Analysis of treatment means from factorial experiments with unequal replication

Michael P. Meredith[2]

ABSTRACT

A two-step procedure is proposed for the analysis of factorial experiments with unequal replication. The procedure entails a check for interaction in the general means model, followed by estimation of either main effects or simple effects. The use of a set of contrasts which will address the hypotheses of interest is advocated over a set which is orthogonal and dependent on the number of replications. The problem of no replication for some treatments is briefly discussed along with the inherent difficulties.

The proposed approach to data analysis is applied to the results of a multiple cropping experiment. Care is exercised when invoking a statistical computing package so that the pitfalls of the default analyses are avoided. The aim of the data analysis is to allow the experimenter to specify the contrasts of research interest rather than rely upon the default options of a computing package.

---

[1] Contribution from the Biometrics Unit, Cornell University, Ithaca, NY 14853 as Technical Report BU-819-M.

[2] Teaching Support Specialist, Dept. of Plant Breeding and Biometry, Biometrics Unit, Cornell University, Ithaca, NY 14853.

Additional index words: Unbalanced data; Factorial experiments; Statistical computing packages.

INTRODUCTION

The analysis of treatment means from factorial experiments with unequal replication is a problem commonly confronted by some agricultural researchers. Unequal replication may arise due to topographic or economic constraints at the onset of an experiment, or due to the destruction or loss of experimental units while the experiment is being conducted. Unequal replication is sometimes termed unbalanced or messy data in the literature.

There is an abundance of statistical literature addressing the problem of analyzing data from experiments with unequal replication (Searle, 1971; Speed, Hocking and Hackney, 1978). However, it is precisely this wealth of literature that may make the task of finding the appropriate procedures for the problem at hand a difficult one. To assist in this task, the "Instructions to Authors" in the Agronomy Journal (1982) gives some indication of how researchers may go about reporting the results of experiments with well-defined treatment structures. The following is a quote from the Statistical Methods section of the "Instructions to Authors":

> "Whenever possible, treatment comparisons that are logical from a scientific standpoint should be made as single degree of freedom contrasts as part of the analysis of variance. Orthogonality of these contrasts is desirable because information from one test is independent of others but such orthogonality is not necessary. A more important criterion is whether the particular contrasts are meaningful and/or were planned before the data were examined."

With the above suggestions in mind the present article proposes a systematic approach to the analysis of factorial experiments with unequal replication. As an example, the proposed approach is applied to the results of a multiple cropping experiment.

METHODOLOGY

Typically, a researcher is interested in estimating sample means and the associated standard errors. If the treatments are in a factorial arrangement, then well-defined single degree-of-freedom contrasts may be estimated from the sample means. The standard errors associated with each contrast need to be calculated as well.

For example, consider a 2×3 factorial experiment where each of the three levels of factor A occur with each of the two levels of factor B. Schematically, the statistical layout appears as:

	A1	A2	A3
B1	μ_1	μ_2	μ_3
B2	μ_4	μ_5	μ_6

.

Interest lies in estimating the μ_j's as well as linear combinations of the μ_j's.

A statistical model useful in such a situation is termed the general means model (Allen and Cady, 1982) and is written as

$$y_{j\ell} = \mu_j + \epsilon_{j\ell} \qquad \text{where} \quad j = 1,2,\cdots,t \quad \text{and} \quad \ell = 1,2,\cdots,n_j \ ,$$

$$n_j \geq 1 \quad \text{for all } j \ .$$

The jth treatment combination has n_j replications. In the above example $t = 6$, the number of treatment combinations. The general means model asserts that the $y_{j\ell}$th observation is independently drawn from a distribution with mean μ_j and common variance σ^2. The above model is appropriate for a completely randomized design and extensions for other designs are straightforward.

In the case of equal replication, i.e., $n_j = n$ for all j, then it is

relatively simple to write down a meaningful, complete orthogonal set of contrasts. If c_{ij} represents the coefficient of the jth mean for the ith contrast then the following relationships are true:

$$\sum_{j=1}^{t} c_{ij}\mu_j = L_i \ , \qquad \sum_{j=1}^{t} c_{ij} = 0$$

and

$$\sum_{j=1}^{t} c_{ij}c_{i'j} = 0 \qquad i \neq i', \ i = 1,2,\cdots,t-1 \ .$$

Examples of complete sets of orthogonal contrasts may be found in many textbooks (Cochran and Cox, 1957). However, when there is unequal replication, i.e., $n_j \neq n$ for some j, then the problem of determining a complete orthogonal set of contrasts which is meaningful to the researcher becomes a difficult if not fruitless pursuit. The problem lies in the fact that the contrast coefficients are now dependent upon the individual n_j. Thus, a treatment combination that had more replication may receive more weight in the orthogonal contrast than in the natural contrast (the term natural contrast will be used to denote the coefficients that would arise if equal replication was the case). Unless unequal replication was designed into the experiment for reasons of precision, it is typically the natural set of contrasts that answer the questions of research interest.

In using the natural set of contrasts when there is unequal replication the orthogonality is, in general, lost. But if the orthogonal set fails to address the questions of interest, then little is gained by strictly adhering to the principle of orthogonality. An example of choosing contrasts of subject matter interest in the area of animal

science is discussed in Urquhart and Weeks (1978).

An approach which employs natural contrasts in the unequal replication setting is the analysis of unweighted sample means. Snedecor and Cochran (1980) caution that this approach will yield reasonable approximations to the F distribution only if the ratio of the largest to the smallest n_j is no greater than two. If this ratio exceeds two, or if the analysis of unweighted means is unsatisfactory, then the two-step approach to be given below may be used.

In what follows a main effect is defined to be the comparison of levels of one factor averaged over all levels of the other factors. A simple effect is defined to be the comparison of levels of one factor at fixed levels of all other factors.

Step 1: Analysis of the general means model.

In this step the importance of interaction is assessed. The interaction between treatment factors may be assessed by using the composite F-test. Main effects due to treatment factors are not evaluated in this step. A residual analysis should be performed at this step.

Step 2: a) If the interaction is deemed to be unimportant, then proceed to evaluate main effects using the reduced model. The reduced model is the general means model with the restriction that all interactions are zero. This is equivalent to the practice of "pooling" interaction sum-of-squares with experimental error when the composite test for interaction is not significant. Some guidance for

choosing a type I error level to assess interactions may be found in Bancroft (1964).

b) If the interaction is found to be important, then retain the general means model and proceed to evaluate simple effects.

The above two-step procedure should be coupled with plots of the cell means to visually display the outcome of the experiment. If interactions are present, then such a plot assists in elucidating their whereabouts. Standard error bars should also be included about the estimated means.

Although the simple two-step procedure outlined above generally suffices to approach the analysis of many unequal replication factorial experiments, there are several special notes worthy of mention.

If main effects are to be assessed compositely to determine if the sums-of-squares due to a particular factor should be pooled or not, then some additional guidelines are required. The reader is referred to Table 18.9 of Allen and Cady (1982) for such an approach.

If the factorial arrangement of treatments includes a control level of each factor, then careful consideration should be given to the test for interaction. Often the behavior of the control responses are quite disparate from the remainder of the experiment. Such a situation will potentially result in a significant F-statistic in the composite test for interaction, even though there is no interaction between the treatment factors other than that introduced by the control treatment. In this case the single degree-of-freedom contrast associated with the control treatment should be partitioned from the interaction sums-of-

squares. Then, test this individually, and test the remaining interaction sums-of-squares via a composite F test. Under this approach the two-step procedure may be rewritten in the following way.

Step 1': Assess the importance of the single degree-of-freedom interaction contrast and the remaining composite interaction by way of the general means model.

Step 2': a) Same as in Step 2a.

b) If the single degree-of-freedom interaction contrast is significant and the remaining composite test is not, then proceed to estimate main effects for that portion of the experiment free of interaction. Evaluate simple effects for those combinations with the control.

c) Both the single degree-of-freedom and remaining composite tests are important. Proceed to evaluate simple effects in the general means model.

d) The single degree-of-freedom contrast is unimportant and the remaining composite test is significant. Proceed to estimate simple effects using the general means model.

One important difference between the general means model and the reduced (no interaction) model needs to be discussed. The estimated means, $\hat{\mu}_j$, in the general means model are simply the sample means, $\bar{y}_j$. However, in the reduced model this is no longer the case since we have imposed the restriction that interactions are defined to be zero. Thus, the estimated means $\hat{\mu}_j$ under the reduced model will be such that any contrast among the estimated column (row) means is the same for each row

(column). That is, the no interaction criteria is met.

DISCUSSION

In the preceding it has been assumed that each n_j was non-zero. Suppose now that some of the treatment combinations have no replication (i.e., $n_j = 0$) due to either missing data or lack of interest in the particular treatment combination(s). The general means model still applies, but the appropriate choice of a set of contrasts is no longer obvious.

As before, the analysis should be directed to address the hypotheses of research interest. The underlying complete factorial treatment structure should be regarded more loosely now. The absence of some treatments clearly alters the usual notions of interactions and main effects in a complete factorial. If a meaningful set of contrasts is not forthcoming, then it is often fruitful to seek a subset of the treatments available which do form a complete factorial experiment. If such a subset (or several subsets) may be found, then the procedures described above may be directly applied. As an example, consider what was originally a 3 x 3 factorial experiment. Suppose that the (1,3) and (3,2) treatment combinations are missing as indicated below:

	A1	A2	A3
B1	μ_1	μ_2	X
B2	μ_3	μ_4	μ_5
B3	μ_6	X	μ_7

In this example, two complete 2 x 2 factorials may be recognized. They are as follows:

	B1	A2
B1	μ_1	μ_2
B2	μ_3	μ_4

and

	A1	A3
B2	μ_3	μ_5
B3	μ_6	μ_7

.

One difficulty should be pointed out. If the same data were used to estimate μ_3 in each 2 x 2 experiment, then the separate analyses will not be independent. Although the lack of independence is unsavory, the construction of the two orthogonal interaction contrasts for the original table are, at best, difficult to understand.

The above discussion should help emphasize the need for both careful treatment design and conduct of an experiment. Haphazard experiments tend to admit less than fruitful results when a convolute analysis must be performed.

EXAMPLE

An experiment was conducted on a low nitrogen field soil to determine the effect of growing peas and barley in monoculture versus mixture. The experiment was a 3 x 3 factorial laid out in a completely randomized design with three replications. Pea and barley monocultures were planted at 100%, 150%, and 200% of the normal planting rate by increasing the seeding rate within rows. Polycultures were formed at each of these densities by substituting alternate rows of one crop for the other. Consequently, at each of the three densities two monocultures and a 50:50 alternate row polyculture were planted. The plots were harvested at dry maturity and dry seed yield reported as grams/quadrat.

During the growing season several complications arose which altered the original balanced 3 x 3 factorial layout. At harvest the plant

densities within plots varied from the desired planted densities. It was decided that samples would be grouped into either high (>150% of normal) or low (≤150% of normal) density based upon the number of plants per plot at final harvest. Thus, the experiment was analyzed as a 2 x 3 factorial with unequal replication of the six treatment combinations. It should be noted that five plots were lost during the course of the experiment, yielding a total of 22 responses at final harvest.

The statistical layout of the final harvest is shown below. The number of replications for each treatment combination is reported.

		System		
		Peas	Barley	Peas and Barley
Density	Low	2	3	4
	High	5	4	4

.

The actual data and computer code used to analyze the experiment is included in the appendix.

The composite test for interaction indicates that interaction is present ($p = 0.04$). Upon fitting the general means model, the following table of predicted treatment means is computed:

		System		
		Peas	Barley	Peas and Barley
Density	Low	82.955	68.883	91.663
	High	88.154	78.315	127.663

.

From this table of cell means two single degree-of-freedom interaction contrasts were examined. The 2 x 2 portion of the experiment associated with the monocultures appears to be free of interaction ($p = 0.74$).

However, the difference in yield between densities for the polyculture is significantly greater than the average difference between densities for the monocultures ($p = .01$). Thus, the significance of the composite test for interaction is due almost entirely to the single degree-of-freedom associated with the polyculture vs. averaged monocultures interaction contrast. Note that these are natural, not orthogonal interaction contrasts.

Simple effects are now estimated to assess the difference in yield due to density for each of the cropping systems. For peas and barley in polyculture the yield is 36.00 ± 7.87 g greater for the high density ($p < 0.001$); for peas alone this difference is 5.20 ± 9.31 g ($p = 0.58$) and for barley alone this difference is 9.43 ± 8.50 g ($p = 0.28$). Alternatively, the main effects for monoculture may be estimated since the 2×2 monoculture portion of the experiment appears free of interaction. The high density yields are 7.32 ± 6.30 g greater for monocultured peas and barley ($p = 0.26$).

Thus, the densities studied do not significantly affect yields of peas or barley grown in monoculture. However, the yield is significantly greater for the polyculture grown at the higher density.

In the unlikely event that the composite F-test for interaction between cropping system and density is felt to be unimportant ($p = 0.04$), then the reduced model is fitted. The treatment means are now estimated with the restriction that interactions are zero. For the sake of completeness, the estimated means are given in the following table:

		System		
		Peas	Barley	Peas and Barley
Density	Low	73.391	63.651	100.368
	High	91.979	82.239	118.956

Note that the difference between rows is the same for each column. Alternatively, note that any contrast among the columns is the same for each row.

ACKNOWLEDGEMENTS

Matt Liebman of the Department of Ecology and Systematics at Cornell University was kind enough to supply the data set used in the example.

LITERATURE CITED

1. Allen, D. M., and F. B. Cady. 1982. Analyzing experimental data by regression. Lifetime Learning Publications, Belmont, Calif.

2. Bancroft, T. A. 1964. Analysis and inference for incompletely specified models involving the use of preliminary test(s) of significance. Biometrics 20, 427-442.

3. Buxton, D. 1982. Instructions to authors. Agron. J. 74, 1100-1102.

4. Cochran, W. G., and G. M. Cox. 1957. Experimental designs, 2nd ed. John Wiley and Sons, Inc., N. Y.

5. SAS Institute, Inc. 1982. SAS user's guide: Statistics. SAS Institute, Inc., Cary, NC.

6. Searle, S. R. 1971. Linear models. John Wiley and Sons, Inc., N. Y.

7. Snedecor, G. W., and W. G. Cochran. 1980. Statistical methods, 7th ed. The Iowa State University Press, Ames, Iowa.

8. Speed, F. M., R. R. Hocking, and O. P. Hackney. 1978. Methods of analysis of linear models with unbalanced data. J. Am. Stat. Assoc. 73, 105-112.

9. Urquhart, N. S., and D. L. Weeks. 1978. Linear models in messy data: Some problems and alternatives. Biometrics 34, 696-705.

APPENDIX

The SAS (Statistical Analysis System, 1982) program used to analyze the data in the example is given below. Selected annotations follow the program.

```
DATA CROP;
INPUT SYSTEM $ DENSITY $ YIELD;
CARDS;
   P L 89.13
   P L 76.78
   P H 109.67
   P H 89.67
   P H 75.44
   P H 89.6
   P H 76.39
   B L 75.59
   B L 70.63
   B L 60.43
   B H 80.8
   B H 77.45
   B H 79.05
   B H 75.96
   X L 88.28
   X L 104.5
   X L 84.9
   X L 88.97
   X H 125.44
   X H 128.96
   X H 108.51
   X H 147.74
PROC PRINT N;

PROC PLOT DATA=CROP;
     PLOT YIELD*SYSTEM=DENSITY;
     PLOT YIELD*DENSITY=SYSTEM;

PROC GLM; CLASSES SYSTEM DENSITY;
     MODEL  YIELD = SYSTEM DENSITY SYSTEM*DENSITY;
     LSMEANS SYSTEM DENSITY SYSTEM*DENSITY / STDERR;
     MEANS   SYSTEM DENSITY SYSTEM*DENSITY / DEPONLY;
     OUTPUT OUT=NEW1 PREDICTED=YHAT1 RESIDUAL=RESID1;

PROC PLOT;
     PLOT RESID1*YHAT1 / VREF=0;
```

① ② ③ ④ ⑤

APPENDIX (cont.)

```
PROC GLM; CLASSES SYSTEM DENSITY;                                    } (6)
    MODEL  YIELD = SYSTEM*DENSITY / NOINT SS1;                       }
    ESTIMATE 'MONO * DENSITY'
        SYSTEM*DENSITY -1 1 1 -1 0 0 / DIVISOR=2;                    } (7)
    ESTIMATE 'MONOPOLY * DENSITY'
        SYSTEM*DENSITY -1 1 -1 1 2 -2 / DIVISOR=4;                   }
    ESTIMATE 'DENSITY W/I POLY'
        SYSTEM*DENSITY 0 0 0 0 1 -1;
    ESTIMATE 'DENSITY W/I PEAS'                                      } (8)
        SYSTEM*DENSITY 0 0 1 -1 0 0;
    ESTIMATE 'DENSITY W/I BARLEY'
        SYSTEM*DENSITY 1 -1 0 0 0 0;
    ESTIMATE 'PEAS VS BARLEY FOR MONO'
        SYSTEM*DENSITY -1 -1 1 1 0 0 / DIVISOR=2;                    } (9)
    ESTIMATE 'DENSITY FOR MONO'
        SYSTEM*DENSITY 1 -1 1 -1 0 0 / DIVISOR=2;                    }
    LSMEANS SYSTEM*DENSITY / STDERR;


PROC GLM; CLASSES SYSTEM DENSITY;                                    } (10)
    MODEL  YIELD = SYSTEM DENSITY / SS1 SS2 P CLM;                   }
    LSMEANS SYSTEM DENSITY / STDERR;

    ESTIMATE 'POLY VS MONO'                                          }
       SYSTEM -1 -1 2 / DIVISOR=2;                                   }
    ESTIMATE 'BARLEY VS PEAS'                                        } (11)
       SYSTEM -1 1 0;                                                }
    ESTIMATE 'DENSITY MAIN EFFECT'                                   }
       DENSITY 1 -1;                                                 }
    OUTPUT OUT=NEW2 PREDICTED=YHAT2 RESIDUAL=RESID2;                 } (12)


PROC PLOT;                                                           }
    PLOT RESID2*YHAT2 / VREF=0;                                      }
```

Annotations:

1. P = peas monoculture, B = beans monoculture, X = peas and beans mixed in polyculture, H = high density and L = low density.
2. Plots of the observed responses.
3. The composite test for interaction is given by the F-statistic associated with the SYSTEM*DENSITY term.
4. The LSMEANS are the unweighted means whereas the MEANS give the weighted means. The cell means are the same for both in the full model.
5. Residual analysis for the full model.
6. Fitting the general means model.
7. The two "natural" interaction single degree-of-freedom contrasts. In general, these will not be orthogonal. These contrasts test for interaction between monoculture and density, and between mono- vs. polyculture and density.
8. These are the single degree-of-freedom simple effects contrasts to assess how yields differ due to density for each of the cropping systems.
9. These are the two main effects contrasts for the 2 x 2 factorial of density by monocultures.
10. Fitting the reduced model with interaction defined to be zero. The Type II sums of squares are equivalent to the Type III and IV and are less expensive to compute. The P and CLM options print the predicted cell means and a 95% confidence interval for each observation. The cell means in the last table are gleaned from the results of the P option.

11. Main effect contrasts for cropping systems and density with no interaction.

12. Residual analysis for the reduced model.