# PRAGMATIC METHODOLOGY FOR ANALYSIS OF FACTORIAL EXPERIMENTS WITH UNEQUAL REPLICATION

Michael P. Meredith[1] and Foster B. Cady[2]

BU-969-M[3] March 1988

## 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 treatment mean comparisons which address the hypotheses of interest is advocated over a set which is orthogonal and dependent on the number of replications. Emphasis is on estimation of means, appropriate mean comparisons, and standard errors rather than upon hypothesis testing. The problem of no replication for some treatments is briefly discussed along with the inherent difficulties. No replication is considered for the distinct cases where interaction is of concern or interest and where it is defined to be zero in the case of many cross-classified factors in observational studies.

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 a default analysis are avoided. The aim of data analysis is to allow the experimenter to specify mean comparisons of research interest rather than rely upon the default options of computing packages.

---

[1]Assistant Professor, Biometrics Unit, Cornell University, Ithaca, NY

[2]Adjunct Professor, Dept. of Plant Breeding and Biometry, Cornell University, Ithaca, NY

[3]In the Biometrics Unit technical report series, Cornell University, Ithaca, NY, 14853.

## INTRODUCTION

The analysis of treatment means from factorial experiments with unequal replication is a problem that often confronts researchers from many areas. Unequal replication may arise due to economic constraints at the onset of an experiment, or due to the loss of experimental units while the experiment is being conducted. Unequal replication has sometimes been termed "unbalanced" or "messy" data in the statistical literature.

There is abundant statistical literature addressing the problem of analyzing data from experiments with unequal replication (Searle, 1971, 1987a,b; Hocking, 1984; Freund, Littell, and Spector, 1987; Speed, Hocking and Hackney, 1978). For over a decade focus has been upon the appropriate sums-of-squares to use when analyzing unbalanced cross-classified data as in factorial experiments. The question never seemed to arise whether or not the F-tests associated with these sums-of-squares were addressing hypotheses of interest to the investigator. A consequence of this focus is that many practical investigators are now well versed in "the analysis of unbalanced data" and therefore fail to provide analyses of their experiments that are as useful or powerful as they could be. The questions inherent in most factorial experiments are best addressed via estimates of meaningful contrasts amongst the observed treatment means and their associated standard errors.

Given the focus of the past decade, and its wealth of literature, the task of finding the appropriate procedures for the problem at hand has become an unnecessarily difficult one. To assist in this task some subject matter journals have tried to specify guidelines for their prospective authors to follow in presenting the results of data analysis. For example, "Instructions to Authors" in the *Agronomy Journal* (1982) give some indication of how agricultural researchers should report results of experiments with well-defined treatment structures. The following is quoted 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 df 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."

It would seem that with recommendations such as the above appearing in subject matter journals that the instructor of statistical methods to an audience that

conducts (or shall conduct) designed research should provide their students with a methodology that is tenable in the real world. The past focus on F-tests associated with various sums-of-squares has led such methodology astray.

With the above discussion in mind the present article proposes a systematic approach to the analysis of factorial experiments having unequal replication. Emphasis is placed on estimating meaningful treatment mean comparisons and their standard errors. The recommended procedure uses available statistical computing packages with general linear model or regression programs including options that easily handle (i) continuous and categorical factors, and (ii) estimation of treatment mean comparisons and their standard errors. As an example, the proposed approach is applied to data from a multiple cropping experiment.

## DISCUSSION AND METHODOLOGY

Typically, a researcher is interested in estimating sample means and their associated standard errors. If the treatments are in a factorial arrangement then well-defined single degree-of-freedom (df) contrasts may be estimated from the sample means. The standard errors associated with each contrast need to be calculated as well. It should be noted that these contrasts and standard errors are not supplied in the default output of any statistical package *since the contrasts are dictated by the objectives of the researcher in designing the experiment.*

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. The statistical layout and expected cell means appear as:

| | | A = 1 | A = 2 | A = 3 |
|---|---|---|---|---|
| B | 1 | $\mu_{11}$ | $\mu_{12}$ | $\mu_{13}$ |
| | 2 | $\mu_{21}$ | $\mu_{22}$ | $\mu_{23}$ |

Interest lay in estimating the $\mu_{ij}$'s as well as linear combinations of the $\mu_{ij}$'s.

A statistical model useful in such a situation is termed the general means model and is written as: $Y_{ijk} = \mu_{ij} + \varepsilon_{ijk}$, where $i=1, 2,..., b$, $j=1, 2,..., a$, and $k=1, 2,..., n_{ij}$, and $n_{ij} \geq 1$ for all i,j .

The i,jth treatment combination has $n_{ij}$ replications. In the above example ab=6, the number of treatment combinations. The general means model asserts that the $Y_{ijk}$th observation is independently drawn from a distribution with mean $\mu ij$ and common variance $\sigma^2$. The above model is appropriate for a completely randomized design and extensions for other designs are straightforward. A transformation may be required to meet the common variance assumption. The assumption of normality is necessary to form F-ratios or use t-critical values to construct confidence intervals about sample means or estimated contrasts.

In the case of equal replication, i.e., $n_{ij}=n$ for all i,j, it is relatively simple to write down a meaningful, complete orthogonal set of contrasts. If $c_{ijm}$ represents the coefficient of the i,jth cell mean for the mth contrast then the following are true:

$$\sum_{i=1}^{b}\sum_{j=1}^{a} c_{ijm}\mu_{ij} = L_{ij}\,, \quad \sum_{i=1}^{b}\sum_{j=1}^{a} c_{ijm} = 0, \text{ and}$$

$$\sum_{i=1}^{b}\sum_{j=1}^{a} c_{ijm}c_{ijm'} = 0, \text{ for all } m \neq m',\ m = 1, 2,..., ab-1\,.$$

This set may reflect structure in qualitative treatment combinations or response curves or surfaces when the treatments are combinations of quantitative and qualitative factors.

Examples of complete sets of orthogonal contrasts may be found in many textbooks (Cochran and Cox, 1957, section 3.4). However, when there is unequal replication, i.e., $n_{ij} \neq n$ for some i,j, the problem of determining a complete orthogonal set of contrasts which is also meaningful to the researcher becomes arduous, if not impossible. The problem lay in the fact that the contrast coefficients are now dependent upon the individual $n_{ij}$ (see Allen and Cady, 1982, Unit 18). Thus, a treatment combination that had more replications 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 set of natural contrasts that answer questions of research interest in designed factorial experiments and cross-classified observational studies.

When using the set of natural contrasts under unequal replication the orthogonality is, in general, lost (i.e., the contrasts are no longer statistically independent of one another). But if an 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). A forestry example was discussed by Warren (1979).

An approach which employs natural contrasts in the unequal replication setting is the analysis of unweighted sample means (AUWM). Snedecor and Cochran (1980), Section 20.4, caution that this approach will yield reasonable approximations to the F-distribution only if the ratio of the largest to the smallest $n_{ij}$ is no greater than two. If this ratio exceeds two, or if the AUWM is unsatisfactory, then the two-step procedure to be given below may be used. The analysis of unweighted sample means arose largely for computational convenience prior to the availability of computers. Today there is little need for this method of analysis except perhaps for inexpensive exploratory analyses of large scale studies having unbalanced data. Speed and Monlezun (1979) point out that the AUWM is exact in the simple case of a $2 \times 2$ factorial experiment.

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. Single df interaction contrasts, guided by meaningful contrasts among the levels of each treatment factor, may be evaluated. In general, however, a practical alternative for assessing interaction between treatment factors is the composite F-test (Snedecor and Cochran, 1980, Sections 16.6 and 20.3). Main effects due to treatment factors are not evaluated in this step. A residual analysis should be performed to determine if underlying assumptions are tenable.

**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" the interaction sum-of-squares with experimental error when the composite test for interaction is not significant.

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

**Note:** Step 2a corresponds to the problem of conditional model specification since our final model is conditioned upon the outcome of the interaction assessment. In order to avoid biasing the Type I error levels of subsequent inferences this preliminary test of interaction should be evaluated at a more liberal $\alpha$–level such as $\alpha=0.15$ or $\alpha=0.25$ . Some additional guidance for choosing a Type I error level to perform a preliminary test of interaction may be found in Bancroft (1964). A general review of the conditional specification problem and associated literature may be found in Bancroft and Han (1977).

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 aids in elucidating their whereabouts. Standard error bars should also be included about the estimated means to indicate the associated variability.

Even though the procedure described above involves the estimation and evaluation of main effects or simple effects, there is truly only one set of comparisons that are of inherent interest to the investigator. The only real distinction is whether or not these comparisons are to be made within levels of another factor (a simple effect comparison) or if the comparisons shall be made after averaging over the levels of another factor (a main effect comparison).

Although the simple two-step procedure outlined above is generally sufficient to analyze many unequally replicated factorial experiments, there are several special notes worthy of mention.

If main effects are to be assessed compositely to determine if the sum-of-squares due to a particular factor should be pooled or not, then some additional guidelines are required. The reader is referred to Table 7.4 of Searle (1971) for such an approach.

If the factorial arrangement of treatments includes a control or zero 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 may 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 df contrast associated with the control treatment should be partitioned from the interaction sums-of-squares. Then 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 manner.

**Step 1′**: Assess the importance of the single df 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 df 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 df and remaining composite tests are important. Proceed to evaluate simple effects in the general means model.

d) The single df 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 should be discussed. The estimated means, $\mu_{ij}$, in the general means model are simply the sample means, $Y_{ij}$. However, in the reduced model this is no longer the case because the interactions are restricted to be zero. Thus, the estimated means $\mu_{ij}$ 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 interaction criterion is met.

## GENERALIZATIONS

The two-step procedure developed above can be generalized. The procedure addresses treatment design, specifically factorial experiments, and therefore is not affected by replication or restrictions on the randomization of the treatments to the experimental units. Experimental design considerations do, of course, affect the structure and magnitude of the experimental error(s) used for model evaluation and standard error calculations.

The number and nature of the treatment factors are important. For example, if there were three factors the assessment of interaction would begin with either (i) the pooled two- and three-way interaction, or (ii) the three-way interaction alone.

The next step in the analysis depends upon the outcome of the assessment chosen above.

If both factors of a two factor experiment are continuous factors with quantitative levels, rather than categorical factors with discrete states, then interest resides in response surface considerations. Thus, contrasts of interest are quantitative contrasts; e.g., the contrast coefficients can be orthogonal polynomial coefficients. With just one factor continuous and one factor categorical, interest of the researcher now turns toward the comparison of curves which, again, can be handled by appropriate contrasts within the recommended two-step procedure.

Introduction of measured covariates into the data analysis can present additional complexity in calculation and interpretation. However, the two-step procedure is easily generalized for estimating adjusted means and contrasts as shown in Snedecor and Cochran (1980), Section 18.4, or Allen and Cady (1982), Unit 19.

## TYPES OF MEANS

For the evaluation of main effects using the reduced model, two alternative methods are available for estimating the row and column means in Step 2a of the recommended procedure.

**Method 1**: Sum all of the observations in a given row (column) and divide by the number of observations in the given row (column). This is equivalent to calculating a weighted average of sample means from the general means model where the weights are the number of replications for each mean. The row and column means calculated by Method 1 are simply called row and column means (MEANS).

**Method 2**: Calculate the unweighted average of cell means for a given row (column) where the cell means are estimated from the reduced model. These row and column means calculated by Method 2 are called the least squares means (LSMEANS; also, see Searle, Speed, and Milliken, 1980) and are estimated from the reduced model where the interactions are restricted to be zero.

## MISSING CELLS: INTERACTION

In the preceding it has been assumed that each $n_{ij}$ was non-zero. Suppose now that some of the treatment combinations have no replication (i.e., $n_{ij}=0$ for some i,j) due to either missing data or lack of interest in the particular treatment combination(s). Strictly speaking, this is no longer a factorial combination of treatments originally envisioned. However, 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 select 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 using the error mean square from the general means model to maximize precision of the tests.

As an example, consider what was originally a 3 × 3 factorial experiment. Suppose that the (1,3) and (3,2) treatment combinations are missing as indicated below:

| | | A 1 | A 2 | A 3 |
|---|---|---|---|---|
| | 1 | $\mu_{11}$ | $\mu_{12}$ | × |
| B | 2 | $\mu_{21}$ | $\mu_{22}$ | $\mu_{23}$ |
| | 3 | $\mu_{31}$ | × | $\mu_{33}$ |

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

| | A | |
|---|---|---|
| | 1 | 2 |
| 1 | $\mu_{11}$ | $\mu_{12}$ |
| 2 | $\mu_{21}$ | $\mu_{22}$ |

| | | A | |
|---|---|---|---|
| | | 1 | 3 |
| B | 2 | $\mu_{21}$ | $\mu_{23}$ |
| | 3 | $\mu_{31}$ | $\mu_{33}$ |

One difficulty should be pointed out. If the same data were used to estimate $\mu_{21}$ in each $2 \times 2$ subset, 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 not assessing useful hypotheses. When a statistical package's default analysis is used on data with missing cells, and interaction is included in the model, then the F-test associated with the line labelled "interaction" is typically testing hypotheses of little use. Thus, the experimenter should define and estimate the contrasts of direct interest.

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 satisfactory results when a convoluted analysis must be performed.

## MISSING CELLS: NO INTERACTION

Surveys and observational studies are often undertaken with the intent to study many factors that are cross-classified. One problem, however, is that there are almost never enough data to even come close to having all cells filled. It is often argued that for lack of anything better such data are analyzed using a main effects only model, provided that the design is connected (Searle, 1971). Due to the unbalanced nature of the data there are k! possible ANOVA tables (i.e., k! sets of sequential sums-of-squares) that may be considered when there are k factors under study. The usual suggestion is that none of these k! are of intrinsic interest and thus focus should be on the F-tests associated with the sum-of-squares for a factor that has been adjusted for all other factors in the model. These are the F-tests associated with the Type II sums-of-squares in SAS®.

However, it is rare for all factors considered initially in a study to be present in the final model from which decisions (inferences) are to be drawn. Many of the factors will act as proxies for other factors and render the suggested F-test of little or no worth. Consequently these factors should be removed before further analysis (viz., inference) is undertaken. This is directly analogous to the full rank multiple regression case where the desire is to assess the need of several predictors simultaneously using the additional sums-of-squares due to these predictors. Factors which act as proxies for other factors can be expected to arise whenever a broad survey is undertaken with post-classification of the responses into levels of many factors. A consultant should have the investigator provide an intelligent (in terms of the subject matter) ordering of the factors and proceed to investigate the corresponding analysis of variance table. Once a reasonable subset of factors has been selected then one can advance to use the suggested F-test. Granted, there are again problems of conditional specification and bias in the level of significance, but one can minimize these problems by performing the preliminary F-test at a liberal Type I error level.

## EXAMPLE

An experiment was conducted on a low nitrogen field soil to determine the effect of growing peas and barley in monoculture versus polyculture. The experiment was a $3 \times 3$ 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 per quadrat.

During the growing season several complications arose which altered the original balanced $3 \times 3$ 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 $2 \times 3$ factorial with unequal replication of the six resultant treatment combinations. In addition, 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 data and computer code used to analyze the experiment are included in the appendix.

The composite test for interaction (Step 3 of the appendix) indicates that interaction is present (p=0.04). Upon fitting the general means model, the following table of predicted treatment means is computed (Steps 4 or 6 of the appendix):

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

Error Mean Square = 123.8685 with 16 df

In order to further elucidate the interaction two additional single df interaction contrasts were examined. The 2 × 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 df associated with the polyculture versus averaged monocultures interaction contrast. Note that these are natural, not orthogonal interaction contrasts which reflect the intent of the researcher's original objectives.

Simple effects are now estimated (Step 8 of the appendix) to assess the difference in yield due to density for each of the cropping systems. For peas and barley in polyculture the yield difference ± the standard error of the yield difference is 36.00 ± 7.87g in favor of the high density (p<0.001); for peas alone this difference is 5.20 ± 9.31g (p=0.58) and for barley alone this difference is 9.43 ± 8.50g (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 ± 6.30g greater for the 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 poly-culture 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 (Steps 10, 11 and 12). 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 |

Error Mean Square = 166.0038 with 18 df

Note that the difference between the rows is the same for each column, 18.588g. Alternatively, note that any contrast among the columns is the same for each row.

## ACKNOWLEDGEMENT

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

## LITERATURE CITED

Allen, D. M., and Cady, F. B. (1982), *Analyzing experimental data by regression,* New York, NY: Van Nostrand Reinhold.

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

Bancroft, T.A., and Han, Chien-Pae. (1977), "Inference Based on Conditional Specification: A Note and a Bibliography," *International Statistical Review,* 45, 117-127.

Buxton, D. (1982), "Instructions to authors," *Agronomy Journal* 74, 1100-1102.

Cochran, W. G., and Cox, G. M. (1957), *Experimental designs, 2nd ed.,* NY: John Wiley and Sons, Inc.

Freund, R.J., Littell, R., and Spector, P.C. (1987), *SAS®System for Linear Models, 1986 Edition,* Cary, NC: SAS Institute, Inc.

Hocking, R.R. (1985), *The Analysis of Linear models,* Monterey, CA: Brooks/Cole Publishing Co.

SAS Institute, Inc. (1985), *SAS user's guide: Statistics, Version 5 Edition,1986 Edition,* Cary, NC: SAS Institute, Inc.

Searle, S. R. (1971), *Linear models,* NY: John Wiley and Sons, Inc.

Searle, S. R. (1987a), *Linear models for Unbalanced Data,* NY: John Wiley and Sons, Inc.

Searle, S. R. (1987b), "Linear models for some-cells-empty data: The cell means formulation, a consultant's best friend," in *Design, Data, and Analysis: by some Friends of Cuthbert Daniel,* Editor, C.L. Mallows, New York: John Wiley and Sons.

Searle, S. R., Speed, F. M., and G. A. Milliken. (1980), "Population Marginal Means in the Linear Model: An Alternative to Least Squares Means," *The American Statistician* 34, 216-221.

Snedecor, G. W., and W. G. Cochran. (1980), *Statistical Methods, 7th ed.*, Ames, IA: The Iowa State University Press.

Speed, F. M., R.R. Hocking, and Hackney, O. P. (1978), "Methods of Analysis of Linear Models with Unbalanced Data," *Journal of the American Statistical Association* 73, 105-112.

Speed, F. M., and Monlezun, C.J. (1979), "Exact F-tests for the Method of Unweighted Means in a $2^k$ Experiment," *The American Statistician* 33, 15-18.

Urquhart, N. S., and Weeks, D. L. (1978), "Linear Models in Messy Data: Some Problems and Alternatives," *Biometrics* 34, 696-705.

Warren, W.G. (1979), "Analysis and Interpretation of an Experiment with a Heterogeneous Mixture of Treatment Types," *Biometrics* 35, 869-872.

## APPENDIX

The SAS® (Statistical Analysis System, 1985) program used to analyze the data in the example is given below. The same analysis could have been handled by other statistical computing packages including those for microcomputers. The analysis could have also been performed by defining the six treatment combinations and fitting a one-way classification model with appropriate contrasts corresponding to the factorial structure of the six treatments. The necessary requirements include those mentioned in the Introduction. Selected annotations follow the program.

```
DATA CROP;
INPUT SYSTEM $ DENSITY $ TMT YIELD;
CARDS;
P L 1 89.13
P L 1 76.78
P H 4 109.67
P H 4 89.67
P H 4 75.44
P H 4 89.60
P H 4 76.39
```

```
B L 2  75.59
B L 2  70.63
B L 2  60.43
B H 5  80.80        [1]
B H 5  77.45
B H 5  79.05
B H 5  75.96
X L 3  88.28
X L 3 104.50
X L 3  84.90
X L 3  88.97
X H 6 125.44
X H 6 128.96
X H 6 108.51
X H 6 147.74
PROC PRINT;
PROC PLOT DATA=CROP;
     PLOT YIELD*SYSTEM=DENSITY;    [2]
     PLOT YIELD*DENSITY=SYSTEM;
PROC GLM; CLASSES SYSTEM DENSITY;              [3]
     MODEL  YIELD = DENSITY SYSTEM DENSITY*SYSTEM;
     LSMEANS DENSITY SYSTEM DENSITY*SYSTEM / STDERR;  [4]
     MEANS   DENSITY SYSTEM DENSITY*SYSTEM / DEPONLY;
     OUTPUT OUT=NEW1 PREDICTED=YHAT1 RESIDUAL=RESID1;
PROC PLOT;                          [5]
     PLOT RESID1*YHAT1 / VREF=0;
PROC GLM; CLASSES DENSITY SYSTEM;              [6]
MODEL  YIELD = DENSITY*SYSTEM / NOINT;
     ESTIMATE 'MONO * DENSITY'
     DENSITY*SYSTEM 1 -1 0 -1 1 0;
     ESTIMATE 'MONOPOLY * DENSITY'               [7]
     DENSITY*SYSTEM -1 -1 2 1 1 -2 / DIVISOR=2;
     ESTIMATE 'DENSITY W/I POLY'
     DENSITY*SYSTEM 0 0 1 0 0 -1;
     ESTIMATE 'DENSITY W/I PEAS'               [8]
     DENSITY*SYSTEM 0 1 0 0 -1 0;
     ESTIMATE 'DENSITY W/I BARLEY'
     DENSITY*SYSTEM 1 0 0 -1 0 0;
     ESTIMATE 'PEAS VS BARLEY FOR MONO'
     DENSITY*SYSTEM -1 1 0 -1 1 0 / DIVISOR=2;
```

```
        ESTIMATE 'DENSITY FOR MONO'                [9]
        DENSITY*SYSTEM 1 1 0 -1 -1 0 / DIVISOR=2;
PROC GLM;  CLASSES DENSITY SYSTEM;
        MODEL  YIELD = DENSITY SYSTEM / P CLM;     [10]
        LSMEANS DENSITY SYSTEM / 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;
PROC PLOT;  PLOT RESID2*YHAT2 / VREF=0;            [12]
```

**Annotations:**

1. P = peas monoculture, B = barley monoculture, X = peas and barley 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 DENSITY*SYSTEM term.
4. The LSMEANS are the unweighted means whereas the MEANS give the weighted means as discussed in the section on Types of Means. The cell means are the same for both MEANS and LSMEANS in the full model when there are no missing cells.
5. Residual plot for the full model.
6. Fitting the general means model. This could also have been accomplished by using PROC GLM; CLASS TMT; MODEL YIELD = TMT; and constructing the ESTIMATE statements accordingly.
7. The two "natural" interaction single df contrasts. In general, these will not be orthogonal. These contrasts assess the interaction between monoculture and density, and between mono- vs. polyculture and density.
8. These are the single df 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 $2 \times 2$ factorial of density by monocultures.
10. Fitting the reduced model with interaction defined to be zero. The P and CLM options print the predicted cell means and a 95% confidence interval for each observation. Cell means can be gleaned from the output of the P option. The

LSMEANS are those discussed in the section on Types of Means.

11. Main effect contrasts for cropping systems and density with no interaction.
12. Residual plot for the reduced model.