqrt{2s^2/r}$, where $i = 1,2,\ldots,p$ and $d_{v,f,\alpha/2}$, is the two-sided α point of a p-variate t-distribution (p = v-1) with f degrees of freedom and common correlation $\rho = 1/2$. These values have been tabulated by Dunnett (1964) and are also given by others, e.g., Chew (1977) and Miller (1966).

Ranking and selection methods are more recent in origin than the other multiple comparison procedures and are considerably less known and used. The recently proposed confidence intervals by Hsu (1981) give a subset of the treatments that is guaranteed (with probability $1-\alpha$) to contain the population which has the largest mean and, at the same time, a set of simultaneous confidence intervals for the differences between the population means and the largest mean. Specifically, if we let $\mu_{[v]}$ denote the largest population mean then the procedure is:

(vi) Choose the i^{th} population to be in the selected subset if

$$\overline{Y}_{i\cdot} \geq \max_{i' \neq i} \overline{Y}_{i'\cdot} - d_{v,f,\alpha}\sqrt{\frac{2s^2}{r}} .$$

$d_{v,f,\alpha}$ is the one-sided α point of a (v-1)-variate t-distribution with f degrees of freedom and common correlation coefficient $\rho = 1/2$. These values have been tabulated by Dunnett (1955) and are also given by others, e.g., Miller (1966), Krishnaiah and Armitage (1966) or Gupta and Sobel (1957). (To use the tables in Dunnett or Miller, enter with $k = v-1$. The entries in Gupta and Sobel need to be divided by $\sqrt{2}$ before use. The tables in Gupta and Sobel or Krishnaiah and Armitage need to be entered with $\rho = 1/2$.) The simultaneous confidence intervals for $\mu_{[v]} - \mu_i$ are given by $[0,D_i]$, where

$$D_i = \max\{0, \max_{i' \neq i} \overline{Y}_{i'\cdot} - \overline{Y}_{i\cdot} + d_{v,f,\alpha}\sqrt{\frac{2s^2}{r}} .$$

Thus, Dunnett's one-sided comparisons with a control are related to simultaneous confidence intervals with the best in the following way. If the control treatment has the largest sample mean, the upper confidence intervals for the difference between the control mean and the other treatment means will be the same as those for $\mu_{[v]} - \mu_i$.

In the following we show how to apply these procedures to more complex designs of the split plot and split block nature.

2. SPLIT PLOT DESIGNS

There are many split plot designs (e.g. see Federer, 1975) but we shall confine our attention to one which has a whole plot treatments laid out in a randomized complete block design with r blocks. Then, each whole plot is divided into b split plots to which the b split plot treatments are randomly allocated within each whole plot. Thus, there are r randomizations for the whole plot treatments and ra randomizations for the split plot treatments. This is the design considered in most textbooks. Let Y_{hij} be the observation in block h (h = 1,2,...,r), on whole plot treatment i (i = 1,2,...,a), and on split plot treatment j (j = 1,2,...,b) and μ_{hij} be the expected value of Y_{hij}. Two analyses of variance (ANOVA) tables for this design are presented in Table 2.1. The first one is the standard one given in most textbooks that discuss an analysis of variance for a split plot design. The second one given in Table 2.1b is a useful one in many situations, especially where possibly different split plot treatments are to be recommended for each whole plot treatment. It is also appropriate when the split plot treatments differ from whole plot to whole plot; it should be used as a prior and/or supplementary ANOVA to the first one. Some of the sums of squares in the second ANOVA may be used to obtain those in the first. As examples, $T_{\cdot}$ is the total sum of squares with rab degrees of freedom, $E_{\cdot}$ is the error (b) sum of squares with a(r-1)(b-1) degrees of freedom, and $B_{\cdot}$ is the sum of the B and A × B sums of squares.

If the error rate base is comparisonwise, then the lsd procedure, (i), is appropriate. The procedure is straightforward using error (a), error (b), or possibly the individual $E_i/(r-1)(b-1)$ mean squares to compute the lsd. If, however, one wishes to use some sort of family error rate (see Tukey, 1953, and Miller, 1966), several procedures come to mind. These are:

a) Make comparisons among the a whole plot treatment means $\{\bar{\mu}_{\cdot i \cdot}-\bar{\mu}_{\cdot i' \cdot},\ i \neq i'\}$, and among the b split plot treatment means $\{\bar{\mu}_{\cdot\cdot j}-\bar{\mu}_{\cdot\cdot j'},\ j \neq j'\}$. These comparisons would be pertinent if there were no A × B interaction; the error terms are error (a) and error (b) from 2.1a.

b) Make comparisons among the a whole plot treatment means $\{\bar{\mu}_{\cdot i \cdot}-\bar{\mu}_{\cdot i' \cdot},\ i \neq i'\}$, and among the split treatment means separately for each whole plot treatment $\{\bar{\mu}_{\cdot ij}-\bar{\mu}_{\cdot ij'},\ j \neq j',\ i = 1,2,\ldots,a\}$. The error terms to be used are error (a) and error (b) for the whole plot treatments and split plot treatments respectively.

c) Make comparisons as in b) above, but use $E_i/(b-1)(r-1)$ within each whole plot treatment to compare the split plot treatment means. This would be appropriate if the variances of observations under different whole plot treatments are unequal. If possibly different split plot treatments are to be recommended for each whole plot treatment, the experimenter might wish to control the family error rate for each whole plot treatment.

d) Make comparisons among the b split plot treatment means $\{\bar{\mu}_{\cdot\cdot j} - \bar{\mu}_{\cdot\cdot j'}, j \neq j'\}$, and among the a whole plot treatment means separately for each split plot treatment, $\{\bar{\mu}_{\cdot ij} - \bar{\mu}_{\cdot i'j}, i \neq i', j = 1,2,\ldots,b\}$.

e) Make comparisons among the b split plot treatment means $\{\bar{\mu}_{\cdot ij} - \bar{\mu}_{\cdot ij'}, j \neq j', i = 1,2,\ldots,a\}$ separately for each of the whole plot treatments. Use error (b) as an error estimate.

To see how the standard procedures can be adapted, consider the following model for a split plot design:

$$Y_{hij} = \mu + \alpha_i + \rho_h + \delta_{hi} + \beta_j + (\alpha\beta)_{ij} + \epsilon_{hij}, \qquad (2.1)$$

where

Y_{hij} = observation in block h with whole plot treatment i and split plot treatment j,

μ = overall mean,

α_i = effect of whole plot treatment i,

ρ_h = effect of block h,

δ_{hi} = whole plot error term,

β_j = effect of split plot treatment j,

$(\alpha\beta)_{ij}$ = interaction of whole plot treatment i and split plot treatment j,

and

ϵ_{hij} = split plot error term.

We will assume $\delta_{hi} \sim \text{iid } N(0,\sigma_\delta^2)$ independently of $\epsilon_{hij} \sim \text{iid } N(0,\sigma_\epsilon^2)$ and that $\sum_i \alpha_i = \sum_j \beta_j = \sum_i (\alpha\beta)_{ij} = \sum_j (\alpha\beta)_{ij} = 0$. For situation c) we allow the possibility that the ϵ_{hij} have different variances, $\sigma^2_{\epsilon_i}$, for each whole plot treatment. Under model (2.1), comparisons among the whole plot means are based on

$$\hat{\alpha}_i - \hat{\alpha}_{i'} - (\alpha_i - \alpha_{i'}) = \bar{Y}_{\cdot i\cdot} - \bar{Y}_{\cdot i'\cdot} - (\alpha_i - \alpha_{i'})$$

$$= \bar{\delta}_{\cdot i} - \bar{\delta}_{\cdot i'} + (\bar{\epsilon}_{\cdot i\cdot} - \bar{\epsilon}_{\cdot i'\cdot}).$$

Table 2.1a

Standard ANOVA for the split plot design

Source of variation	Degrees of freedom	Sums of squares	Mean square
Total	rab	$\sum_{h=1}^{r}\sum_{i=1}^{a}\sum_{j=1}^{b} Y_{hij}^2$	-
Correction for mean	1	$\frac{Y_{\cdot\cdot\cdot}^2}{rab}$	-
Blocks = R	r-1	$\sum_{h=1}^{r}\frac{Y_{h\cdot\cdot}^2}{ab} - \frac{Y_{\cdot\cdot\cdot}^2}{rab}$	-
Whole plot treatments = A	a-1	$\sum_{i=1}^{a}\frac{Y_{\cdot i\cdot}^2}{rb} - \frac{Y_{\cdot\cdot\cdot}^2}{rab}$	MSA
R × A = error (a)	(r-1)(a-1)	$\sum_{h=1}^{r}\sum_{i=1}^{a}\frac{Y_{hi\cdot}^2}{b} - \sum_{i=1}^{a}\frac{Y_{\cdot i\cdot}^2}{rb} - \sum_{h=1}^{r}\frac{Y_{h\cdot\cdot}^2}{ab} + \frac{Y_{\cdot\cdot\cdot}^2}{rab}$	MSE(a)
Split plot treatments = B	b-1	$\sum_{j=1}^{b}\frac{Y_{\cdot\cdot j}^2}{ra} - \frac{Y_{\cdot\cdot\cdot}^2}{rab}$	MSB
A × B	(a-1)(b-1)	$\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{Y_{\cdot ij}^2}{r} - \sum_{i=1}^{a}\frac{Y_{\cdot i\cdot}^2}{rb} - \sum_{j=1}^{b}\frac{Y_{\cdot\cdot j}^2}{ra} + \frac{Y_{\cdot\cdot\cdot}^2}{rab}$	MSAB
B × R within A = error (b)	a(r-1)(b-1)	by subtraction	MSE(b)

Table 2.1b

Alternative ANOVA

Source of variation	Degrees of freedom	Whole plot (sums of squares) 1	2	.... a	Sum
Total	rb	T_1	T_2	T_a	T.
Correction for mean	1	C_1	C_2	C_a	C.
Blocks = R	r-1	R_1	R_2	R_a	R.
Split plot treatments = B	b-1	B_1	B_2	B_a	B.
R × B	(r-1)(b-1)	E_1	E_2	E_a	E.

Denoting $\mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$ by μ_{ij}, comparisons among the subplot means within whole plot i'' are based on

$$\hat{\mu}_{i''j} - \hat{\mu}_{i''j'} - (\mu_{i''j} - \mu_{i''j'}) = \overline{Y}_{\cdot i''j} - \overline{Y}_{\cdot i''j'} - (\mu_{i''j} - \mu_{i''j'})$$

$$= \overline{\epsilon}_{\cdot i''j} - \overline{\epsilon}_{\cdot i''j'} \;.$$

The covariance between these is

$$\mathrm{Cov}(\overline{\epsilon}_{\cdot i\cdot} - \overline{\epsilon}_{\cdot i'\cdot}, \overline{\epsilon}_{\cdot i''j} - \overline{\epsilon}_{\cdot i''j'}) \qquad (2.2)$$

since δ_{hi} and ϵ_{hij} are independent. (2.2) equals

$$\mathrm{Cov}(\overline{\epsilon}_{\cdot i\cdot}, \overline{\epsilon}_{\cdot i''j} - \overline{\epsilon}_{\cdot i''j'}) - \mathrm{Cov}(\overline{\epsilon}_{\cdot i'\cdot}, \overline{\epsilon}_{\cdot i''j} - \overline{\epsilon}_{\cdot i''j'})$$

which is zero since $\mathrm{Cov}(\overline{\epsilon}_{\cdot i\cdot}, \overline{\epsilon}_{\cdot i''j})$ is independent of j. Thus, inferences about the whole plot means and the split plot means are independent. The same is true for error estimates based on them (MSE(a) and MSE(b)). So, for example, considering situation (a)

$$P\{\left|\hat{\alpha}_i - \hat{\alpha}_{i'} - (\alpha_i - \alpha_{i'})\right| \le q_{[\alpha_1, a, (r-1)(a-1)]}\sqrt{\frac{\mathrm{MSE(a)}}{rb}} \quad i,i'$$

and

$$\left|\hat{\beta}_j - \hat{\beta}_{j'} - (\beta_j - \beta_{j'})\right| \le q_{[\alpha_2, b, a(r-1)(b-1)]}\sqrt{\frac{\mathrm{MSE(b)}}{ra}} \quad j,j'\}$$

$$= (1 - \alpha_1)(1 - \alpha_2).$$

Also, inferences among subplots within a particular whole plot are independent of those in a different whole plot.

We illustrate some of these situations with an example from Federer (1955). The means and two ANOVA's are presented in Table 2.2. The data are number of seeds germinating out of 100. Several transformations of the data were attempted but none affected the analysis appreciably so the analysis on the original scale is presented. In Federer (1955), single degree of freedom contrasts were presented for comparisons among the split plot treatment means, seed treatments, and for partitioning the seed treatment x variety (whole plot) interaction. If any three contrasts among seed treatment means were the only ones of interest, one could use the esd procedure (iii) described above. For the three contrasts one could use $t_{\alpha/3,f}\sqrt{\frac{2\ \text{error (b)}}{ra}} = t_{.05/3,48}\sqrt{2(24.23)/8(3)}$ for $\alpha = .05$. If one wished to make these three contrasts within each whole plot treatment, there would be $3a = 3(8) = 24 = m$ contrasts. In this case use $t_{\alpha/24,48}\sqrt{2(24.23)/8(3)}$. If p contrasts among the varietal means were desired, one could add these to the preceding to obtain $m + p$ contrasts or one could consider m and p separately.

Table 2.2a

Variety and seed treatment means

Seed treatment	Variety a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	Means
b_1	66	63	65	50	49	58	46	48	56
b_2	12	18	13	10	16	8	15	20	14
b_3	12	14	26	14	10	30	22	31	20
b_4	8	11	11	10	12	17	10	12	11
Means	25	27	29	21	22	28	23	28	25

We will illustrate the hsd and esd procedures and Hsu's simultaneous confidence intervals with the best for situations a) and c). For situation a) we take the error rate for whole plot comparisons and the error rate for split plot comparisons to be .05 so that the overall error is $.95^2 \doteq .90$. The hsd procedure uses $q_{.05,8,14}\sqrt{\frac{\text{MSE(a)}}{3(4)}} = 4.99\sqrt{\frac{98.38}{12}} = 14.29$ for whole plot

Table 2.2b.

Classical textbook ANOVA

Source of variation	Degrees of freedom	Sum of squares	Mean square
Total	96	98,195.00	-
Correction for mean	1	61,458.76	-
Blocks	2	38.58	19.29
Varieties	7	763.16	109.02
Error (a)	14	1,377.25	98.38
Seed treatments	3	30,774.28	10,258.09
Seed treatments x varieties	21	2,620.13	124.77
Error (b)	48	1,162.84	24.23

comparisons and $q_{.05,4,48}\sqrt{\frac{MSE(b)}{3(8)}} = 3.77\sqrt{\frac{24.23}{24}} = 3.79$ for the split plot comparisons. The esd procedure uses $t_{.05/28,14}\sqrt{\frac{2(MSE(a))}{3(4)}} =$ $3.85\sqrt{\frac{2(98.38)}{12}} = 15.57$ for whole plot comparisons and $t_{.05/6,48}\sqrt{\frac{2(MSE(b))}{3(8)}} = 2.76\sqrt{\frac{2(24.23)}{24}} = 3.92$ for split plot comparisons Since the hsd always gives smaller intervals when comparing all pairs of means, it is to be preferred in this situation. Tables 2.3a and 2.3b list the means, differences between the means and the hsd length to be used for confidence intervals:

$$\bar{Y}_{\cdot i \cdot} - \bar{Y}_{\cdot i' \cdot} \pm 14.3$$

$$\text{(for } \alpha_i - \alpha_{i'}\text{)}$$

and

$$\bar{Y}_{\cdot\cdot j} - \bar{Y}_{\cdot\cdot j'} \pm 3.8$$

$$\text{(for } \beta_j - \beta_{j'}\text{)}.$$

Calculating simultaneous confidence intervals with the best, we need $d_{v,f,\alpha}\sqrt{\frac{2(MSE)}{r}}$. For the whole plot treatments, we use $d_{8,14,.05}=2.59$

Table 2.2c

ANOVA as suggested in Table 2.1b

Whole plot treatment (variety)

Sum of squares

Source of variation	Degrees of freedom	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	Sum
Total	12	14326	14956	16184	9112	8889	13972	9107	11649	98195
Correction for mean	1	7301.3	8640.3	9976.3	5292.0	5764.1	9520.3	6486.8	9240.8	62221.9
Blocks = R	2	2.2	763.2	338.2	8.0	4.2	83.2	120.5	96.5	1416
Treatments = B	3	6972.0	5403.0	5645.7	3476.7	3066.2	4271.0	2350.9	2208.9	33394.4
R × B	6	50.5	149.5	223.8	335.3	54.5	97.5	148.8	102.8	1162.7

and for split plots we use $d_{4,48,.05}=2.11$. Thus the selected subsets which contain the population with the largest mean (with confidence coefficient .90) are

$$\{a_1,a_2,a_3,a_4,a_5,a_6,a_7,a_8\} \quad \text{for the whole plots}$$

and

$$\{b_1\} \quad \text{for the split plots.}$$

Table 2.4 lists the simultaneous confidence intervals from the best calculated as $[0,D_i^{(\alpha)}]$ for the whole plots, where

$$D_i^{(\alpha)} = \max\{0, \max_{i'\neq i} \bar{Y}_{\cdot i'\cdot} - \bar{Y}_{\cdot i\cdot} + 10.5\}$$

and as $[0,D_j^{(\beta)}]$ for the split plots, where

$$D_j^{(\beta)} = \max\{0, \max_{j'\neq j} \bar{Y}_{\cdot\cdot j'} - \bar{Y}_{\cdot\cdot j} + 3.0\}.$$

Table 2.3a

Whole plot hsd comparisons for the data of Table 2.2 for situation (a). The body of the table gives pairwise mean differences. (hsd = 14.3 for confidence coefficient .90)

		Variety						
		a_3	a_6	a_8	a_2	a_1	a_7	a_5
Variety	Mean	29	28	28	27	25	23	22
a_4	21	8	7	7	6	4	2	1
a_5	22	7	6	6	5	3	1	
a_7	23	6	5	5	4	2		
a_1	25	4	3	3	2			
a_2	27	2	1	1				
a_8	28	1	0					
a_6	28	1						

Notice that the "yardstick" used for confidence intervals from the best are much shorter than the hsd (10.5 compared to 14.3 and 3.0 compared to 3.8), reflecting the smaller number of comparisons made.

Table 2.3b

Split plot hsd comparisons for the data of Table 2.3b for situation (a). The body of the table gives pairwise mean differences. (hsd = 3.8 for confidence coefficient .90)

		Seed Treatment		
		b_1	b_3	b_2
Seed Treatment	Mean	56	20	14
b_4	11	45	9	3
b_2	14	42	6	
b_3	20	36		

Thus confidence intervals from the best is clearly advantageous when the goal is to make statements about which treatments have the largest means.

Table 2.4

Simultaneous confidence intervals from the best for the data of Table 2.2 using situation a) (confidence coefficient .90)

confidence interval	parameter
[0,14.5]	$\alpha_{[8]} - \alpha_1$
[0,12.5]	$\alpha_{[8]} - \alpha_2$
[0, 9.5]	$\alpha_{[8]} - \alpha_3$
[0,18.5]	$\alpha_{[8]} - \alpha_4$
[0,17.5]	$\alpha_{[8]} - \alpha_5$
[0,11.5]	$\alpha_{[8]} - \alpha_6$
[0,16.5]	$\alpha_{[8]} - \alpha_7$
[0,11.5]	$\alpha_{[8]} - \alpha_8$
[0,0]	$\beta_{[4]} - \beta_1$
[0,45]	$\beta_{[4]} - \beta_2$
[0,39]	$\beta_{[4]} - \beta_3$
[0,48]	$\beta_{[4]} - \beta_4$

Next we consider situation c). We will choose an error rate of .05 for the whole plot comparisons and each of the within-whole-plot comparisons. This will give an overall error rate of $.95^9 \doteq .63$. The whole plot comparisons are the same as above. We will illustrate the calculations for the comparisons within whole plot treatment a_6. The hsd procedure uses $q_{.05,4,6}\sqrt{\frac{E_6/6}{3}} = 4.90\sqrt{\frac{97.5/6}{3}} = 11.39$ and the confidence intervals from the best uses $d_{4,6,.05}\sqrt{\frac{2(E_6/6)}{3}} = 2.56\sqrt{\frac{2(97.5/6)}{3}} = 8.43$. Thus, for the hsd, the intervals are

$$\overline{Y}_{.6j} - \overline{Y}_{.6j'} \pm 11.39$$

and the intervals from the best are $[0, \max\{0, \max_{j' \neq j} \overline{Y}_{.6j'} - \overline{Y}_{.6j} + 8.43\}]$ which are [0,0], [0,58.4], [0,36.4], and [0,49.4] for b_1, b_2, b_3 and b_4 respectively. Doing the calculations for the other whole plot treatments shows that the split plot treatment b_1 is the best within each of the whole plot treatments. This, as well as other statements, e.g., that all whole plot treatments are within 18.5 of the best, can be made with confidence coefficient .63.

3. EXTENSIONS

In a split block design, there are two sets of whole plot treatments with each set having a different error variance. For our purpose, let us consider the split block design for which each set of whole plots is in a randomized complete block design, i.e. there would be r randomizations for each set of whole plots (e.g. see Federer, 1955 and 1975). Thus, within a block, each plot in each of the two sets of whole plots is divided into subplots by the plots of the other set. An analysis of variance table for this situation is given in Table 3.1. There are three separate error mean squares, MSE(a), MSE(b), and MSE(ab), for this design. Resulting F ratios are MSA/MSE(a), MSB/MSE(b), and MSAB/MSE(ab). In setting up an "experimentwise" error rate, one could consider several situations, one of which is

f) Consider comparisons among each of the two sets of whole plot treatments. These will be independent (with different error terms) and could be handled as in Section 2.

Table 3.1

Analysis of variance table for a split block design.

Source of Variation	Degrees of freedom	Sum of squares	Mean square
Total	rab	$\sum_{h=1}^{r} \sum_{i=1}^{a} \sum_{j=1}^{b} Y_{hi}^2$	-
Correction for mean	1	$\frac{Y_{\cdot\cdot\cdot}^2}{rab}$	-
Blocks	r-1	$\sum \frac{Y_{h\cdot\cdot}^2}{ab} - \frac{Y_{\cdot\cdot\cdot}^2}{rab}$	-
One set of whole plots = A	a-1	$\sum \frac{Y_{\cdot i\cdot}^2}{rb} - \frac{Y_{\cdot\cdot\cdot}^2}{rab}$	MSA
Blocks × A	(r-1)(a-1)	subtraction	MSE(a)
Second set of whole plots = B	b-1	$\sum \frac{Y_{\cdot\cdot j}^2}{ra} - \frac{Y_{\cdot\cdot\cdot}^2}{rab}$	MSB
Block × B	(r-1)(b-1)	subtraction	MSE(b)
A × B	(a-1)(b-1)	$\sum\sum \frac{Y_{\cdot ij}^2}{r} \sum \frac{Y_{\cdot i\cdot}^2}{rb} - \sum \frac{Y_{\cdot\cdot j}^2}{ra} + \frac{Y_{\cdot\cdot\cdot}^2}{rab}$	MSAB
Blocks × A × B	(r-1)(a-1)(b-1)	subtraction	MSE(ab)

For a split-split plot design let us consider the design in Section 2 for which the split plot experimental units are split into c split-split plot experimental units (sspeu) and the c split-split plot treatments are randomly allocated to these c sspeus within each of the rab split plot experimental units resulting in rab randomizations for split-split plot treatments. An analysis of variance table for this design may be found in several places, one of which is Table X-8 of Federer (1955). There are nine separate variances for contrasts listed as formulae (X-11) through (X-19) in Federer (1955). These involve three error mean squares E_a, E_b and E_c. One could

g) Consider comparisons among whole plot treatments, among split plot treatments and among split-split plot treatments. Again, these are independent (with different error variances) and can be treated as in Section 2.

A multitude of other possibilities are conceivable.

4. DISCUSSION

We have illustrated three techniques for multiple comparisons, each with its advantages. When desiring statements about the best treatment, an investigator should use the procedure of Hsu (1981, 1982) which yields shorter intervals for the comparisons of interest. The hsd is a reasonable procedure for making comparisons among all means and is applicable for some situations in the multiple-error-term designs we considered here. The esd always gives intervals which are wider than the hsd when making comparisons among all means. However, it is extremely versatile and can be used easily even in the most complex situations.

SUMMARY

Three multiple comparisons procedures, least significant difference, Studentized range, and Bonferroni, and simultaneous confidence intervals containing the best population are demonstrated for comparing means from split plot designs. A numerical example is used to illustrate two of the five situations for a split plot design.

BIBLIOGRAPHY

Bechhofer, R.E, (1954). A single-sample multiple decision procedure for ranking means of normal populations with known variances. Ann. Math. Statist. 25, 16-39.

Bechhofer, R.E., Kiefer, J. and Sobel, M. (1968). Sequential identification and ranking procedures. University Chicago Press, Chicago.

Bryan-Jones, J., and Finney, D.J. (1983). On an error in "Instructions to Authors." HortScience 18, 279-282.

Carmer, S.G. and Walker, W.M. (1982). Baby bear's dilemma: a statistical tale. Agron. J. 74(1), 122-124.

Chew, V. (1977). Comparisons among treatment means in an analysis of variance. USDA, ARS/H/6.

Duncan, D.B. (1947). Significance tests for differences between ranked variates drawn from normal populations. Ph.D. thesis, Iowa State University.

Duncan, D.B. (1955). Multiple range and multiple F tests. Biometrics 11, 1-42.

Dunnett, C.W. (1955). A multiple comparisons procedure for comparing several treatments with a control. J. Amer. Statist. Assoc. 50, 1096-1121.

Dunnett, C.W. (1964). New tables for multiple comparisons with a control. Biometrics 20, 482-491.

Federer, W.T. (1955). Experimental Design-Theory and Application. MacMillan, NY (Republished by Oxford and IBH Company, New Delhi 1967, 1974), Chapter II.

Federer, W.T. (1975). The misunderstood split plot. In Applied Statistics (Ed. R.P. Gupta) North Holland Publishing Co. Amsterdam.

Gibbons, J.D., Olkin, I. and Sobel, M. (1977). Selecting and Ordering Populations: a new statistical methodology. Wiley: New York.

Gupta, S.S. (1965). On some multiple decision (selection and ranking) rules. Technometrics 6, 225-245.

Gupta, S.S. and Sobel, M. (1957). On a statistic which arises in selection and ranking problems. Ann. Math. Statist. 28, 957-967.

Gupta, S.S. and Sobel, M. (1958). On Selecting a subset which contains all populations better than a standard. Ann. Math Statist. 29, 235-244.

Hsu, J.C. (1981). Simultaneous confidence intervals for all distances from the "best". Ann. Statist. 9, 1026-1034.

Hsu, J.C. (1982). Simultaneous inference with respect to the best in block designs. J. Amer. Statist. Assoc. 77, 461-467.

Keuls, M. (1952). The use of the "studentized range" in connection with the analysis of variance. Euphytica 1, 112-122.

Krishnaiah, P.R. and Armitage, J.V. (1966). Tables for the multivariate t distribution. Sankhya, Ser.B. 28, 31-56.

Little, T.M. (1978). If Galileo published in HortScience. HortScience 13, 504-506.

Miller, R.G. Jr. (1966). Simultaneous Statistical Inference. McGraw-Hill, New York.

Miller, R.G. Jr. (1977). Developments in multiple comparisons 1966-1976. J. Amer. Statist. Assoc., 72, 779-788.

Peterson, R.G. (1977). Use and misuse of multiple comparisons procedures. Agron. J. 69, 205-208.

Paulson, E. (1964). A sequential procedure for selecting a population with the largest mean from k normal populations. Ann. Math. Statist. 35, 174-180.

Scott, A.J. and M. Knott (1974). A cluster analysis method for grouping means in an analysis of variance. Biometrics 30, 507-512.

Steel, R.G.D. and Federer, W.J. (1955). Yield-stand analyses. J. Indian Soc. Agri. Statist. VII (1&2), 27-45.

Tukey, J.W. (1953). The problem of multiple comparisons. Unpublished report in private circulation.

Waller, R.A. and D.B. Duncan (1969). A Bayes rule for the symmetric multiple comparisons problem. J. Amer. Statist. Assoc. 64, 1484-1503.