

Multiple Comparisons in Split Block and Split-Split Plot Designs*

by

Walter T. Federer and Charles E. McCulloch

BU-1132-MA

November 1991

Abstract

Multiple comparisons in split block and split-split plot designs encounter statistical, conceptual, and philosophical difficulties. Multiple error terms and distributional properties of some of the error terms account for the statistical problems. Structure or lack of it in treatment design is what causes the conceptual and philosophical difficulties. When there is structure in the treatment design, specific contrasts in the form of orthogonal single degree of freedom, or sets of degrees of freedom, comparisons are required. When the different entries of a factor have a nominal scale of measurement and no structure, multiple comparisons procedures usually will be indicated. The error rate base will need to be addressed regardless of the nature and structure of the treatment design. Selection of the error rate base is the important consideration. Then, the selection the particular multiple comparisons procedure to be used follows. The term multiple comparisons needs to be precisely defined. The above situations are addressed for these more complex designs.

* In the Technical Report Series of the Biometrics Unit, Cornell University, 337 Warren Hall, Ithaca, New York 14853

1. Introduction

This topic was selected for presentation at this Conference because of the contributions of Charles E. Dunnett in this area, especially his papers dealing with experimentwise error rates for comparisons with a control. The paper is a sequel to the ones by Federer (1975), Federer and McCulloch (1984) and Federer and Meredith (1991) on analyses for these designs. In this paper, we shall consider the complete set of all possible comparisons or some subset thereof as falling in the realm of multiple comparisons. Some authors appear to consider multiple comparisons to be confined to all possible pairwise comparisons of means, some as all comparisons with a control, some as all possible comparisons, and some as a selected subset of all possible comparisons. We consider all these situations to be in the realm of multiple comparisons.

In many situations involving the use of split block and split plot designs, there is considerable structure associated with the treatment design. Hence the problem of considering all possible comparisons does not arise even though the problem of selecting the error rate will remain. As Tukey (1991) has so aptly pointed out in his thoughtful article, multiple comparisons are not always appropriate, especially when there is structure in the treatment design. When structure is absent, multiple comparisons are appropriate. One particular case in which multiple comparisons will be appropriate for a split-split plot design is when the main or whole plot treatments are populations (grandparents, species, hospitals, e.g.), the split plot treatments are subpopulations (parents, crosses or families, procedures), and the split-split plot treatments are individuals within a subpopulation (children, lines or selections, technicians). Many such situations can be found in practice for the simple split plot design. A variety of multiple comparisons may be desired among the whole plot treatments, among split plot treatments, and/or among the split-split plot treatments. In any event, every experiment involves a set of comparisons and the need for selecting an error rate base.

In the next section, a particular member of the family of split block designs is selected to illustrate the procedures for multiple comparisons. The first situation considered is for the case when there is considerable structure among the levels of the two factors which form the two-way whole plots.

Then, the situation for which all possible comparisons among pairs of means is desired, is presented. Multiple error terms cause complications for both cases and the differing sums of squares of coefficients for the different contrasts necessitates the computation of a different range for each contrast.

In the third section a third situation is considered for split block design where there is no structure among the levels of the two factors. A simulation-based approach is illustrated.

The subject of multiple comparisons for a split-split plot design is discussed in the fourth section. The particular split-split plot design selected to illustrate multiple comparisons procedures is the one which has populations as the whole or main plots, subpopulations as the split plots, and individuals within subpopulations as the split-split plots.

Three particular and different error rate bases are to be used in the following. These are comparisonwise or per comparison, per experiment, and experimentwise. There are more bases (See, e.g., Tukey, 1953, 1991, Hartley, 1955, Federer, 1961, Chew, 1977, Saville, 1990, 1991, Holland, 1991, and Lea, 1991, for a discussion of error rates and multiple comparisons procedures.), but our attention is confined to these three. An error rate base needs to be selected prior to selecting a multiple comparisons procedure. This means that it will not be possible to use only one multiple comparisons procedure for all situations as advocated by some. (See, e.g., Saville, 1990, 1991.) A per comparison or comparisonwise error rate is defined to be

$$\begin{aligned} \text{Error rate per comparison} &= \frac{[\text{expected number of erroneous inferences of comparisons}]}{[\text{number of inferences on comparisons attempted}]} \\ &= \text{proportion of all comparisons expected to be erroneous when the null hypothesis is true.} \end{aligned}$$

A per experiment error rate is said to be

$$\begin{aligned} \text{error rate per experiment} &= \frac{[\text{expected number of erroneous inferences}]}{[\text{number of experiments}]} \\ &= \text{the expected number of erroneous statements per experiment when the null hypothesis is true} \end{aligned}$$

An experimentwise error rate is defined to be

$$\begin{aligned} \text{Experimentwise error rate} &= \frac{[\text{expected number of experiments with one or more erroneous statements}]}{[\text{number of experiments}]} \\ &= \text{expected proportion of experiments with one or more erroneous inferences when the null hypothesis is true.} \end{aligned}$$

Following are four multiple comparisons procedures we consider which, for illustrative purposes, are described only for comparing means in the one-way layout.

lsd: The comparisonwise confidence interval for a pair of means is computed as $\bar{y}_i - \bar{y}_{i'} \pm t_{\alpha, f} (2s^2/r)^{1/2}$, where s^2 is an estimate of the experimental error mean square with f degrees of freedom, \bar{y}_i is the sample treatment mean $i \neq i' = 1, 2, \dots, v$, r is the number of replicates for the i^{th} treatment mean, and $t_{\alpha, f}$ is the tabulated value for the two-tailed Student's t at the α percent level for f degrees of freedom. This is the least significance difference procedure

esd: The per experiment confidence interval for m pairs of means is computed as $\bar{y}_i - \bar{y}_{i'} \pm t_{\alpha/m, f} (2s^2/r)^{1/2}$. This is sometimes called the Bonferroni procedure.

hsd: The experimentwise confidence interval on pairs of means is computed as $\bar{y}_i - \bar{y}_{i'} \pm q_{\alpha, f, v} (s^2/r)^{1/2}$, where $q_{\alpha, f, v}$ is the tabulated value of the Studentized-range statistic at the α percent level for f degrees of freedom and v treatments in the experiment. This is commonly called the honestly significant difference or Tukey's range procedure.

multiple comparisons with the best and subset selection: Choose the i^{th} population to be in the selected subset if

$$\bar{y}_i \geq \max_{i' \neq i} \{\bar{y}_{i'}\} - d_{v, f, \alpha} (2s^2/r)^{1/2},$$

where $d_{v, f, \alpha}$ is the one-sided α point of a $(v-1)$ -variate t -distribution with f degrees of freedom and common correlation $\rho = 1/2$. These values have been tabulated by Dunnett (1955), for example, where the tables are entered using $k = v - 1$. Simultaneous intervals which correspond to the subset selection procedure confidence intervals for $\mu_{[v]} - \mu_i$ are given by $[0, D_i]$, where

$$D_i = \max \{0, (\max_{i' \neq i} \{\bar{y}_{i'}\} - \bar{y}_i + d_{v, f, \alpha} (2s^2/r)^{1/2})\}$$

and $\mu_{[i]}$, $i = 1, 2, \dots, v$, represent the ordered population means ($\mu_{[1]} \leq \mu_{[2]} \leq \dots \leq \mu_{[v]}$) (Hsu, 1981). Dunnett's one-sided comparisons with a control are related to simultaneous confidence intervals with the best. If the control treatment has the largest sample mean, the upper one-sided confidence limits for the differences between the control mean and the other treatment means will be the same as those for $\mu_{[v]} - \mu_i$. The error rate is experimentwise.

2. Split Block Design

The particular split block, or two-way whole plot, design we use as an example will have the levels of one factor, say A, laid out in a latin square design with the rows corresponding to complete blocks and the columns corresponding to orders within the complete blocks. The experimental units for the levels of the second factor, say B, are laid out across the experimental units for the levels of factor A and are designed as a randomized complete block design. Let r represent the number of blocks for A and let a be the number of levels of A, then for $r = 4 = a$, a schematic plan would appear like this:

Block (row)		Order (column)			
		1	2	3	4
1	B ₁	A ₁	A ₂	A ₃	A ₄
	B ₂				
	B ₃				
2	B ₃	A ₂	A ₃	A ₄	A ₁
	B ₁				
	B ₂				
3	B ₁	A ₃	A ₄	A ₁	A ₂
	B ₂				
	B ₃				
4	B ₂	A ₄	A ₁	A ₂	A ₃
	B ₁				
	B ₃				

There are $a = 4$ experimental units for the levels of factor A and $b = 3$ experimental units for levels of factor B in each block. There will be r randomizations for the a levels of factor A, restricted to form a latin square, and another r randomizations (unrestricted) for the b levels of factor B. The experimental unit for the combination of level i of factor A and level j of factor B is $1/ab$ of the block size. The fact that there are three levels of randomization and three different sizes of experimental units leads to the consideration of three different error mean squares. A linear model for the above designed experiment would be

$$Y_{ghij} = \mu + \rho_g + \gamma_h + \tau_i + \pi_j + \delta_{ghi} + (\tau\pi)_{ij} + \lambda_{gj} + \epsilon_{ghij}, \quad (1)$$

where Y_{ghij} is the response for the $ghij^{th}$ observation, μ is a mean effect common to all observations, ρ_g is the g^{th} row effect, γ_h is the h^{th} column effect, τ_i is the effect of the i^{th} level of factor A, π_j is the effect of the j^{th} level of factor B, δ_{ghi} is a random error effect for the subdivision within row g and column h associated with the levels of factor A, λ_{gj} is a random error effect for the subdivision within

row g associated with the levels of factor B, $(\tau\pi)_{ij}$ is an interaction effect of the i^{th} level of factor A and the j^{th} level of factor B, and ϵ_{ghij} is a random error effect associated with the $ghij^{\text{th}}$ observation. Given that there is sufficient structure in the treatment design to construct orthogonal single degree of freedom contrasts, a partitioning of the factor and interaction degrees of freedom may be made as given in Table 2.1. The multiple error mean squares is a feature of the experiment design and not of multiple comparisons procedures.

Given the $ab - 1$ single degree of freedom factor interaction contrasts in Table 2.1, the error rate base for constructing confidence intervals (or making tests of significance) could be per contrast (comparisonwise), per the $ab - 1$ set of contrasts (per experiment error rate base), or some other base. The problem of determining an error rate is a feature of multiple comparisons and not of the statistical design of either the treatments or the experiment.

For the $a - 1$ contrasts A_1 to A_{a-1} for factor A, we use E_a to construct the $a - 1$ confidence intervals. For the $b - 1$ contrasts B_1 to B_{b-1} for factor B, use is made of E_b to construct the $b - 1$ confidence intervals. For the $(a - 1)(b - 1)$ contrasts $A_1 \times B_1$ to $A_{a-1} \times B_{b-1}$ for the interaction, we make use of E_{ab} to construct the $(a - 1)(b - 1)$ confidence intervals. For a comparisonwise error rate, use

$$t_{f_a, \alpha} (E_a \sum_i c_i^2 / rb)^{1/2},$$

$$t_{f_b, \alpha} (E_b \sum_j d_j^2 / ra)^{1/2}, \text{ and}$$

$$t_{f_{ab}, \alpha} (E_{ab} \sum_i \sum_j \{c_i d_j\}^2 / r)^{1/2},$$

for the A, B, and $A \times B$ contrasts, respectively. In the above, c_i represent the coefficients in one of the contrasts among the means for levels of factor A, i.e.,

$$\sum_i c_i \bar{y}_i.$$

d_j are the coefficients in a contrast among levels of factor B. For a per experiment error rate, we replace α with $\alpha / (ab - 1)$. According to Ghosh (1955), this is what would be done for an experimentwise error rate for this set of contrasts. A more efficient approach would be to use percentage points of the multivariate t-distribution, though tables of those percentage points are less accessible.

Table 2.1. An analysis of variance for a split block design.

Source of variation	Degrees of freedom	Mean square
Total	rab	
Correction for mean	1	
Block (row)	$r - 1$	
Column	$r - 1$	
Factor A	$a - 1 = r - 1$	
Contrast A_1	1	
Contrast A_2	1	
\vdots	\vdots	
Contrast A_{a-1}	1	
Error (a)	$f_a = (r - 1)(r - 2)$	E_a
Factor B	$b - 1$	
Contrast B_1	1	
Contrast B_2	1	
\vdots	\vdots	
Contrast B_{b-1}	1	
Error (b)	$f_b = (r - 1)(b - 1)$	E_b
A \times B	$(a - 1)(b - 1)$	
$A_1 \times B_1$	1	
$A_1 \times B_2$	1	
\vdots	\vdots	
$A_2 \times B_1$	1	
$A_2 \times B_2$	1	
\vdots	\vdots	
$A_{a-1} \times B_{b-1}$	1	
Error (ab)	$f_{ab} = (r - 1)(a - 1)(b - 1)$	E_{ab}

If multiple comparisons were made among pairs of the ab means $\bar{y} \dots ij'$, five different variances of a difference between two means would be involved (See Federer and Meredith, 1991). These are

$$V(\bar{y} \dots i. - \bar{y} \dots i'.) = 2E_a/rb ,$$

$$V(\bar{y} \dots .j - \bar{y} \dots .j'.) = 2E_b/ra ,$$

$$V(\bar{y} \dots ij - \bar{y} \dots ij'.) = 2[E_a + (b-1)E_{ab}]/rb ,$$

$$V(\bar{y} \dots .ij - \bar{y} \dots .ij'.) = 2[E_b + (a-1)E_{ab}]/ra , \text{ and}$$

$$V(\bar{y} \dots .ij - \bar{y} \dots .ij'.) = 2[aE_a + bE_b(ab - a - b)E_{ab}]/rab .$$

Note that the five different standard errors is a feature of the experiment design and not of multiple comparisons procedures. These multiple variances will result in five different ranges for comparing pairs of means. The number of degrees of freedom, f_{f^*} , for the last three mean squares must be approximated. (See Grimes and Federer, 1984, and also the next section for an alternate approach.) The resulting comparisonwise error rate and the per experiment error rate $(1 - \alpha)\%$ confidence intervals are computed as

comparisonwise:

$$\text{A mean: } \bar{y} \dots i. - \bar{y} \dots i'. \pm t_{f_a, \alpha} (2E_a/rb)^{1/2}$$

$$\text{B mean: } \bar{y} \dots .j - \bar{y} \dots .j'. \pm t_{f_b, \alpha} (2E_b/ra)^{1/2}$$

$$\text{AB mean: } \bar{y} \dots ij - \bar{y} \dots ij'. \pm t_{f^*, \alpha} (2[E_a + \{b-1\}E_{ab}]/rb)^{1/2}$$

$$\bar{y} \dots .ij - \bar{y} \dots .ij'. \pm t_{f^*, \alpha} (2[E_b + \{a-1\}E_{ab}]/rb)^{1/2}$$

$$\bar{y} \dots .ij - \bar{y} \dots .ij'. \pm t_{f^*, \alpha} (2[aE_a + bE_b + \{ab - a - b\}E_{ab}]/rab)^{1/2}$$

per experiment: In the above confidence interval formulae, replace the value for α by

$$\alpha/[a(a-1)/2 + b(b-1)/2 + ab(ab-1)/2] ,$$

where the denominator is the total number of pairwise comparisons of means made in comparing the A, the B, and the $A \times B$ means.

3. Simulation-based Methods for the Split-Block Design

In complicated experiment designs, such as the split-block or split-split plot experiments illustrated here, it is quite likely that the experimenter will wish to control the error rate for a subset of

all the possible comparisons. In such cases the lsd procedure is not applicable, since it only controls the comparisonwise error rate. The esd procedure may be easily adapted by appropriately adjusting the α level, but can be overly conservative (Edwards and Berry, 1987). Given the variety of error terms and correlations between contrasts that are possible, a simulation approach similar to that used by Edwards and Berry (1987) seems the most feasible. We sketch such an approach for the split-block experiment and model (1) using the hsd and subset selection approaches. We assume that the experimenter is interested in simultaneously controlling the error rate for both comparisons among A and B. This would be appropriate, for example, when A and B are applied together and our inferences depend on the combination of the two.

For the hsd approach let S_A be the set of indices, (i, i') , representing all of the comparisons of interest among levels of A. That is, $(i, i') \in S_A$ implies we wish to make inferences about $\tau_i - \tau_{i'}$. Let S_B have a similar definition for factor B. Simultaneous confidence intervals for all the comparisons in S_A and S_B would be of the form:

$$\bar{y} \dots i. - \bar{y} \dots i'. \pm d \sqrt{\frac{2E_a}{rb}} \text{ for all } (i, i') \in S_A$$

and

$$\bar{y} \dots j - \bar{y} \dots j' \pm d \sqrt{\frac{2E_b}{ra}} \text{ for all } (j, j') \in S_B.$$

The problem is to determine d . This can be achieved by a simulation approach. Straightforward calculations show that $\bar{y} \dots i. - \bar{y} \dots i'.$ and $\bar{y} \dots j - \bar{y} \dots j'$ are uncorrelated. An algorithm follows.

1. Generate a set of normal variates, $Z_{ii'}$, with the same variance-covariance structure as

$$\frac{\bar{y} \dots i. - \bar{y} \dots i'.}{\sqrt{\frac{2\sigma_a^2}{rb}}} \quad (i, i') \in S_A.$$

2. Generate an independent (of $Z_{ii'}$) set of normal variates, $X_{jj'}$, with the same variance-covariance structure as

$$\frac{\bar{y} \dots j - \bar{y} \dots j'.}{\sqrt{\frac{2\sigma_b^2}{ra}}} \quad (j, j') \in S_B.$$

3. Generate a χ^2 variate, U , with f_a degrees of freedom.
4. Generate a χ^2 variate, V , with f_b degrees of freedom.

5. Calculate $W = \max \left\{ \max_{(i, i') \in S_A} \frac{Z_{ii'}}{\sqrt{U/f_a}}, \max_{(j, j') \in S_B} \frac{X_{jj'}}{\sqrt{V/f_a}} \right\}$
6. Repeat steps 1 – 5 N times (a large number).
7. Order the values in 6 from smallest to largest and use the αN largest one as d .

This approach is fully efficient in that it incorporates the exact number of comparisons used and the exact correlation structure.

For the subset selection approach suppose we wish to derive simultaneous subsets and multiple comparisons with the best for both τ_i (factor A) and π_j (factor B). This would be appropriate if we simultaneously needed the level of A and the level of B to be within a certain distance of the “best”. Since contrasts among the factor A means are independent of the factor B means (assuming normality) inferences can straightforwardly be made by using simple independence arguments. We will, however, illustrate the simulation approach to show its flexibility. The i^{th} level of factor A is in the first selected subset if

$$\bar{y} \dots i \dots \geq \max_{i' \neq i} \{ \bar{y} \dots i' \dots \} - d \left(\frac{2E_a}{rb} \right)^{\frac{1}{2}}.$$

the j^{th} level of factor B is in the second selected subset if

$$\bar{y} \dots j \dots \geq \max_{j' \neq j} \{ \bar{y} \dots j' \dots \} - d \left(\frac{2E_b}{ra} \right)^{\frac{1}{2}}.$$

The corresponding simultaneous confidence intervals with the best are given by

$$[0, D_i] \text{ for } \tau_{[a]} - \tau_i \text{ and}$$

$$[0, G_j] \text{ for } \pi_{[b]} - \pi_j$$

where

$$D_i = \max \left\{ 0, \max_{i' \neq i} \{ \bar{y} \dots i' \dots \} - \bar{y} \dots i \dots + d \left(\frac{2E_a}{rb} \right)^{\frac{1}{2}} \right\}$$

and

$$G_j = \max \left\{ 0, \max_{j' \neq j} \{ \bar{y} \dots j' \dots \} - \bar{y} \dots j \dots + d \left(\frac{2E_b}{ra} \right)^{\frac{1}{2}} \right\}.$$

Again the problem is to determine d . An algorithm follows:

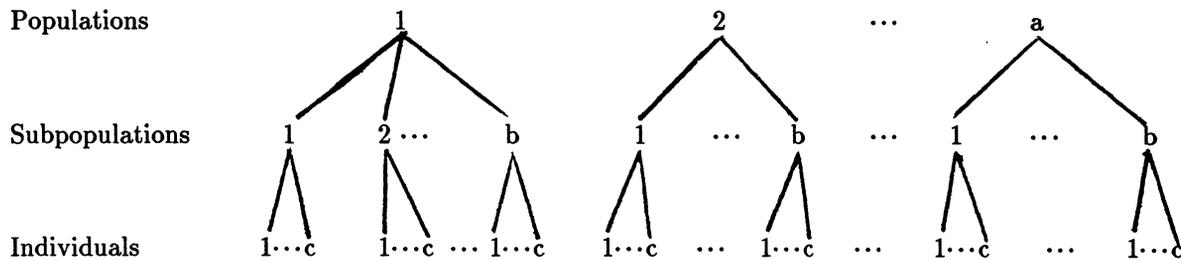
1. Generate a set of standard normal variates, Z_i , with correlation $\frac{1}{2}$ ($i = 1, 2, \dots, a - 1$).

2. Generate an independent (of Z_i) set of standard normal variates, X_j , with correlation $\frac{1}{2}$ ($j = 1, 2, \dots, b - 1$).
3. Generate a χ^2 variate, U , with f_a degrees of freedom.
4. Generate a χ^2 variate, V , with f_b degrees of freedom.
5. Calculate $W = \max \left\{ \max_i \frac{Z_i}{\sqrt{U/f_a}}, \max_j \frac{X_j}{\sqrt{V/f_b}} \right\}$
6. Repeat steps 1 - 5 N times (a large number).
7. Order the values in 6 from smallest to largest and use the αN largest one as d .

For an error rate of $\alpha = .05$ Edwards and Berry (1987) recommend choosing N in the range 3,200 to 320,000. Even the upper value is easily implemented on a fast personal computer. We used an 80486 IBM-PC compatible computer running at 33MHz and the matrix language GAUSS. As an example, the subset-selection problem with $a = 4$, $b = 5$ and $N = 320,000$ required 22 minutes. Note this is 3 times faster than the mainframe results reported by Edwards and Berry (1987)! Problems with complicated correlation structures do not require significantly more time since the Cholesky decomposition required to simulate the correlation structure needs to be computed only once.

4. Split-Split Plot Design

To illustrate multiple comparisons procedures for a split-split plot design, we select a particular member of the family of these designs. The treatment design selected for this experiment design may be schematically represented as follows:



The set of a populations will be the whole or main plots which will be arranged in a randomized complete block design of r blocks. Each population has b subpopulations which are nested within

populations. These b subpopulations of a population will be randomly assigned to the split plot experimental units within each whole plot or population. Note that subpopulation 1 from population 1 has nothing in common with subpopulation 1 from population 2 except the number 1. Each subpopulation has c individuals which are randomly allocated to the split-split plot experimental units within each split plot experimental unit or a subpopulation. There are r randomizations performed for the populations in the whole plots, ra randomizations on the subpopulations in the split plots, and rab randomizations on the individuals in the split-split plots within a subpopulation. A linear model for the response from the $ghij$ th split-split plot experimental unit is

$$Y_{ghij} = \mu + \rho_g + \alpha_h + \delta_{gh} + \beta_{hi} + \pi_{ghi} + \gamma_{hij} + \epsilon_{ghij}, \quad (2)$$

where μ is an effect common to every observation, ρ_g is the g^{th} block effect, α_h is the h^{th} population effect, δ_{gh} is a random error effect associated with whole plot experimental units, β_{hi} is the effect of subpopulation i from population h , π_{ghi} is a random error effect associated with split plot experimental units, γ_{hij} is the effect for individual j from subpopulation i from population h , and ϵ_{ghij} is a random error effect associated with split-split plot experimental units. An analysis of variance for this response model is outlined in Table 3.1. In the table, additional partitioning of various sets of degrees of freedom are made. A comparison of the error terms making up E_b and E_c may be made if desired. Likewise the relative sizes of the mean squares making up the pooled mean square for subpopulations within populations and of the individuals within subpopulations are of interest in certain situations.

The various variances of a difference between two means that are of interest in this design are

$$V(\bar{y}_{.h..} - \bar{y}_{.h'..}) = 2E_a/rbc$$

$$V(\bar{y}_{.hi.} - \bar{y}_{.hi'.}) = 2E_b/rc$$

$$V(\bar{y}_{.hi.} - \bar{y}_{.h'i.}) = V(\bar{y}_{.hi.} - \bar{y}_{.h'i'.}) = 2[E_a + (b-1)E_b]/rbc$$

$$V(\bar{y}_{.hij} - \bar{y}_{.hij'.}) = 2[E_c]/r$$

$$V(\bar{y}_{.hij} - \bar{y}_{.hi'j'}) = V(\bar{y}_{.hij} - \bar{y}_{.hi'j'}) = 2[E_b + (c-1)E_c]/rc$$

$$V(\bar{y}_{.hij} - \bar{y}_{.h'ij'}) = V(\bar{y}_{.hij} - \bar{y}_{.h'ij'}) = V(\bar{y}_{.hij} - \bar{y}_{.h'ij'}) = V(\bar{y}_{.hij} - \bar{y}_{.h'ij'}) =$$

$$2[E_a + (b-1)E_b + b(c-1)E_c]/rbc$$

Table 3.1. An analysis of variance for a split-split plot design.

Source of variation	Degrees of freedom	Mean square
Total	rbc	
Correction for mean	1	
Block	$r - 1$	
Whole plot (populations = Factor A)	$r - 1$	
Error (a)	$f_a = (r - 1)(a - 1)$	E_a
Split plots (subpopulations within populations)	$a(b - 1)$	
Subpop'ns within pop'n 1	$b - 1$	
Subpop'ns within pop'n 2	$b - 1$	
⋮		
Subpop'ns within pop'n a	$b - 1$	
Error (b)	$f_b = a(r - 1)(b - 1)$	E_b
Subpop'n 1 × block	$(r - 1)(b - 1)$	
⋮	⋮	
Subpop'n b × block	$(r - 1)(b - 1)$	
Individuals within subpopulations	$ab(c - 1)$	
Ind. wn subpop 1, pop 1	$c - 1$	
Ind. wn subpop 2, pop 1	$c - 1$	
⋮	⋮	
Ind. wn subpop b, pop 1	$c - 1$	
Ind. wn subpop 1, pop 2	$c - 1$	
⋮	⋮	
Ind. wn subpop b, pop 1	$c - 1$	
Error (c)	$f_c = ab(r - 1)(c - 1)$	E_c
Ind. wn subpop 1, pop 1 × block	$(r - 1)(c - 1)$	
Ind. wn subpop 2, pop 1 × block	$(r - 1)(c - 1)$	
⋮	⋮	
Ind. wn subpop b, pop a × block	$(r - 1)(c - 1)$	

Here again, the multiple error mean squares is a feature of the experiment design and not of the fact that multiple comparisons procedures are being considered.

For each of the above variances of a difference between two means, a batchwise error rate could be set. For an experimentwise error for the batch of comparisons pertaining to pairwise comparisons among the a population means, v would be set equal to a , the number of populations, in $q_{\alpha, f, v}$. The corresponding confidence interval would be

$$\bar{y} \cdot h \cdot \cdot - \bar{y} \cdot h' \cdot \cdot \pm q_{\alpha, f_a, a} (E_a / rbc)^{1/2} .$$

The value for v using the second variance and the resulting set of pairwise comparisons above, would be $v = ab$, and the corresponding confidence interval would be

$$\bar{y} \cdot hi \cdot - \bar{y} \cdot hi' \cdot \pm q_{\alpha, f_b, ab} (E_b / rc)^{1/2}, i \neq i' .$$

For the third variance, $v = ab$, and the corresponding confidence intervals are

$$\bar{y} \cdot hi \cdot - \bar{y} \cdot h'i' \cdot \pm q_{\alpha, f^*, ab} \{ [E_a + (b-1)E_b] / rb \}^{1/2}, h \neq h' \text{ and } i \neq i' ,$$

where the degrees of freedom, f^* , for this variance may be approximated via a Cochran type (Cochran and Cox, 1957, p. 101) approximation (but see the previous section for an alternate approach). For the fourth variance, $v = abc$ and the corresponding confidence interval is

$$\bar{y} \cdot hij - \bar{y} \cdot hij' \pm q_{\alpha, f_c, abc} (E_c / rb)^{1/2}, j \neq j' .$$

For the last two variances, $v = abc$ and f^* needs to be approximated for each variance. The corresponding confidence intervals are

$$\bar{y} \cdot hij - \bar{y} \cdot hi'j' \pm q_{\alpha, f^*, abc} \{ [E_b + (c-1) E_c] / rc \}^{1/2}, i \neq i' \text{ and } j \neq j'$$

and

$$\bar{y} \cdot hij - \bar{y} \cdot h'i'j' \pm q_{\alpha, f^*, abc} \{ [E_a + (b-1) E_b + b(c-1)E_c] / rbc \}^{1/2} ,$$

$$h \neq h', i \neq i' \text{ and } j \neq j' .$$

For an approximately experimentwise error rate, the v used for all six variances would be $v = abc$.

For the subset selection procedure, the rules in the preceding paragraph would apply except that in each case, $v - 1 = k$ would be used in place of v when entering Dunnett's tables. The appropriate degrees of freedom would be used in each case.

5. Discussion

We have considered both traditional and simulation-based approaches to multiple comparisons in complicated experiment designs with multiple error terms. Procedures such as the lsd or esd generalize in a straightforward manner. However, because of the multiple error terms and multiple degrees of freedom, tables for the hsd and subset selection are not available. In such cases, simulation-based approaches are straightforward to implement, though computationally intensive.

6. Acknowledgements

This research was partially supported by the Math Sciences Institute at Cornell University and Hatch grants 151-401 and 151-406. We wish to thank the two referees for constructive comments.

7. Literature Cited

- Chew, V. (1977). Comparisons among treatment means in an analysis of variance. Agricultural Research Service, USDA, ARS/H/6, iv + 64 pp.*
- Cochran, W.G. and G.M. Cox (1957). *Experimental Designs*, second edition. John Wiley & Sons, Inc., New York.
- Dunnett, C.W. (1955). A multiple comparisons procedure for comparing several treatments with a control. *Journal of the American Statistical Association* 50, 1096-1121.
- Edwards, D. and Berry J.J. (1987). The efficiency of simulation-based multiple comparisons. *Biometrics* 43, 913-928.
- Federer, W.T. (1975). The misunderstood split plot. In *Applied Statistics* (Editor: R.P. Gupta), North Holland Publishing Company, Amsterdam pp. 9-39.
- Federer, W.T. (1961). Experimental error rates. *American Society for Horticultural Science* 78, 605-615.
- Federer, W.T. and C.E. McCulloch (1984). Multiple comparisons procedures for some split plot and split block designs. In *Design of Experiments - Ranking and Selection* (Editors: T.J. Santner and A.C. Tamhane), Marcel Dekker, Inc., New York and Basel.*
- Federer, W.T. and M.P. Meredith (1991). Covariance analysis for split plot and split block designs.

(To appear in the American Statistician, in May, 1992.)

- Ghosh, M.N. (1955). Simultaneous tests of linear hypotheses. *Biometrika* 42, 441-449.
- Grimes, B.A. and W.T. Federer (1984). Comparison of means from populations with unequal variances. In *W.G. Cochran's Impact on Statistics* (Eds. P.S.R.S. Rao and J. Sedransk) John Wiley & Sons, Inc., New York, pp. 353-374.*
- Hartley, H.O. (1955). Some recent developments in analysis of variance. *Communications in Pure and Applied Mathematics* 8, 47-72.*
- Holland, B. (1991). Comment on Saville. *The American Statistician* 45, 165.
- Hsu, J.C. (1981). Simultaneous confidence intervals for all distances from the "best". *Annals of Statistics* 9, 1026-1034.
- Lea, P. (1991). Multiple confusions. *The American Statistician* 45, 165-166.
- Saville, D.J. (1990). Multiple comparisons procedures: The practical solution. *The American Statistician* 44, 174-180.
- Saville, D.J. (1991). Reply to Holland and Lea. *The American Statistician* 45, 166-168.
- Tukey, J.W. (1991). The philosophy of multiple comparisons. *Statistical Science* 6, 100-116.*

*See these papers for additional references.