A METHOD FOR HYPOTHESIS TESTING IN TWO-WAY, CROSS-CLASSIFIED DESIGN WITH MISSING CELLS

Mattana Sanguanruang and Kenneth L. Koonce
Louisiana State University

ABSTRACT

Missing cells in a two-way, cross-classified fixed effects model with interaction create problems in the estimation of treatment means and in the tests of hypotheses regarding main effects. This paper presents an alternative method for analyzing such data in two situations: (1) unobserved but existent treatment combinations, and (2) nonexistent populations. For unobserved treatment combinations, modified marginal means can be used to test hypotheses about mean parameters. For nonexistent populations, linear relationships of means of all filled cells should be constructed, and the hypotheses tested according to the purposes of the researcher. This alternative method provides a rationale for selecting meaningful hypotheses when missing cells occur rather than depending upon the standard computer packages, since the result of this proposed method is based on the reason for the missing cells and not on computational conveniences. In this paper, the full rank cell means model is presented as a tool for evaluating the hypotheses being tested.

INTRODUCTION

In the analysis of linear models for designed experiments with balanced data, there is general agreement on the appropriate analysis of variance (ANOVA) table. But in the analysis of cross-classified data with missing cells, the appropriate ANOVA table has been a well-recognized problem for many years. One of the difficulties encountered in the missing cells problem is the most common non-estimable conditions such as

(i) $\alpha_{\cdot} = \beta_{\cdot} = (\alpha\beta)_{i\cdot} = (\alpha\beta)_{\cdot j} = 0$, or

(ii) $\sum_i n_{i\cdot}\alpha_i = \sum_j n_{\cdot j}\beta_j = \sum_{i \text{ or } j} n_{ij}(\alpha\beta)_{ij} = 0$

are not sufficient to give a full rank model because there are more parameters than can uniquely estimated from sample data without assumptions about the relationships among the parameters. This situation, referred to as "overparameterization" causes problem concerned with estimability, so the "usual" tests of hypotheses of main effects cannot be carried out unless one is willing to make nontestable assumptions about the model. Although, a number of specialized methodologies have been developed to solve these problems, there is no single universally acceptable methodology for handling missing cells. In more recent years, the availability of statistical software packages has permitted increasing use of general linear model methodology to test the hypotheses associated with unbalanced data, for example, SAS, SPSS, and BMDP. With these major packages, the user can obtain any number of ANOVA tables by specifying different options and/or changing the order of terms in the model. This flexibility has left many users quite confused as to which set of sums of squares, if any, is appropriate. The SAS GLM procedure, one of the most widely used packages, offers the user four different types of sums of squares. Type III and type IV sums of squares can handle missing cells but only in certain cases. However, for convenience, users frequently calculate F-statistics associated with missing data without understanding which hypotheses are being tested. Not only is this logically backwards, but the hypothesis is often neither of use nor of interest.

The purpose of this paper is to examine simple procedure that may be used with missing cells in two situations: (1) the missing cells occurred by chance and not as a result of the "main" or "interaction" effects involved (unobserved but existent treatment combinations), and (2) the apparent loss occurred as a result of the "main" or "interaction" effects involved (nonexistent populations). In addition, a criterion is provided for obtaining the reasonable hypotheses. The cell means model is presented as a tool for evaluating hypotheses being tested in terms of linear combinations of population cell means. In this paper we confine attention to the two-way, cross-classified model, 3 x 3 factorial arrangement with interactions and missing cells.

THE CELL MEANS MODEL

Over the years, statisticians have tried to solve problems with the overparameterization of the model. Hocking and Speed (1975) developed the "cell means" model approach which yields a full rank model. The cell means model is an approach to the analysis of linear models in which it is assumed that each observation is drawn from a separate population having its own mean and variance. By utilization of this model, the population means are the parameters of interest and hypothesis testing is defined in terms of those mean parameters.

For a two-way, cross-classified fixed effects model with interaction, the cell means model would be:

$$Y_{ijk} = \mu_{ij} + \varepsilon_{ijk} \qquad (1)$$

where

Y_{ijk} = value for dependent variable for kth observation from ith level of A and jth level of B

μ_{ij} = population cell mean of the ijth combination of AB

ε_{ijk} = random error assumed NID$(0,\sigma^2)$

The parameter of the cell means model relate to the parameters of the standard analysis of variance model according to the equation:

$$\mu_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} \qquad (2)$$

where

μ = overall mean
α_i = fixed effect of ith level of A
β_j = fixed effect of jth level of B
$(\alpha\beta)_{ij}$ = fixed interaction effect of AB

In matrix notation the cell means model may be written as:

$$Y = W\mu + \varepsilon \quad (3)$$

subject to the restrictions,

$$G\mu = 0 \quad (4)$$

where

Y = n x 1 vector of observations
W = n x (axb) design matrix of ones and zeros denoting from which population the sample was taken
μ = (axb) x 1 vector of unknown parameters
ε = n x 1 vector of random errors assumed to be distributed $(0,\sigma^2 I)$
G = r x (axb) constraint matrix of rank r displaying some assumed or known relationship between population parameters, for example, no interaction effect.

These restrictions (G) are not always present. Typically, they reflect the assumption of no interaction in the model, but other types of restrictions may arise. We emphasize the fact that the restrictions represent assumptions about the mean μ_{ij} which may be based on the researcher's knowledge of the experimental situation. When no restriction is imposed on μ, then $G \equiv 0$.

The least squares solution of the normal equations is direct. The W'W matrix is diagonal and of full rank with the elements being the number of times the ijth population was sampled; therefore, the $(W'W)^{-1}$ matrix exists. The least squares normal equations are given by

$$(W'W)\mu = W'Y \quad (5)$$

where

μ = (axb) x 1 vector of unknowns
$(W'W)$ = (axb) x (axb) incidence matrix
$(W'Y)$ = (axb) x 1 vector of cross product of totals.

ADVANTAGES OF CELL MEANS MODEL

1. An investigator is led to consider hypotheses about relations among the μ_{ij}, whereas the overparameterized model encourages consideration of the α_i, β_j, and $(\alpha\beta)_{ij}$. Jennings and Ward (1982) believed that it is much easier to generate useful functions of the μ_{ij} than it is to generate useful estimable functions of the parameters of overparameterized model.

2. The cell means model avoids the imposition of non-estimable conditions on the population parameters as required by the overparameterized model.

3. There is no problem with the concept of estimability since the cell means model is a full rank model. Every μ_{ij} and any linear function of such μ_{ij} are estimable if that population is observed (Hocking and Speed 1979). In addition, if the model is constrained, we may be able to obtain estimates even if the population is not directly observed. This allows the researcher to test any linear hypothesis about the cell means of interest.

4. The hypotheses must be specified by the user, but of course, certain "standard" hypotheses could be "pre-programmed" (Hocking and Speed 1979).

5. The interpretation of the main effects and interactions are more meaningful since they are from population means (Casanova 1981).

6. There is no complex computation when dealing with missing cells.

Because of these advantages, the cell means model is used here as a tool for developing sums of squares for linear models with missing cells.

HYPOTHESIS TESTING

The theoretical basis of hypothesis testing is quite simple. If the model is stated in terms of a parameter vector μ and an hypothesis of the form $H\mu = 0$ is proposed, then the test statistic is given by

$$F(q,n-p) = SS(H\mu=0|G\mu=0)/(qs^2) \quad (6)$$

where

q = rank of H
s^2 = residual mean square with n-p degree of freedom for the model.

The numerator sum of squares SS(H) is computed according to Bryce et al.(1980), and Hocking and Speed (1979):

$$SS(H\mu=0|G\mu=0) = (H\mu)'[HA(W'W)^{-1}A'H']^{-1}(H\mu) \quad (7)$$

where

$$\mu = A(W'W)^{-1}W'Y \quad (8)$$

when $G \equiv 0$

$$A = I - (W'W)^{-1}G'[G(W'W)^{-1}G']^{-1}G \quad (9)$$

when $G \neq 0$

$$A = I \quad (10)$$

The confusion that currently exists among the various computer programs arises from the fact that the hypothesis matrix H is often not specified and SS(H) is not computed directly from the hypothesis. On the contrary, a quantity to be used for the numerator sum of squares is arrived at as a consequence of a particular computational procedure. There is no objection to this approach if the effective hypothesis is known and reasonable. Frequently, it is not the case and hence the test statistic is of dubious value. Thus, a researcher should first specify a hypothesis which is of interest, and then calculate the corresponding sum of squares.

Hypotheses that usually interest researchers in cross-classified design are those concerning main effects and the interaction effect. Nine hypotheses that are commonly used for testing the main effects and interaction effect

Table 1. Hypotheses in the two-way, cross-classified model.
(adopted from Speed et al. 1978)

Hypotheses	
Row Effects	
H_1 :	$\bar{\mu}_{i.} = \bar{\mu}_{i'.}$
H_2 :	$\sum_j n_{ij}\mu_{ij}/n_{i.} = \sum_j n_{i'j}\mu_{i'j}/n_{i'.}$
H_3 :	$\sum_j (n_{ij})\mu_{ij} = \sum\sum_{i'j} n_{ij}n_{i'j}\mu_{i'j}/n_{.j}$
H_4 :	$\mu_{ij} = \mu_{i'j}$
Column Effects	
H_5 :	$\bar{\mu}_{.j} = \bar{\mu}_{.j'}$
H_6 :	$\sum_i n_{ij}\mu_{ij}/n_{.j} = \sum_i n_{ij'}\mu_{ij'}/n_{.j'}$
H_7 :	$\sum_i (n_{ij})\mu_{ij} = \sum\sum_{ij'} n_{ij}n_{ij'}\mu_{ij'}/n_{i.}$
H_8 :	$\mu_{ij} = \mu_{ij'}$
Interaction Effect	
H_9 :	$\mu_{ij}-\mu_{i'j}-\mu_{ij'}+\mu_{i'j'} = 0$

are shown in Table 1 (Speed et al. 1978).

These nine hypotheses will be related to the different methods. A typical method might refer to either of H_1, H_2, H_3, or H_4 as "main effect A." That is, the associated sums of squares are used to measure, in some sense, the effect of factor A. Inspection of these hypotheses reveals that the sense in which this effect is measured is different for the four hypotheses. Hypotheses H_1, H_2, and H_3 agree if $n_{ij} = n$ for all i and j but not otherwise. The hypothesis H_1 is quite easy to interpret as it states that there is no difference in the levels of factor A when averaged over all levels of factor B. The hypotheses H_2 and H_3 are not easy to understand as they represent comparisons of weighted averages with the weights being a function of the cell frequencies. The hypothesis H_4 is distinctly different from the first three as it is concerned only with comparisons of the cell means in the specific column. While H_4 is easy to understand, it is not of general interest, unless we are willing to assume that there is no interaction effect:

$$\mu_{ij} - \mu_{i'j} = \mu_{ij'} - \mu_{i'j'} \qquad (11)$$

for all i, i', j, and j'. The hypotheses H_5-H_8 are the counterparts of H_1-H_4 generally associated with "main effect B." The hypothesis of no interaction is H_9.

THE MISSING CELLS SITUATIONS

Whether the incomplete data set is due to unobserved treatment combinations or nonexistent populations, serious problems with estimability are created, and the "usual" tests of hypotheses of main effects cannot be used. Casanova (1981) stated some examples of missing cells resulting from unobserved treatment combinations:

(1) in agriculture, a particular crop or field in a study that may be damaged or destroyed by flooding,

(2) in dairy science, animals may become very ill or possibly die due to diseases before the completion of the experiment,

(3) in social science, human subjects may be absent when scheduled to participate in an experiment, and

(4) equipment malfunction such as in fisheries a net may break on a cast at a particular location.

Missing cells may be caused by nonexistent populations due to problems involving the "main" effects or "interaction" effect. Some examples of missing cells resulting from nonexistent populations are:

(1) in agriculture, a particular species of plants in a study dies at a particular level of fertilizer. In this case the researcher should use 3x2 factorial arrangement,

A \ B	1	2	3
1	μ_{11}	μ_{12}	X
2	μ_{21}	μ_{22}	X
3	μ_{31}	μ_{32}	X

(2) in dairy science, suppose A and B are two nutrients, and at least one of which is necessary to animals, then animals at (0,0) level would be missing,

A \ B	0	1	2
0	X	μ_{01}	μ_{02}
1	μ_{10}	μ_{11}	μ_{12}
2	μ_{20}	μ_{21}	μ_{22}

(3) in marine biology, interaction effect of 2 factors such as temperature and salinity can cause oysters to die at a particular level (Garton and Koonce 1982),

Temp \ Salinity	1	2	3	4
1	μ_{11}	μ_{12}	μ_{13}	X
2	μ_{21}	μ_{22}	μ_{23}	μ_{24}

(4) in animal physiology, suppose we have 2 levels of treatment A and 2 levels of treatment B, and we want to compare the body weight increases within each treatment and to compare increases between treatment groups and control group. Researchers usually use 2x2 factorial arrangement and one cell of control instead of using 3x3 factorial with 6 missing cells.

		B		
		1	2	0
A	1	μ_{11}	μ_{12}	X
	2	μ_{21}	μ_{22}	X
	0	X	X	μ_{00}

When one or more cells are missing in the cross-classified situation and the model contains interactions, there appear to be two primary issues to be considered (Hocking et al. 1980):

(1) Is there enough data to estimate all of the parameters?

(2) What general hypotheses should be tested in the preliminary ANOVA?

The first question is related to the concept of connectedness in the two-way ANOVA (Hocking et al. 1980). In general, we say that a design is connected if all cell means are estimable. Note this will be the case if all cells are filled but may also be true with empty cells if sufficient relations are assumed on the means, e.g. lack of certain interactions.

The question of ANOVA hypotheses with missing cells has no clear answer because some of the means are non-estimable. The common ANOVA hypotheses as shown in Table 1 are not valid if some $n_{ij} = 0$ unless some assumptions about the model are made. Statisticians handle this problem differently. Yates (1934) developed a formula for estimating the missing cells. This does not supply additional information to the experiment, it only facilitates the analysis of the remaining data. One degree of freedom is removed from error for each missing cell. If there is more than one missing cell, then an iterative process is required until all estimates are stabilized. In order to estimate these cells, the assumption of additivity must be made which places a restriction on the model. This technique can be used only if the cause of missing cells results from unobserved treatment combinations, not from the nonexistent populations.

When cells are missing, under certain models, GLM does not give an estimate of the least squares or marginal means; instead the user is told they are non-estimable (NON-EST). Although GLM will not estimate some of the means, four types of sums of squares may be obtained along with their expectations (E1, E2, E3, E4 options). Koonce and Speed (1980) stated that these estimable functions are not always satisfactory to the user since most hypotheses are formulated as mean parameters and interpretation of estimable functions when cells are missing is apparently poorly understood by many users. GLM Type III and PROC HARVEY handle the missing cell problem by assuming the interaction in the missing cell is zero, whereas interaction may occur in all other cells. This generally leads to estimates of means, however, their validity depends on the true lack of interaction in that cell in the population (Koonce and Speed 1980 and Harvey 1977). GLM Type IV sums of squares are computed without making assumption of no interaction, but rather try to obtain a full or balanced set of data. This may result in deleting data from a factor level which contains the missing cell. Therefore, with the same set of data, Type IV sums of squares can vary by altering the location of the missing cells (Goodnight 197B and Freund 1980).

Hocking et al. (1980) proposed that a reasonable hypothesis to be tested when there are missing cells should be based on the premise that the researcher has in mind hypotheses that are appropriate if all cells are filled.

Searle et al. (1980) modified the least squares means for the missing cell situation. That modified least squares mean is called "modified marginal mean" (MM^*) which is a linear combination of the appropriate cell means. For row i it would be an average of the cell means for cells for which μ_{ij} is estimable, and could be formally expressed as:

$$MM^*(\alpha_i) = \Sigma \delta_{ij}\mu_{ij} / \Sigma \delta_{ij} \qquad (12)$$

where

$\delta_{ij} = 1$ if μ_{ij} is estimable

$\delta_{ij} = 0$ if μ_{ij} is non-estimable.

In all cases MM^* will be estimable, with best linear unbiased estimate (b.l.u.e).

$$\widehat{MM^*(\alpha_i)} = \sum_{j=1}^{a} \delta_{ij}\hat{\mu}_{ij} / \sum_{j=1}^{a} \delta_{ij} \qquad (13)$$

Similarly the estimated MM^* of column j would be

$$\widehat{MM^*(\beta_j)} = \sum_{i=1}^{b} \delta_{ij}\hat{\mu}_{ij} / \sum_{i=1}^{b} \delta_{ij} \qquad (14)$$

THE EXAMPLE

The concepts developed in the preceding discussion were applied to a data base. This study was designed to compare the offspring of three sires for average daily gains on three rations. There are 50 steers available to the researcher and he randomly allocates the rations to steers from each sire so that there are not an equal number of steers of each sire being fed each ration. Due to logistical problems, the cells 12 and 31 were not sampled, so the experimental design is a 3 x 3 factorial with 2 missing cells. The layout of the data is given in Table 2.

Case 1. Unobserved treatment combinations

The results of least squares means (LSM) from GLM and the modified marginal means (MM^*) are shown in Table 3. The modified marginal mean will give "some" estimate which would appear to be reasonable when no best linear unbiased estimate (b.l.u.e.) exists due to missing cells.

The modified marginal mean can be used to test hypotheses about mean parameters. The hypothesis for differences between levels of A

would be $H_0 : \mu_{1.} = \mu_{2.} = \mu_{3.}$ and hypothesis for difference between levels of B would be $H_0 : \mu_{.1} = \mu_{.2} = \mu_{.3}$.

Table 2. Layout of experimental design showing factor levels, location of missing cells.

		B		
		1	2	3
A	1	μ_{11}	X	μ_{13}
	2	μ_{21}	μ_{22}	μ_{23}
	3	X	μ_{32}	μ_{33}

Factor A = sire of steers
Factor B = ration

Table 3. Results of least squares means (LSM) from GLM* and the modified marginal means (MM).

Mean parameter	LSM	MM*
Overall	-	0.53
A_1	NON-EST	0.64
A_2	0.50	0.50
A_3	NON-EST	0.47
B_1	NON-EST	1.17
B_2	NON-EST	0.60
B3	0.06	0.06

Results of sums of squares obtained for main effects and interactions under the different methods are given in Table 4.

Table 4. Sums of squares for main effects and interactions for missing cell example for three different computational methods.

Source	GLM Type III	GLM Type IV	Modified Marginal Mean
A	0.402	0.396	0.204
B	B.685	8.742	8.217
AB	0.263	0.263	0.263

Only the A x B interaction sum of squares is the same for all three computational methods because these three methods test the same hypotheses, $H_0 : \mu_{11} - \mu_{21} = \mu_{13} - \mu_{23}$ and $\mu_{22} - \mu_{32} = \mu_{23} - \mu_{33}$. The main effect sum of squares vary. An examination of the cell means actually compared by each method sum of squares is given in Table 5.

Case 2. Nonexistent populations

When the missing cells result from nonexistent populations, test of main effects and interaction effect are no longer meaningful. The linear relationship among means of all filled cells should be constructed, and hypotheses that interest the researcher should be tested. For example, a researcher wants to test if the mean of A1 population means is equal to the mean of the A2 population means and at the same time the mean of the B1 population means is equal to the mean of the B2 population means, the hypotheses could be:

$$3\mu_{11} + 3\mu_{13} = 2\mu_{21} + 2\mu_{22} + 2\mu_{23}$$

$$\mu_{11} + \mu_{21} = \mu_{22} + \mu_{32}$$

Another example of missing cells result from nonexistent populations occurs when a researcher wants to compare the body weight increases among levels of two treatments, and also to make comparisons between the treatment group and the control group. For example, he wants to test if the mean of the A2 population means is equal to the mean of the A3 population means and the mean of the B1 population means is equal to the mean of the B3 population means, and at the same time the mean of the treatments means is equal to the mean of the control population means, hypotheses could be:

$$\mu_{21} + \mu_{22} + \mu_{23} = \mu_{31} + \mu_{32} + \mu_{33}$$

$$\mu_{11} + \mu_{21} + \mu_{31} = \mu_{13} + \mu_{23} + \mu_{33}$$

$$\mu_{11} + \mu_{12} + \mu_{13} + \mu_{21} + \mu_{22} + \mu_{23} + \mu_{31} + \mu_{32} + \mu_{33} = 9\mu_{00}$$

This alternative method provides the researcher reasonable results because he can test any hypothesis that interests him and all cell means are estimable.

Table 5. Hypotheses tested for A and B main effects for three different methods of calculation of sums of squares for a 3x3 cross-classified model with missing cells.

GLM Type III

A: $\mu_{11}+2\mu_{13}+\mu_{22} = \mu_{21}+\mu_{32}+2\mu_{33}$

$2\mu_{11}+7\mu_{32}+8\mu_{33} = 2\mu_{13}+2\mu_{21}+7\mu_{22}+6\mu_{23}$

B: $7\mu_{11}+8\mu_{21}+2\mu_{32} = 7\mu_{13}+2\mu_{22}+6\mu_{23}+2\mu_{33}$

$2\mu_{11}+8\mu_{22}+7\mu_{32} = 2\mu_{13}+2\mu_{21}+6\mu_{23}+7\mu_{33}$

GLM Type IV

A: $\mu_{13} = \mu_{33}$

$\mu_{22}+\mu_{23} = \mu_{32}+\mu_{33}$

B: $\mu_{11}+\mu_{21} = \mu_{13}+\mu_{23}$

$\mu_{22}+\mu_{32} = \mu_{23}+\mu_{33}$

Modified Marginal Mean

A: $3\mu_{11}+3\mu_{13} = 2\mu_{21}+2\mu_{22}+2\mu_{23}$

$\mu_{11}+\mu_{13} = \mu_{32}+\mu_{33}$

B: $\mu_{11}+\mu_{21} = \mu_{22}+\mu_{32}$

$3\mu_{11}+3\mu_{21} = 2\mu_{13}+2\mu_{23}+2\mu_{33}$

DISCUSSION

Problems encountered when cells are missing in the two way cross-classified situation where interaction exists are considered in this paper. The cell means model has been shown to be an effective tool for testing the hypotheses involving missing cells since the cell means model avoids the confusion regarding estimability and estimable functions. Also the user need not work in the overparameterized model and can be concerned only with which hypotheses being tested. In particular, this model simplifies interpretation of parameters and clarifies hypothesis statements. Finally, it forces the user to specify hypotheses to be tested.

Two cases of missing cells are considered in this paper: (1) unobserved treatment combinations, and (2) nonexistent populations. There are several ways to test hypotheses with missing cells. The widely used methods are GLM Type III sum of squares and PROC HARVEY, which draw inferences concerning "main effects" by assuming that the interaction effects are zero for the missing cells. This assumption will often be unrealistic because the interaction effects are really not known. Another method is GLM Type IV which computes sum of squares without making the assumption of no interaction, but tries to obtain a full or balanced set of data. But altering the location of the missing cells produces different results (Goodnight 1978 and Freund 1980). There are several statisticians who have approached the problems with missing cells such as Speed and Hocking (1976), Harvey (1977), Goodnight (1980), Henderson and McAllister (1978), Hocking and Speed (1979), Freund (1980), Searle (1980), Jennings and Ward (1982). But most of the methods assume that the loss of cells occurred by chance, and the populations do exist.

In this paper, we examine both cases. For unobserved treatment combinations, modified marginal means are used to test the hypotheses which are associated with the hypotheses that the researcher has in mind for the all filled cells case. For those hypotheses which started out with filled cells, but for which observations were lost due to external forces, there may be no reason to alter the hypotheses. For nonexistent populations, linear relationships of means of all filled cell have been constructed, and the hypotheses can be tested according to the purpose of researcher.

This alternative method provides a rationale for selecting meaningful hypotheses when missing cells occur rather than depending upon the standard computer packages, since the results of this proposed method is based on the reason for the missing cells without making any assumptions about possible main effects or interaction effect. The concept used with this method is defined in terms of those parameters that can be estimated based on filled cells.

BIBLIOGRAPHY

Bryce, G.R., D.T. Scott and M.W. Carter. 1980. Commun. Statist., A9, 131-150.

Casanova, D.L. 1981. The 6th Ann. SUGI Conf., 191-194.

Freund, R.J. 1980. The Amer. Statist., 34, 94-98.

Garton, D.W. and K.L. Koonce. 1982. The 6th Ann. SUGI Conf., 185-190.

Goodnight, J.H. 1980. Commun. Statist., A9, 167-180.

Harvey, W.R. 1977. Proc. Stat. Comp. Sec., Amer. Statist. Assoc., 22-26.

Henderson, C.R. and A.J. McAllister. 1978. J. Animal Science, 46, 1125-1137.

Hocking, R.R. and F.M. Speed. 1975. J. Amer. Statist. Assoc., 70 (351), 709-712.

Hocking, R.R. and F.M. Speed. 1979. The 35th Ann. Conf. Appl. Statist..

Hocking, R.R., F.M. Speed and A.T. Coleman. 1980. Commun. Statist., A9, 117-129.

Jennings, E. and J.H. Ward, Jr. 1982. The Amer. Statist., 36, 25-27.

Koonce, K.L. and F.M. Speed. 1980. The 5th Ann. SUGI Conf., 230-235.

Sas Institute, Inc. 1985. SAS/IML Guide for Personal Computers, Cary, NC.

Searle, S.R. 1980. Commun. Statist., A9, 181-200.

Searle, S.R., F.M. Speed and G.A. Milleken. 1980. The Amer. Statist., 34, 216-221.

Speed, F.M. and Hocking, R.R. 1976. The Amer. Statist., 30, 30-33.

Speed, F.M., R.R. Hocking and O.P. Hackney. 1978. J. Amer. Statist. Assoc., 73, 105-112.

Yates, F. 1934. J. Amer. Statist. Assoc., 29, 52-66.