A STUDY OF FOUR METHODS OF COMPUTING ANALYSIS OF VARIANCE ON A TWO-WAY DESIGN FIXED-MODEL WITH DISPROPORTIONATE CELL FREQUENCIES

DISSERTATION

Presented to the Graduate Council of the
North Texas State University in Partial
Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

By

Kenneth U. Black, B.A., M.A.

Denton, Texas

August, 1978

wKb

Black, Kenneth U., *A Study of Four Methods of Computing Analysis of Variance on a Two-Way Design Fixed-Model with Disproportionate Cell Frequencies*. Doctor of Philosophy (Educational Research), August, 1978, 123 pp., 15 tables, 1 illustration, bibliography, 39 titles.

This study sought to determine the effect of varying degrees of disproportionality of four methods of handling disproportionality cell frequencies in two-way analysis of variance. A Monte Carlo simulation procedure was employed. Two multiple linear regression techniques and two "approximate" techniques were compared.

In each case 1000 F values were calculated for each method under eleven levels of disproportionality. Forty numbers per run were used for each design. Probability distributions of F values for the four methods were compared to an equal cell method by using Kolmogorov-Smirnov tests. Chi-square values were used to measure disproportionality.

Five cases were examined: the no effects case, the row effects case, the column effects case, the interaction effects case, and the row and column effects case. These effects were generated through the use of noncentral F distributions. The cases were used to provide information concerning Type I and Type II errors.

In each case, several seed numbers and their effects on results were examined. In cases with built-in effects, power values were changed in order to examine the effect of power on the results.

Several conclusions were reached within the given parameters of this study. For small levels of disproportionality, all four methods will yield similar nonspurious results. For moderate levels of disproportionality, the complete linear-model regression method and the unweighted means analysis committed fewer Type I errors; and the method of expected frequencies committed fewer Type II errors. For extreme levels of disproportionality, all four methods yielded spurious results. The complete linear-model regression methods and the unweighted means analysis produced similar results at all levels.

TABLE OF CONTENTS

LIST OF TABLES

Table		Page

LIST OF ILLUSTRATIONS

CHAPTER I

INTRODUCTION

In using a factorial analysis of variance design, "one of three possible situations can exist with respect to the numbers of observations within the various cells of the design." "The n's can be (a) equal to one another, (b) unequal but proportional, or (c) unequal and disproportional" (9, p. 281). Traditional analysis of variance techniques can be used when cell memberships are equal in number or if cells have proportionate n's. Tsao (23, p. 195) says that with analysis of variance, "the applicable equations are generally concerned with the case of equal or proportionate numbers of observations in the subclasses." However, if cell memberships (n's) are disproportionate, traditional methods fail.

Roscoe (19, p. 348) says that with disproportionality of cell membership, it is impossible to partition the sum of squares for total into independent and nonoverlapping sums of squares. Mood (14, p. 358) underscores that by saying that "when cell frequencies are not equal, . . . tests become nonorthogonal so that simple successive partition of the total sum of squares is no longer possible." Snedecor (21, p. 285) states "another startling characteristic of disproportionality in a two-way table is the failure of the addition theorem for sums of squares."

Ostle states that the various sums of squares calculated in the usual fashion do not sum up to agree with the total sum of squares (15, p. 381). He goes on to say that this causes the different comparisons with which the sums of squares are associated to be nonorthogonal which leads to biased test procedures. He adds that simple treatment means are biased estimates of the true effects and that serious errors can be made if inferences are made based on such biased estimates (15). Wert, Neidt, and Almann (24, p. 211) write:

> The ordinary methods of computing the analysis of variance with multiple classification are applicable only when the number of cases in the subclasses are proportional. When disproportionality exists among the subclasses, ordinary methods of computation of the sums of squares yield biased results for all sources of variation except that for within subgroups.

Kendall (11, p. 220) further states that when disproportionate numbers in the subclass, the row and column effects are no longer independent; and thus, they cannot be summed and subtracted from the total to get a residual or interaction term to be used as an unbiased estimator.

Anderson and Bancroft (1, p. 278) say that if subclass numbers are not proportional, row, column, and interaction effects are confounded. Snedecor (21, p. 285) says that, clearly, no proper estimate of interaction can be given. "In fact, all estimates and tests of significance may be biased by the disproportion of subclass numbers, and the appropriate statistical methods are thereby complicated" (21, p. 285). Scheffe' states that tests for interactions are more difficult to compute in the case of unequal numbers of observations in the cells (20, p. 112). Overall and Klett (16, p. 445) state "the effect of unequal and disproportionate cell frequencies is to introduce correlation between columns of the design matrix."

Ostle (15, p. 381) relates that disproportionality would lead to biased test procedures unless some adjustment were made. Marks (13, p. 351) says that difficulties arise in interpreting results of unbalanced data analyses because the estimatable functions involved in the tests of hypotheses are not orthogonal. Dixon and Massey (6, p. 134) state that the analysis of variance must be modified for disproportionate numbers of measurements in cells.

In many areas of research, disproportionate cell numbers in two-way analysis of variance occur. Tsao says that "in fields connected with human beings such as education and psychology, unequal representation in each cell of the multiple-classification of data is of common occurrence" (22, p. 107). Johnson and Jackson (10, p. 234) state "unfortunately, in the social sciences the appearance of unequal subclass numbers is the rule rather than the exception." Cohen (4, p. 426) says that in nonexperimental research, it frequently occurs that some subjects are missing data on one or more of the independent variables under study. Bessent (2, p. 1) says, "an unequal number of observations in subgroups (unbalanced data) is the rule rather than the exception for experiments in some areas of research, especially the social and biological sciences."

There are several possible reasons why disproportionate cell frequencies occur. Subjects may fail to appear for all or part of an experiment and must therefore be excluded from the data analysis. With the variables being manipulated or observed, different sample sizes may occur naturally (classroom A may be larger than classroom B). An experimenter might purposefully use an unbalanced design to represent variables in their natural, correlated state. In general, field samples lead to unequal n's or unbalanced designs (7, p. 132). Proger (18, p. 2) mentions three reasons why disproportionality might occur: (1) there is an inherent dearth of some types of subjects, (2) there is inadvertant experimental mortality, and (3) there is forced experimental mortality (some subjects who are inappropriate are dropped). Cochran and Cox (3, p. 72) say that some reasons for misssing data might include: failure to record, gross errors in recording, and accidents. Keppel suggests that unequal sample sizes may result from subjects failing to complete the experimental sequence due to illness or a conflicting appointment (12, p. 77). He says that sometimes studies may require subjects to reach a performance criterion. Those that fail are eliminated from the experiment (12, p. 78).

It would seem that the field of education is one area that is especially prone to the possibility of disproportionality of cell membership occurring. Proger (18, p. 2) says that in large public school situations, unequal n's are the rule in using analysis of variance. Furthermore, Tsao (23, p. 195) says that differences between grades or schools are almost always going to be different sizes unless subsamples are taken.

There are several ways mentioned in the literature of handling this problem. Among these techniques are "approximate solutions" and "regression solutions."

Dalton (5, p. 2) states that "several investigators have compared the various regression solutions and clarified the hypotheses tested by each. Yet, despite this clarification, no one has empirically compared the best known regression solutions to the more popular approximate ones." Marks (13, p. 351) says that the diversity of purpose in the various solutions "combined with the relative narrowness of the individual efforts, has resulted in a fragmented treatment of the problem of unbalanced data and in some cases confusion and controversy regarding methodology." Tsao states that "therefore, the need is very urgent for a systematic formulation of general methods of attacking the problems under such conditions" (22, p. 107). Overall and Spiegal (17, p. 316) say that " . . . theoretical statisticians provide few specific recommendations for handling of unequal cell frequencies . . ." Godbout (8, p. 5) says that special techniques have been derived to eliminate confounding as a result of unbalanced designs but that it is unclear which of these techniques should be used for a particular research question. Dalton (5, p. 2) has said that a computer simulation (Monte Carlo) study investigating the major techniques involved in handling disproportionate cell frequencies would be an important study.

Statement of the Problem

The problem of this study will be to determine the effect of varying degrees of disproportionality on four methods of handling disproportional cell frequencies in two-way analysis of variance.

Purposes of the Study

The purposes of this study will be (1) to determine if four methods of handling disproportionate cell frequencies in two-way analysis of variance differ in the results they produce, (2) to determine if the "approximate solutions" diverge from the "regression solutions", (3) to determine if the two "regression solutions" give different results, and (4) to determine if there is a point of disproportionality at which the four solutions begin to give spurious results.

Hypotheses

The following hypotheses have been formulated to carry out the purpose of this study.

1. Method 1 and Method 2 (two "regression solutions") will give diverging results as disproportionality increases.

2. The unweighted means analysis and the method of expected frequencies will give diverging results as disproportionality increases.

3. For moderate levels of disproportionality, Method 1 and the unweighted means analysis will give less spurious results than Method 2 and the method of expected frequencies.

4. For extreme levels of disproportionality, all four methods will yield results that tend to converge on each other.

5. For extreme levels of disproportionality, all four methods will give results that are spurious.

6. There will be a point of disproportionality at which one or more of the four methods will give spurious results.

Definition of Terms

"*a priori*" - Presupposed by experience.

Cell - All observations in a factorial design taken under one level of each independent variable of the design simultaneously.

Disproportionate Cells - Cell frequencies which are not proportionate with each other in a design.

Factorial Design - The simultaneous evaluation of two or more Factors (Independent Variables) in one experiment.

Fixed Model - A factorial design in which all treatment levels about which inferences are to be drawn are included in the design.

Method 1 - A multiple linear regression technique used to perform analysis of variance. It involves an estimation of independent effects of each factor adjusted for all others included in the model.

Method 2 - A multiple linear regression technique used to perform analysis of variance. It involves an estimation of main effects disregarding interactions and then an estimation of interactions adjusted for main effects.

Monte Carlo Simulation - A procedure in which random samples are drawn from populations having specified parameters and then a specific statistic is computed.

Proportionate Cells - Cells of a factorial design in which the number of observations is in a constant ratio with other cells in that design.

Subclass Number - The number of observations in a cell of a factorial design.

Delimitations

This study will be limited to experimental conditions simulated with the following conditions.

1. Factorial designs other than two-way are not being considered.

2. Only fixed models are being considered.

3. Selected methods of handling disproportionality are being considered.

CHAPTER BIBLIOGRAPHY

1. Anderson, R. L. and T. A. Bancroft, Statistical Theory in Research, New York, McGraw-Hill Book Company, 1952.

2. Bessent, Authella, "A Monte Carlo Study of the Analysis of Variance by Unweighted Means," unpublished Dissertation, The University of Texas at Austin, Austin, Texas, August, 1974.

3. Cochran, William G. and Gertrude M. Cox, Experimental Designs, New York, John Wiley and Sons, Inc., 1950.

4. Cohen, Jacob, "Multiple Regression as a General Data-Analytic System," Psychological Bulletin, LXX, no. 6 (1968), 426-443.

5. Dalton, Starrette, "Regression Approaches and Approximate Solutions to Analysis of Variance with Disproportionality Varied," Paper presented at the Annual Meeting of the American Educational Research Association (60th, San Francisco, California, April 19-23, 1976).

6. Dixon, Wilfrid J. and Frank J. Massey, Jr., Introduction to Statistical Analysis, New York, McGraw-Hill Book Company, Inc., 1951.

7. Falzer, Paul, "Representative Design and the General Linear Model," Speech Monographs, XLI (June, 1974), 127-138.

8. Godbout, Robert C., "Analysis of Unbalanced Data," Paper presented at the Annual Meeting of the American Educational Research Association (Washington, D.C., March 30 - April 3, 1975).

9. Huck, Schuyler W. and Benjamin H. Layne, "Checking for Proportional n's in Factorial Anova's," Educational and Psychological Measurement, XXXIV (Summer, 1969), 281-287.

10. Johnson, Palmer O. and Robert W. B. Jackson, Modern Statistical Methods: Descriptive and Inductive, Chicago, Rand McNally and Company, 1959.

11. Kendall, Maurice G., The Advanced Theory of Statistics, London, Charles Griffin and Company Limited, 1948.

12. Keppel, Geoffrey, Design and Analysis: A Researcher's Handbook, Englewood Cliffs, New Jersey, 1973.

13. Marks, Edmond, "Analysis of Variance Techniques for Unbalanced Data," Review of Educational Research, XLIV, no. 3 (Summer, 1974), 351-364.

14. Mood, Alexander McFarlane, Introduction to the Theory of Statistics, New York, McGraw-Hill Book Company, 1950.

15. Ostle, Bernard, Statistics in Research, Ames, Iowa, The Iowa State College Press, 1954.

16. Overall, John E. and C. James Klett, Applied Multivariate Analysis, New York, McGraw-Hill Book Company, 1972.

17. Overall, John E. and Douglas K. Spiegal, "Concerning Least Squares Analyses of Experimental Data," Psychological Bulletin, LXXII (1969), 311-322.

18. Proger, Barton B., "Unequal Cell Frequencies in Analysis of Variance: A Review and Extension of Methodology for Multiple Missing Observations," Paper published at Montgomery County Schools, King of Prussia, Pennsylvania, Reported in ERIC, 1972.

19. Roscoe, John T., Fundamental Research Statistics for Behavioral Sciences, New York, Holt, Rinehart and Winston, Inc., 1975.

20. Scheffe', Henry, The Analysis of Variance, New York, John Wiley & Sons, Inc., 1959.

21. Snedocer, George W., Statistical Methods, Ames, Iowa, The Iowa State College Press, 1946.

22. Tsao, Fei, "General Solution of the Analysis of Variance and Covariance in the Case of Unequal or Disproportionate Numbers of Observations in the Subclasses," Psychometrika, XI, no. 2 (June, 1946), 107-128.

23. Tsao, Fei, "Tests of Statistical Hypotheses in the Case of Unequal or Disproportionate Numbers of Observations in the Subclasses," Psychometrika, VII, no. 3 (September, 1942), 195-212.

24. Wert, James E., Charles O. Neidt, and J. Stanley Almann, Statistical Methods in Educational and Psychological Research, New York, Appleton-Century-Crofts, Inc., 1954.

CHAPTER II

SURVEY OF RELATED RESEARCH

Analysis of variance is a greatly used tool in educational research. One of the underlying assumptions in the traditional solutions of factorial analysis of variance designs is that there are equal or proportionate cell frequencies. However, in education as in many other fields of study, disproportionate cell frequecies occur quite often. Several methods of handling this situation have been developed.

Williams (27, p. 67) says that there are at least eight different solutions to the problem of disproportionality. He includes two data forcing methods (forced to proportionality): the method of discarding data and the method of estimating missing data. Three approximate methods are considered: the method of unweighted means, the method of expected cell frequencies, and the method of weighted means. Williams says that "the approximate methods were conceived as computational compromises for the method of fitting constants, a full regression solution" (27, p. 67). Overall and Spiegal (19) have defined three regression solutions for analyzing disproportionate data: Method 1, Method 2, and Method 3.

When the original work was done on the disproportionate cell frequency situation, a full regression solution was regarded as computationally too laborious to be of practical use for the research worker. Full regression solutions are now more viable options with the advent of the computer. Thus, the researcher now has many ways to handle disproportionality. The problem is in determining which if any of these solutions is more appropriate.

Data Forcing Methods

Williams (27, p. 68) argues that the method of discarding data is wasteful. Wert, Neidt, and Almann (26, p. 212) say that the procedure of discarding data causes the investigator to lose information. This may be serious, and it is unnecessary. Dalton (7, p. 10) found that data elimination was a poor alternative to other methods due to the strong tendency to yield Type II errors regardless of the presence or absence of an interaction. The method of discarding data is probably not a viable approach to the problem of disproportionality.

The other data forcing method is the method of estimating missing data. Williams (27, p. 69) has said:

> This method might be seen as more appropriate to the hand calculation era; if there are many missing subjects, it would seem that this method would become prohibitive, particularly in view of the relative ease of other solutions by use of the computer. It should also be noted that this approach will yield treatment effects that are slightly inflated. One additional concern is psychological - it seems somewhat unnerving to artifically create data for a statistical analysis.

Cochran and Cox (5, p. 74) state that the method of estimating missing data causes treatment Sum of Squares to be slightly larger than the correct treatment Sum of Squares for an F test for treatments. Godbout (11, p. 26) says that both the methods of artificially balancing a design by discarding data or estimating missing data involves approximate solutions which do not yield exact tests of the hypotheses of interest. He says that neither of these techniques are very satisfying. Thus, it seems that given the data handling methods available today, the data forcing methods should be considered to be among the weaker approaches to the problem of disproportionality.

Approximate Methods

Among the three approximate solutions being considered is the method of unweighted means. "The unweighted means analysis uses cell means to estimate main effects and interaction, and adjusts the error term by a factor which reflects the unequal cell sizes (7, p. 4). Williams (27, p. 67) says that it may be the most widely used technique for handling disproportionate cell frequencies. Anderson and Bancroft (1, p. 279) relate that the method of unweighted means has a minimum of computation and furnishes a short-cut procedure of testing for the existence of interactions. Steinhorst and Miller (23, p. 805) state that there are several assumptions for the unweighted means solution:

(a) no cell is empty, (b) it is for preliminary analysis only, (c) the cell frequencies do not vary greatly from equality, (d) primary interest is whether interaction is or is not present, (e) one wishes to test main effects when interaction is negligible, and (f) exact solutions are prohibitive or not available, and the study or experiment does not warrant an exact solution. They state (23, p. 802), furthermore, that the unweighted means analysis is approximate and that the statistics derived from it are only approximately distributed as F. Myers (17) warns that the experimenter should question the applicability of the unweighted means solution if the n's are very disparate. Both Dayton (8) and Winer (29) indicate that the unweighted means analysis is applicable only if the experimental design called for equal n and is subject loss was essentially random. Glass and Stanley (10, p. 440) write, "the unweighted means analysis is probably the simplest and one of the most justifiable techniques for analyzing disproportional designs." Johnson and Jackson (13, p. 241) state that "of all the possible approximate solutions, the method of unweighted squares of means is the simplest computationally and is to be preferred . . ."

The method of weighted means involves a more complicated algorithm than the unweighted means method. According to Williams (27, p. 72), "this method can be seen as one of the more complex approximate solutions, but that can be accomplished with the aid of a hand calculator." It gives an exact solution with regard to the interaction effect. Steinhorst and Miller (23, p. 806) relate that there are two important considerations: (1) the weighted means method is not applicable beyond the two factor situation and (2) as long as no empty cells appear, the method of unweighted means is more generally usable and offers an analysis similar to what the experimenter is familiar with in the equal or proportional frequency case. Tsao (24, p. 108) says that Yates presented this method assuming that interactions exist. Tsao goes on to say that the method is rather tedious. Dalton (7, p. 5) says that this method is of limited utility. It is seldom recommended when there are two or more missing scores per cell. Keppel (14, p. 356) takes the position that only rarely will one want to consider the weighted means analysis appropriate. He says that it may produce marked distortions and that these distortions do not occur with the unweighted means analysis. Steinhorst and Miller (23, p. 802) claim that the weighted means analysis yields tests for main effects which are not the usual F statistic and which have different power functions.

Another one of the approximate methods is the method of expected frequencies. This method involves multiplication of cell sums by the expected cell frequency to obtain a sum for each cell. Sums obtained in this manner are used in estimating main effects and interactions (7, p. 4). Myers

(17, p. 116) says that the method of expected frequencies is appropriate when proportionality can be assumed and when departure from proportionality is not too great. The method has been used largely when cell frequencies would naturally be disproportionate.

Regression Solutions

Among the "regression solutions" are Method 1, Method 2, and Method 3. Overall and Klett (18, p. 449) call Method 1 the "complete linear-model analysis." It involves an estimation of independent effects of each factor adjusted for all others included in the model. They call Method 2 the "experimental-design analysis." It involves an estimation of main effects <u>disregarding</u> interactions and then an estimation of interactions adjusted for main effects. Method 3 is called the "step-down analysis." It involves an initial ordering of the effects and then estimating each effect adjusting for those preceding it in the ordering and ignoring those following it. Overall and Klett (18, p. 449) state that "quite different results derive from the three methods in applications involving disproportionate cell frequencies." Keren, Gideon, and Lewis (15, p. 817) state that

> Since the use of unequal n's alters variability by itself, it turned out that three different least squares solutions that were presented by Overall and Spiegel yielded different results, although they were identical for the case of equal cell frequencies.

The structural model for Method 1 in a two-way analysis of variance is: $X_{ijm} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + e_{ijm}$ where μ is the grand mean, α_i is the treatment effect for level i of the first factor, β_j is the treatment effect for level j of the second factor, $\alpha\beta_{ij}$ is the interaction term for cell ij, and e_{ijm} is the error for individual m in cell ij. Marks (16, p. 358) elaborates on Method 1 by saying:

> One approach, which has been described and labelled Method 1 by Overall and Spiegel (1969) and is exemplified by the General Linear Hypothesis Program in Dixon (1971), is to compare reductions in sums of squares due to fitting different parameters of the complete model. For example, in a two-way design with interactions, the sum of squares for the α-factor is given as the difference between the sum of squares due to all the parameters except the α_i's, i.e., $SS(\alpha) = SS(\mu, \alpha, \beta, \gamma) - SS(\mu, \beta, \gamma)$.

Carlson and Timm (4, p. 563) believe that Method 1 is the best extension of traditional analysis of variance because the same parameters are estimated and the same hypotheses are tested in the orthogonal and the nonorthogonal cases. Overall and Woodward (21, p. 31) say "from the point of view of interpretation, it was emphasized that this strategy (Method 1) results in estimation of the same effects and tests of the same hypotheses that would be estimated and tested in an equal cell frequency design involving the same factors." Overall and Klett (18, p. 450) suggest that statistical literature says that Method 1 is consistent with the general linear-model analysis described in abstract terms by mathematical statisticians for the equal-cell frequency case. They go on to state, however, that they believe Method 1 is something different from the traditional analysis of variance in the disproportionate case.

The structural model for Method 2 in a two-way analysis of variance is: $X_{ijm} = \mu + \alpha_i + \beta_j + e_{ijm}$ where the terms are defined as with Method 1. The essential difference between the two methods is that Method 2 requires the assumption that no true interaction exists and consequently the interaction is not taken into consideration when estimating main effects (7, p. 3). Overall and Klett (18, p. 451) say that it is clear that Method 2 is the proper generalization of traditional experimental-design statistical texts, in which actual computational procedures are described for analyses of variances involving unequal cell frequencies, provide support for Method 2 as more like the traditional analyses of variance. Overall and Woodward (21, p. 22) suggest that "in the univariate case, Method 2 appears to be preferred by a number of statisticians for analysis of data from reasonably simple designs involving unequal and disproportionate cell frequencies." Overall and Spiegel (19) state that Method 2 seemed to be the most appropriate method for analysis of experimental data involving disproportionate cell frequencies. Later, Overall, Spiegel, and Cohen (20) reversed that stance in favor of Method 1.

The structural model for Method 3 is identical to the model for Method 1. However, Method 3 assumes a priori evidence to justify an ordered entry of vector sets representing α, β, and $\alpha\beta$ into the regression equation (7, p. 4). Method 3, sometimes referred to as the hierarchal model, does not test the same hypotheses as does analysis of variance. Williams and Linden (28, p. 11) state that:

> With this approach, a researcher is required to order the variables in relation to their research interest. For example, a researcher may be most interested in the A, or row effect, less interested in the B, or column effect, and may have little interest in the interaction effect. With this approach, each effect is adjusted only for those effects preceding it to the ordering. Thus, the A effect is found directly, the B effect is adjusted for the combined A and B effect.

The requirement of establishing a priori ordering of variables limits its usage to the researcher (7, p. 4). Below in Table I, Methods 1, 2, and 3 are compared in terms of the Sums of Squares (19, p. 316).

TABLE I

Method 1		
Source	SS	df
A	$SS_T\ [R^2(\alpha_i, \beta_j, \alpha\beta_{ij}) - R^2(\beta_j, \alpha\beta_{ij})]$	a-1
B	$SS_T\ [R^2(\alpha_i, \beta_j, \alpha\beta_{ij}) - R^2(\alpha_i, \alpha\beta_{ij})]$	b-1
AB	$SS_T\ [R^2(\alpha_i, \beta_j, \alpha\beta_{ij}) - R^2(\alpha_i, \beta_j)]$	(a-1) (b-1)
Error	$SS_T\ [1-R^2(\alpha_i, \beta_j, \alpha\beta_{ij})]$	N-ab
Total	SS_T	N-1

TABLE I--Continued

Method 2		
Source	SS	df
A	$SS_T\ [R^2(\alpha_i, \beta_j) - R^2(\beta_j)]$	a-1
B	$SS_T\ [R^2(\alpha_i, \beta_j) - R^2(\alpha_i)]$	b-1
AB	$SS_T\ [R^2(\alpha_i, \beta_j, \alpha\beta_{ij}) - R^2(\alpha_i, \beta_j)]$	(a-1) (b-1)
Error	$SS_T\ [1 - R^2(\alpha_i, \beta_j, \alpha\beta_{ij})]$	N-ab
Total	SS_T	N-1

Method 3		
Source	SS	df
A	$SS_T\ [R^2(\alpha_i)]$	a-1
B	$SS_T\ [R^2(\alpha_i, \beta_j) - R^2(\alpha_i)]$	b-1
AB	$SS_T\ [R^2(\alpha_i, \beta_j, \alpha\beta_{ij}) - R^2(\alpha_i, \beta_j)]$	(a-1) (b-1)
Error	$SS_T\ [1 - R^2(\alpha_i, \beta_j, \alpha\beta_{ij})]$	N-ab
Total	SS_T	N-1

Methods to be Used

In this study, four of the above eight methods of analyzing disproportionality in analysis of variance will be examined. These are (1) the unweighted means solution, (2) the method of expected frequencies, (3) Method 1, and (4) Method 2. The two data forcing techniques will not be included in this study because the literature has already shown them to be the poorest alternative solutions to the problem of disproportionality. The method of discarding data is wasteful and has a strong tendency to yield Type II errors. The method of estimating missing data becomes prohibitive if there are many missing observations; and it yields slightly inflated treatment effects.

The method of weighted means will not be included because it is seldom recommended when there are two or more missing scores per cell. It is not applicable beyond the two factor situation; and when there are no empty cells, the method of unweighted means is more generally usable. Method 3 of the "regression solutions" will not be included in the study because it does not test the same hypotheses as does analysis of variance. Its usefulness is extremely limited.

Advantages of "Least-Squares" Techniques

There are several advantages reported in the literature of using "least-squares" techniques (Method 1, Method 2, and Method 3) over other techniques. One of these advantages is when disproportionality is present. Roscoe says (22, p. 348) that he is partial to the use of multiple regression when disproportionality is present. Appelbaum and Cramer (2, p. 335) state that

> The easy access to computer programs that perform the analysis of variance by a general linear model approach makes possible the computations for this method of dealing with nonorthogonal multifactor designs and eliminates the need for approximate solutions.

Overall and Spiegel relate that "using least-squares regression methods, analyses of variance can be accomplished on data from arbitrary experimental designs in which no attempt is made to control cell frequencies" (19, p. 311). Cochran and Cox (5, p. 73) recommend a least-squares solution as the procedure to use when missing observations exist. Steinhorst and Miller (23, p. 804) state that

> In response to the immediately preceding question, the authors would suggest that with the linear model theory which has been developed to date one can readily analyze disproportionate data with the same theory as one would treat proportionate or equal frequency data.

Cohen (6, p. 438) says that an important aspect of using multiple regression in computing analysis of variance problems is that with multiple regression the researcher has the option of not analyzing all possible aspects of variables. He is particularly referring to not using joint aspects of variables (interaction) if for no other reason than the rapid loss of degrees of freedom for estimating error. Cohen (6, p. 438) goes on to say:

> This goes hand in hand with the flexibility of the MR system, which makes readily possible the representation of the research issues posed by the investigator (i.e., multiple regression in the service of the ego!), rather than the canned issues mandated by AV computational routines.

Anderson and Bancroft (1, p. 279) say that the "method of least squares furnishes an exact test for interactions . . ." and that (1, p. 284) "the exact method (least squares analysis) is somewhat more powerful than the method of unweighted means." Jennings (12, p. 95) states:

> A second purpose is to argue that a regression approach to analysis of variance is a "good" technique in that it offers a major pedagogical advantage and in some cases computational superiority over alternative procedures when computers are available.

Waldberg (25, p. 76) stated:

> The generalized RA model in practice provides comprehensive and useful estimates of magnitudes of effects and their significance. The most obvious instance is the multiple regression coefficient: when squared (R^2) it reveals directly how much variance in the dependent variable is associated with or accounted for by the independent variables; when tested for significance, it reveals the chance probability of overall association between all the independent variables and the dependent variable.

Falzer (9, p. 130) says that "a reliance on both R and F statistics, then, facilitates representative validity and eases data interpretation." However, Marks (16, p. 363) cautions that "although least squares provides a relatively easy and direct method of obtaining a solution and constructing estimable functions for disproportionate (including missing cells) data, the framing, testing, and interpretation of hypotheses are not so simple."

Dalton (7, p. 13) reported that

> A slight divergence of results was found when a moderate degree of nonorthogonality was present, but not along the dimension of regression solutions versus nonregression solutions. Rather Method 1 and the unweighted means analysis appear to be best when results differ.

Method 1 and Method 2 might be expected to give diverging results as disproportionality increases as would the unweighted means analysis and the method of expected frequencies.

Furthermore, for moderate levels of disproportionality, Method 1 and the unweighted means analysis might give less spurious results than Method 2 and the method of expected frequencies.

Dalton also stated that when nonorthogonality was extreme all four solutions led to basically the same results (7, p. 11). Errors were found with all four methods when nonorthogonality was extreme. By utilizing Monte Carlo simulation techniques, an attempt will be made to empirically determine if one of the four methods is superior to the others for a given design as disproportionality is increased.

CHAPTER BIBLIOGRAPHY

1. Anderson, R. L. and T. A. Bancroft, Statistical Theory in Research, New York, McGraw-Hill Book Company, 1952.

2. Appelbaum, Mark I. and Elliot M. Cramer, "Some Problems in the Nonorthogonal Analysis of Variance," Psychological Bulletin, LXXXI (1974), 335-343.

3. Bessent, Authella Ann McGuirre, "A Monte Carlo Study of the Analysis of Variance by Unweighted Means," Doctoral dissertation, The University of Texas at Austin, Austin, Texas, August, 1974.

4. Carlson, James E. and Neil H. Timm, "Analysis of Nonorthogonal Fixed-Effects Designs," Psychological Bulletin, LXXXI (September, 1974), 563-570.

5. Cochran, William G. and Gertrude M. Cox, Experimental Designs, New York, John Wiley and Sons, Inc., 1950.

6. Cohen, Jacob, "Multiple Regression as a General Data-Analytic System," Psychological Bulletin, LXX (1968), 426-443.

7. Dalton, Starrette, "Regression Approaches and Approximate Solutions to Analysis of Variance with Disproportionality Varied," Paper presented at the Annual Meeting of the American Educational Research Association (60th, San Francisco, California, April 19-23, 1976).

8. Dayton, C. M., Design of Educational Experiments, New York, McGraw-Hill Book Company, 1970.

9. Falzer, Paul Robert, "Representative Design and the General Linear Model," Speech Monographs, XLI (June, 1974), 127-138.

10. Glass, Gene V. and Julian C. Stanley, Statistical Methods in Education and Psychology, Englewood Cliffs, New Jersey, Prentice-Hall, 1970.

11. Godbout, Robert C., "Analysis of Unbalanced Data," Paper presented at the Annual Meeting of the American Educational Research Association (Washington, D.C., March 30 - April 3, 1975).

12. Jennings, Earl, "Fixed Effects Analysis of Variance by Regression Analysis," Multivariate Behavioral Research, (January, 1967), 95-107.

13. Johnson, Palmer O. and Robert W. B. Jackson, Modern Statistical Methods: Descriptive and Inductive, Chicago, Rand McNally and Comapny, 1959.

14. Keppel, Geoffrey, Design and Analysis: A Researcher's Handbook, Englewood Cliffs, New Jersey, Prentice Hall, 1973.

15. Keren, Gideon and Charles Lewis, "Nonorthogonal Designs: Sample Versus Population," Psychological Bulletin, LXXXIII (1976), 817-826.

16. Marks, Edmond, "Analysis of Variance Techniques for Unbalanced Data," Review of Educational Research, XLIV, no. 3 (Summer, 1974), 351-364.

17. Myers, Jerome L., Fundamentals of Experimental Design, Boston, Allyn and Bacon, Inc., 1972.

18. Overall, John E. and C. James Klett, Applied Multivariate Analysis, New York, McGraw-Hill Book Company, 1972.

19. Overall, John E. and Douglas K. Spiegal, "Concerning Least Squares Analysis of Experimental Data," Psychological Bulletin, LXXII (1969), 311-322.

20. Overall, John E., Douglas K. Spiegal, and Jacob Cohen, "Equivalence of Orthogonal and Nonorthogonal Analysis of Variance," Psychological Bulletin, LXXXII (March, 1975), 182-186.

21. Overall, John E. and J. Arthur Woodward, "Multivariate Analysis of Variance by Multiple Regression Methods," Psychological Bulletin, LXXXII, no. 1 (1975), 21-32.

22. Roscoe, John T., Fundamental Research Statistics for Behavioral Sciences, New York, Holt, Rinehart and Winston, Inc., 1975.

23. Steinhorst, R. Kirk and C. Dean Miller, "Disproportionality of Cell Frequencies in Psychological and Educational Experiments Involving Multiple Classifications," Educational and Psychological Measurement, XXIX (Winter, 1969), 799-811.

24. Tsao, Fei, "General Solution of the Analysis of Variance and Covariance in the Case of Unequal or Disproportionate Numbers of Observations in the Subclasses," Psychometrika, XI, no. 2 (June, 1946), 107-128.

25. Walberg, Herbert J., "Generalized Regression Models in Educational Research," American Educational Research Journal, VIII (January, 1971), 71-91.

26. Wert, James E., Charles O. Neidt, and J. Stanley Almann, Statistical Methods in Educational and Psychological Research, New York, Appleton-Century-Crofts, Inc., 1954.

27. Williams, John D., "Two Way Fixed Effects Analysis of Variance With Disproportionate Cell Frequencies," Multivariate Behavioral Research, VII (January, 1972), 67-82.

28. Williams, John D. and Alfred C. Lindem, "A Computer Program for Two-Way Fixed Effects Analysis of Variance with Disproportionate Cell Frequencies," Educational and Psychological Measurement, XXXII (1972), 169-172.

29. Winer, B. J., Statistical Principals in Experimental Design, New York, McGraw-Hill Book Company, 1971.

CHAPTER III

PROCEDURE

In order to conduct an investigation of this problem, a Monte Carlo procedure was used. Clark (2, p. 605) says that "typically a Monte Carlo analysis is used only when an analytic solution is not obtainable." Furthermore, he states that "Monte Carlo analysis, as so defined, is almost a general, effective procedure that enables one to solve many problems too complex for mathematical analysis" (2). It is estimated that in this study, 1.5 million random numbers were used for the design examined. Approximately 500,000 F values were calculated. Indeed, without using Monte Carlo techniques, this study would be prohibitive.

According to Clark (2), the term Monte Carlo indicates that one knows explicitly the distribtuions of all the random elements in the problem. In this study, random numbers were generated into a normal distribution thus meeting the above criterion. "In this sense the term Monte Carlo signifies that one could simulate the random process by a desk calculation that used tables of random numbers or by a computer program that generates random numbers" (2). That was true in this study.

Procedures for Collecting Data

Pseudo-random numbers were generated using a Monte Carlo simulation procedure which utilizes a pseudo-random number generator. Data generation was performed on an IBM 360, model 50 computer system at the North Texas State University Computing Center.

For each simulation model, a procedure which employs sub-routines Randu and Gauss were utilized to produce observations for the given condition. Randu computes uniformly distributed random real numbers between zero and 2^{31}. Gauss computes a normally distributed random number with a given mean and standard deviation. In order to produce one normal random observation, Gauss utilizes Randu to generate twelve uniform random numbers.

Two tests were used on selected groups of numbers to test for randomness and normality. Randomness was tested by utilizing the One-Sample Runs Test (11, p. 52). Runs were selected according to whether or not numbers were above or below the mean that was selected for the random number generator. None of the One-Sample Runs Tests conducted proved to be significant at the .05 level. Thus, numbers were generated above and below the mean in a random manner.

Normality was tested by comparing selected groups of the generated numbers with an expected normal distribution of numbers based on the mean and standard deviation used in Gauss. A Chi-square Goodness of Fit test was used to determine if there was a significant difference between the generated numbers and the expected distribution of numbers (3, p. 177). None of the Chi-square Goodness of Fit tests was significant at the .05 level. Thus, the generated numbers were not significantly different from the normal distribution.

In this study, four methods of handling disproportionate data were examined and compared to equal cell two-way analysis of variance. The four methods were: Method 1, Method 2, method of unweighted means, and method of expected frequencies. Computer programs were written by the author to calculate F's based on the method of unweighted means, the method of expected frequencies, and Analysis of Variance - a traditional approach. The computer programs for the methods of unweighted means and expected frequencies were based on those algorithms presented by Williams (13, pp. 69-72). The program computing the traditional Analysis of Variance was based on formulas presented by Ferguson (3, p. 227). The computer program for Method 1 and Method 2 was initially a multiple regression program called REGN (1) that is a part of the North Texas State University Computer library. This program was modified by the author to meet the needs of this study.

Each of these four computer programs was tested to assure that calculations were correct. The method of unweighted means and the method of expected frequencies program results were compared to hand calculations of the same methods. The results were identical. The method of unweighted means program results were compared to the results from a program named ST013 which is a Two-Way Analysis of Variance program in the North Texas State University Computing Center library that utilizes the method of unweighted means to handle disproportionate cell frequencies. The two programs produced F ratios identical to thousandths place.

In order to utilize REGN to compute by Method 1 and Method 2, generating statements were included to produce row, column, and interaction vectors by effect coding. A test run was conducted using data from Overall and Klett (10, p. 445). Results were identical to those calculated by Overall and Klett (10, p. 449). Further testing was done on the Method 1 and Method 2 program by comparing the results to results produced by MULTIVARIANCE (4), a computer program also available from the North Texas State Computer library, which can calculate both by Method 1 and Method 2. Results of the programs were identical.

The computer program written by the author to compute traditional Analysis of Variance for equal cell frequencies was tested by comparing results to those given by ST013 for equal cell sizes. Results were identical. Hand calculations also produced the same results.

All programs were written, modified, and tested separately. Afterwards, the programs were combined and run as one. Thus, the traditional Analysis of Variance for equal cell frequencies along with the four methods of handling disproportionality were one computer program. This program was also tested and checked to assure that it was still giving the same results that the original programs produced.

In this study, five cases were examined: (1) the case of no significant differences in means, (2) the case of significant differences in the rows only, (3) the case of significant differences in the columns only, (4) the case of significant interaction, and (5) the case of significant differences in the rows and columns. The computer program used in all cases was the same except that the random number generator utilized different means for given cells to fit each case.

In all cases, forty numbers were generated and divided up into four cells with ten in each cell. This produced the data for a 2x2 design. F ratios were calculated by the traditional Analysis of Variance for equal cell frequencies. This process of generating numbers and calculating F ratios was repeated one-thousand times. The probability of each F ratio occurring was calculated by using the following formula (12):

For v_1 odd and v_2 even (where v = degrees of freedom):

$$Q(x) = 1 - (1-k)^{\frac{v_1}{2}} \left[1 + \frac{v_1 k}{2} + \frac{v_1(v_1+2)k^2}{2(4)} + \cdots + \frac{v_1(v_1+2)\ldots(v_2+v_1-4)k^{\frac{v_2-2}{2}}}{2(4)\ldots(v_2-2)} \right]$$

where: $k = \dfrac{1}{1+\dfrac{v_1 x}{v_2}}$

A frequency distribution of these probabilities was calculated for row, column, and interaction F's.

Utilizing the same initial seed number, random numbers were then generated again in groups of forty. This time the cell sizes varied depending upon what disproportionality was being examined at the time. For each group of forty numbers, F ratios were produced using the method of unweighted means, the method of expected frequencies, Method 1, and Method 2. The process was repeated one-thousand times. The probability of each F ratio occurring was calculated by the previously mentioned formula (12). Frequency distributions for these probabilities were obtained for row, column, and interaction F's under each of the four methods of handling disproportionality.

For the case of no significant differences, a mean of ten and a standard deviation of two were used to generate the random numbers. In order to determine what means to use in the other four cases, the Non-Central F distribution was used. In this manner, Type II errors could be examined.

Row, column, and interaction effects were calculated by using the following formulas (8, p. 179):

$$\phi = \frac{\sqrt{\sum_{i=1}^{p} \alpha_i^2 / p}}{\sigma_E / \sqrt{nq}}$$

$$\phi = \frac{\sqrt{\sum_{j=1}^{q} \beta_j^2 / q}}{\sigma_E / \sqrt{np}}$$

$$\phi = \frac{\sqrt{\sum_{i=1}^{p} \sum_{j=1}^{q} (\alpha\beta_{ij})^2 / [(p-1)(q-1)+1]}}{\sigma_E / \sqrt{n}}$$

A power of .60 and level of significance of .05 were used. These effects determined the size of the mean. The standard deviation used was two. Other than the differences in means used, the procedures for the five cases were identical.

As was mentioned previously, different cell sizes were examined in the study in order to determine what effect disproportionality had on the four methods. When cell sizes are unequal, there is potential for disproportionality. The following equation is presented by Godbout (6, p. 16) as a test to determine if a design is proportional or disproportional:

$$n_{ij} = \frac{n_{i.}n_{.j}}{n_{..}} \quad \text{for all i and j}$$

If the above equation holds true for a design, then the design is not disproportional. Glass and Stanley (5, p. 434) and Huck and Layne (7, p. 282) also present the same test. In this study, disproportionate conditions were desired. Thus, only designs that failed the above test were examined with the exception of the equal cell designs.

Newman and Oravecz (9) utilized a Chi-square approach to determining how disproportional a design is. They recommend (9, p. 9) that a Chi-square value where $\alpha = .25$ be used as "mild" disproportionality and that a Chi-square value where $\alpha = .05$ be used as "severe" disproportionality. In this study, Chi-square α values were used as a guide to degree of disproportionality. The Chi-square approach used here was recommended by Ferguson (3, p. 238). It is a modified version of the traditional Chi-square test for independence. The Chi-square value is obtained by using the grand mean as the expected value in each cell. In every case in this study, an expected value of ten was used in a cell.

In this study, disproportionality was increased rapidly until spurious results from at least one of the four methods of analysis was found. Disproportionality was then decreased until no spurious results were found. By vacillating the level of disproportionality in this manner, an attempt was made to coverge on the point of disproportionality at which at least one of the four methods of analysis began to give spurious results.

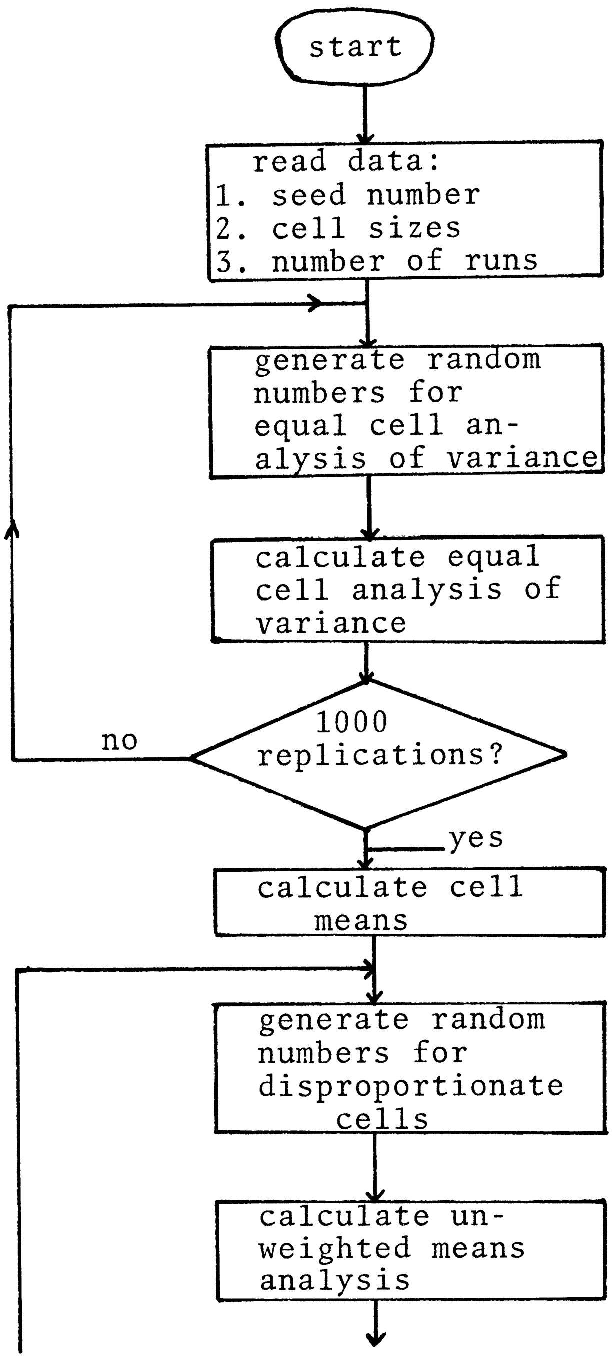
In an attempt to examine the impact of other values of power on the results of this study, power values of .80 and .95 were also used. Several seed numbers were used in the same situation to determine the effect of seed numbers on results.

Procedures for Analysis of Data

The frequency distributions for each of the four methods of handling disproportionality were compared to the equal cell analysis of variance frequency distribution to determine if the distributions of F's were significantly different. The Kolmogorov-Smirnov test was used to determine if significant differences existed between distributions. In the No Effects Case, the frequency distributions of each of the four methods of handling disproportionality were also compared to a theoretical uniform distribution in a similar manner.

Tables are presented displaying the results of the Kolmogorov-Smirnov tests for the row, column, and interaction F probability distributions for all four methods of computing Analysis of Variance in all five cases. The frequency distributions for the F probabilities are presented in table form. Frequencies at the .01, .05, and .10 levels of significance are compared to further aid in determining if Type I and Type II errors have occurred.

FLOWCHART FOR THE COMPUTER PROGRAM



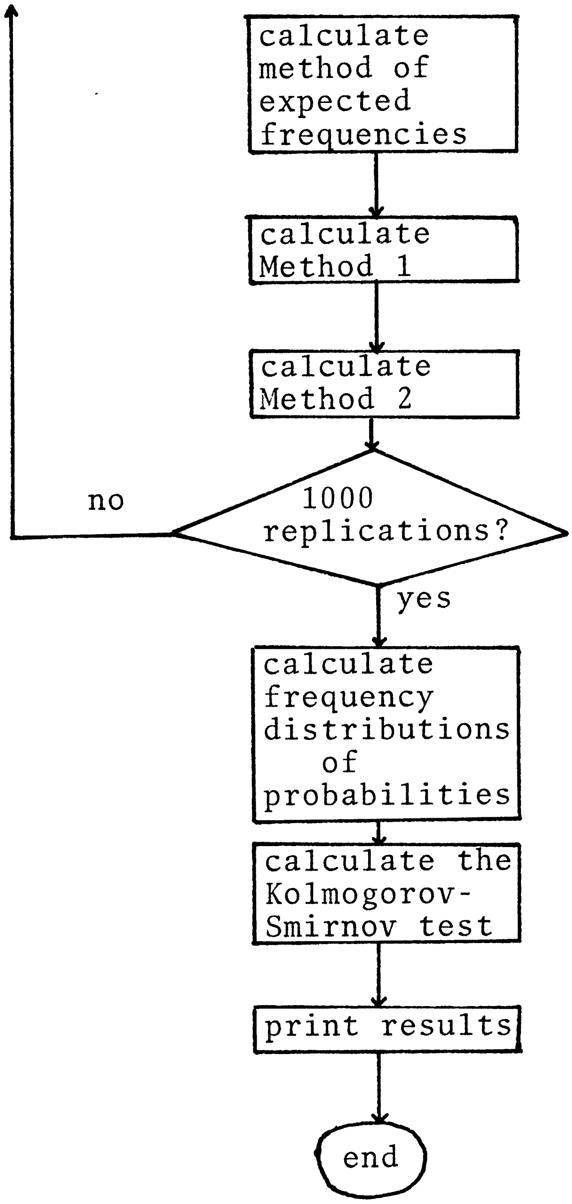


Fig. 1--Flowchart for the computer program

CHAPTER BIBLIOGRAPHY

1. Bolding, J. T., University of Arkansas.

2. Clark, Charles E., "Importance Sampling in Monte Carlo Analyses," Operations Research, IX, no. 5 (September-October, 1961), 603-620.

3. Ferguson, George A., Statistical Analysis in Psychology & Education, New York, McGraw-Hill Book Company, 1971.

4. Finn, Jeremy D., State University of New York.

5. Glass, Gene V. and Julian C. Stanley, Statistical Methods in Education and Psychology, Englewood Cliffs, New Jersey, Prentice-Hall, Inc., 1970.

6. Godbout, Robert C., "Analysis of Unbalanced Data," Paper presented at the Annual Meeting of the American Educational Research Association (Washington, D.C., March 30 - April 3, 1975).

7. Huck, Schuyler W. and Benjamin H. Layne, "Checking for Proportional n's in Factorial Anova's," Educational and Psychological Measurement, XXXIV, no. 2 (Summer, 1974), 281-287.

8. Kirk, Roger E., Experimental Design: Procedures for the Behavioral Sciences, Belmont, California, Brooks/Cole Publishing Company, Inc., 1968.

9. Newman, Isadore and Michael T. Oravecz, "Solutions to the Problem of Disproportionality: A Discussion of the Models," Multiple Linear Regression Viewpoints, VII, no. 3 (1977), 1-25.

10. Overall, John E. and C. James Klett, Applied Multivariate Analysis, New York, McGraw-Hill Book Company, 1972.

11. Siegel, Sidney, Nonparametric Statistics for the Behavioral Sciences, New York, McGraw-Hill Book Company, 1956.

12. Texas Instruments Programmable Calculator Program Manual ST1 Statistic Library 1975, Texas Instruments Incorporated.

13. Williams, John D., "Two Way Fixed Effects Analysis of Variance with Disproportionate Cell Frequencies," Multivariate Behavioral Research, VII (January, 1972), 67-82.

CHAPTER IV

ANALYSIS OF DATA AND FINDINGS

The results of this study are presented in five parts. Each of five cases dealing with row, column, or interaction effects are presented in each part. The first part is the no effects case. The second is the case of row effects only. The third part contains the case of column effects only. The fourth part deals with row and column effects but no interaction. The fifth part is the case of interaction effects only.

An examination of Kolmogorov-Smirnov D values calculated for F probability distributions for each χ^2 level of disproportionality is examined in each case. The analysis includes a presentation of and a discussion of the number of F values at the .10, .05, and .01 levels of α in each case. A discussion of the impact of changing power in the four cases with built-in effects is given. An examination of the effect of changing seed numbers on the results of the analysis is presented in each case.

No Effects Case

In simulating the no effects case, equal cell means were used. Initially, each cell contained ten numbers. Disproportionality was established by generating varying numbers of values in each cell. This disproportionality was measured by Chi-square values. Each level of disproportionality was run 1000 times.

An equal cell analysis of variance method was used to calculate F values before disproportionality was created. From these F values, an F probability distribution was obtained. After disproportionality was established, F values were calculated by the method of unweighted means, the method of expected frequencies, Method 1, and Method 2. F values for each of these four methods were used to calculate F probability distributions.

Kolmogorov-Smirnov D values were calculated between each of the distributions of the four methods of handling disproportionality and the distribution of equal cell analysis of variance under each level of disproportionality (shown by a χ^2 value). Table II contains these results for row, column, and interaction effects.

TABLE II

D VALUES FOR NO EFFECTS CASE COMPARING THE FOUR METHODS OF HANDLING DISPROPORTIONALITY TO THE EQUAL CELL ANOVA AS χ^2 INCREASES FOR ROWS, COLUMNS, AND INTERACTIONS

		D values			
χ^2 value		Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
0.0	row	.000	.000	.000	.000
	col.	.000	.000	.000	.000
	inter.	.000	.000	.000	.000
1.6	row	.030	.024	.020	.020
	col.	.016	.014	.016	.016
	inter.	.026	.021	.026	.026
2.6	row	.011	.026	.011	.013
	col.	.019	.031	.019	.021
	inter.	.026	.021	.026	.026
3.6	row	.021	.041	.021	.021
	col.	.034	.023	.034	.034
	inter.	.036	.017	.036	.036
6.4	row	.034	.057*	.034	.034
	col.	.015	.051*	.015	.015
	inter.	.030	.032	.030	.030
7.4	row	.027	.063*	.027	.025
	col.	.026	.031	.026	.025
	inter.	.021	.041	.021	.021

*Significant at the .05 level.

TABLE II--Continued

χ^2 value		D values: Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
	row	.023	.076*	.023	.028
8.6	col.	.025	.075*	.025	.027
	inter.	.031	.056*	.031	.031
	row	.026	.139*	.026	.035
19.4	col.	.014	.148*	.014	.019
	inter.	.038	.147*	.038	.038
	row	.627*	.409*	.689*	.742*
26.6	col.	.641*	.435*	.693*	.643*
	inter.	.636*	.508*	.684*	.684*
	row	.723*	.490*	.751*	.761*
40.6	col.	.702*	.517*	.741*	.693*
	inter.	.696*	.645*	.732*	.732*
	row	.746*	.489*	.787*	.689*
59.6	col.	.736*	.539*	.788*	.846*
	inter.	.739*	.714*	.789*	.789*

For the no effects case only, Kolmogorov-Smirnov D values were calculated between the distributions of each of the four methods of handling disproportionality and a uniform distribution of an equal number of F values at every .05 interval of proportionality as χ^2 increases. The results of this analysis are presented in Table III

TABLE III

D VALUES FOR NO EFFECTS CASE COMPARING THE FOUR METHODS OF HANDLING DISPROPORTIONALITY TO THE UNIFORM DISTRIBUTION AS χ^2 INCREASES FOR ROWS, COLUMNS, AND INTERACTIONS

		D values			
χ^2 value		Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
0.0	row	.039	.039	.039	.039
	col.	.012	.012	.012	.012
	inter.	.018	.018	.018	.018
1.6	row	.036	.026	.036	.036
	col.	.017	.014	.017	.017
	inter.	.030	.020	.030	.030
2.6	row	.040	.025	.040	.038
	col.	.021	.033	.021	.023
	inter.	.033	.019	.033	.033
3.6	row	.034	.012	.034	.034
	col.	.030	.020	.030	.030
	inter.	.047*	.026	.047*	.047*
6.4	row	.029	.030	.029	.029
	col.	.014	.053*	.014	.014
	inter.	.042	.021	.042	.042

*Significant at the .05 level.

TABLE III--Continued

χ^2 value		Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
		D values			
	row	.030	.030	.030	.027
7.4	col.	.023	.027	.023	.023
	inter.	.034	.034	.034	.034
	row	.031	.042	.031	.035
8.6	col.	.022	.074*	.022	.026
	inter.	.030	.040	.030	.030
	row	.033	.108*	.033	.031
19.4	col.	.019	.146*	.019	.020
	inter.	.025	.135*	.025	.025
	row	.649*	.444*	.700*	.741*
26.6	col.	.648*	.440*	.700*	.650*
	inter.	.649*	.522*	.700*	.700*
	row	.722*	.528*	.750*	.774*
40.6	col.	.711*	.511*	.750*	.700*
	inter.	.714*	.661*	.750*	.750*
	row	.745*	.527*	.800*	.700*
59.6	col.	.745*	.533*	.800*	.850*
	inter.	.750*	.732*	.800*	.800*

In an effort to analyze and interpret these D values in light of Type I errors, several tables are presented containing the number of F values generated by each method under each level of disproportionality at the .10, .05, and .01 level of significance. For each level of disproportionality, an n of 1000 was used. Thus, an expected value for the no effects case at the .10 level of significance is 100. For the .05 level, the expected value is 50; and for the .01 level, it is 10. Table IV contains row, column, and interaction F frequencies for the Row Effects Case at the .10, .05, and .01 levels of significance.

TABLE IV

FREQUENCIES OF ROW, COLUMN, AND INTERACTION F VALUES FOR THE NO EFFECTS CASE AT THE .10, .05, AND .01 LEVELS OF SIGNIFICANCE AS DISPROPORTIONALITY INCREASES

χ^2 value			Equal Cell Method	Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
	row	.10	83	83	83	83	83
		.05	38	38	38	38	38
		.01	6	6	6	6	6
		.10	109	109	109	109	109
0.0	col.	.05	49	49	49	49	49
		.01	7	7	7	7	7
		.10	89	89	89	89	89
	inter.	.05	43	43	43	43	43
		.01	6	6	6	6	6
		.10	83	72	78	72	72
	row	.05	38	36	37	36	36
		.01	6	6	7	6	6
		.10	109	102	108	102	102
1.6	col.	.05	49	51	58	51	51
		.01	7	14	15	14	14

TABLE IV--Continued

χ^2 value			Equal Cell Method	Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
	inter.	.10	89	93	93	93	93
		.05	43	43	45	43	43
		.01	6	13	19	13	13
2.6	row	.10	83	81	82	81	80
		.05	38	39	45	39	40
		.01	6	6	6	6	6
	col.	.10	109	109	122	109	107
		.05	49	60	63	60	58
		.01	7	15	16	15	15
	inter.	.10	89	90	103	90	90
		.05	43	46	52	46	46
		.01	6	10	11	10	10
3.6	row	.10	83	75	93	75	75
		.05	38	37	44	37	37
		.01	6	7	7	7	7
	col.	.10	109	104	111	104	104
		.05	49	55	66	55	55
		.01	7	14	19	14	14
	inter.	.10	89	89	105	89	89
		.05	43	45	51	45	45
		.01	6	9	13	9	9
6.4	row	.10	83	80	102	80	80
		.05	38	40	61	40	40
		.01	6	6	12	6	6
	col.	.10	109	96	134	96	97
		.05	49	47	64	47	47
		.01	7	8	17	8	8
	inter.	.10	89	88	107	98	98
		.05	43	49	65	49	49
		.01	6	8	13	8	8

TABLE IV--Continued

χ^2 value			Equal Cell Method	Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
		.10	83	85	115	85	80
	row	.05	38	46	66	46	49
		.01	6	7	17	7	8
		.10	109	94	113	94	96
7.4	col.	.05	49	55	63	55	46
		.01	7	16	22	16	12
		.10	89	84	120	84	84
	inter.	.05	43	45	63	45	45
		.01	6	8	14	8	8
		.10	83	77	111	77	77
	row	.05	38	39	61	39	41
		.01	6	8	15	8	8
		.10	109	113	158	113	107
8.6	col.	.05	49	48	83	48	51
		.01	7	14	19	14	12
		.10	89	86	125	86	86
	inter.	.05	43	42	69	42	42
		.01	6	8	16	8	8
		.10	83	99	183	99	96
	row	.05	38	41	124	41	42
		.01	6	7	38	7	4
		.10	109	110	221	110	120
19.4	col.	.05	49	63	147	63	64
		.01	7	9	57	9	10
		.10	89	84	179	84	84
	inter.	.05	43	49	115	49	49
		.01	6	8	45	8	8
		.10	83	0	0	0	0
	row	.05	38	0	0	0	0
		.01	6	0	0	0	0
		.10	109	0	0	0	0
26.6	col.	.05	49	0	0	0	0
		.01	7	0	0	0	0

TABLE IV--Continued

χ^2 value			Equal Cell Method	Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
		.10	89	0	0	0	0
	inter.	.05	43	0	0	0	0
		.01	6	0	0	0	0
		.10	83	0	0	0	0
	row	.05	38	0	0	0	0
		.01	6	0	0	0	0
		.10	109	0	0	0	0
40.6	col.	.05	49	0	0	0	0
		.01	7	0	0	0	0
		.10	89	0	0	0	0
	inter.	.05	43	0	0	0	0
		.01	6	0	0	0	0
		.10	83	0	0	0	0
	row	.05	38	0	0	0	0
		.01	6	0	0	0	0
		.10	109	0	0	0	0
59.6	col.	.05	49	0	0	0	0
		.01	7	0	0	0	0
		.10	89	0	0	0	0
	inter.	.05	43	0	0	0	0
		.01	6	0	0	0	0

The results of Table II show that only the method of expected frequencies had significant D values for $\chi^2 \leq 19.4$. For n = 1000, the critical D value for $\alpha = .05$ is .043. This critical D value was first exceeded by the method of expected frequency results at $\chi^2 = 6.4$ for rows, $\chi^2 = 6.4$ for columns, and $\chi^2 = 8.6$ for interaction. The other three methods produced very similar results to each other and all four methods produced extremely significant results at $\chi^2 \geq 26.6$.

Table IV depicts the method of expected frequencies as producing a greater number of large F values than the equal cell method when $3.6 \leq \chi^2 \leq 19.4$. Thus, combined with information from Table II, this is an indication that the method of expected frequencies committed Type I errors for $6.4 \leq \chi^2 \leq 19.4$ for rows; $\chi^2 = 6.4$, $\chi^2 = 8.6$, and $\chi^2 = 19.4$ for columns; and $\chi^2 = 8.6$ and 19.4 for interaction. However, for $\chi^2 \geq 26.6$ even though all four methods had significant D values, Table IV shows that the difference was in the "safe" direction, and Type I errors were not committed at these levels of disproportionality. An examination of frequency distributions (not included here) showed that the F distributions were skewed towards the low levels of probability.

Table III presented the four methods against a uniform distribution. For $\chi^2 \geq 26.6$ significant D values were found for all four methods in the row, column, and interaction effects. An examination of Table IV again shows no significant F values were derived for $\chi^2 \geq 26.6$ for any of the methods. Thus, the significant D values were not a result of Type I errors.

Table III does show that the method of expected frequencies produced significant D values at $\chi^2 > 8.6$ for rows, at $\chi^2 = 6.4$ and $\chi^2 \geq 8.6$ for columns, and $\chi^2 \geq 19.4$ for interaction. Table IV shows that the significant D values were due to excessive numbers of F values at the .10, .05, and .01 levels for the above χ^2 values except when $\chi^2 \geq 26.6$. At $\chi^2 = 3.6$ in Table III, the method of unweighted means, Method 1, and Method 2 all produced significant interaction D values. Table IV shows no excessive numbers of F values at the .10 level, but two at the .05 level, and four at the .01 level for these three methods.

While all four methods produced a greater number of large F values than the equal cell method did at the same level of disproportionality, only the method of expected frequency produced significant D values at these levels except for the interaction case at $\chi^2 = 3.6$. The method of unweighted means, Method 1, and Method 2 did not appear to produce significant enough results to cause Type I errors in rows, columns, or interaction in the no effects case. In all situations presented in Table IV, identical frequencies of F values were found for the method of unweighted means and Method 1. Method 2 differed slightly in a few instances.

Effect of Changing Seed Numbers

An investigation was made into the effects of changing seed numbers in the simulation on the results. For $\chi^2 = 3.6$, a new seed number produced higher D values for rows, columns, and interactions. Overall results of significance were the same except that the method of expected frequencies had significant D values for columns and interaction and not for rows. The other three methods produced no significant D values. At $\chi^2 = 19.4$, five additional seed number results were examined. The results are in Tables V, VI, and VII.

TABLE V

D VALUES FOR NO EFFECTS CASE COMPARING THE FOUR METHODS OF HANDLING DISPROPORTIONALITY TO THE EQUAL CELL METHOD UNDER DIFFERENT SEED NUMBERS FOR $\chi^2 = 19.4$

D values					
Seed Number		Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
1st	row	.026	.139*	.026	.035
	col.	.014	.148*	.014	.019
	inter.	.038	.147*	.038	.038
2nd	row	.039	.130*	.039	.027
	col.	.019	.145*	.019	.030
	inter.	.046*	.171*	.046*	.046*
3rd	row	.028	.155*	.028	.031
	col.	.029	.148*	.029	.029
	inter.	.023	.133*	.023	.023
4th	row	.020	.107*	.020	.036
	col.	.052*	.128*	.052*	.049*
	inter.	.020	.137*	.020	.020
5th	row	.017	.123*	.017	.018
	col.	.052*	.194*	.052*	.045*

*Significant at the .05 level.

TABLE V--Continued

D values					
Seed Number		Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
	inter.	.034	.132*	.034	.034
	row	.021	.133*	.021	.022
6th	col.	.027	.113*	.027	.029
	inter.	.022	.127*	.022	.022

Table V contains D values for the four methods of handling disproportionality at χ^2 = 19.4. The D values for rows yielded the same results in terms of overall significance. The method of expected frequencies yielded significant results and the other three methods did not. For the columns, the fourth and fifth seed numbers yielded significant D values for all four methods. For the interaction D values, there was only one discrepancy in terms of significance and that was on the second seed number where all methods yielded significant D values.

Table VI contains the same general information as Table V except that the distributions of the four methods were compared to the uniform distribution in Table III.

TABLE VI

D VALUES FOR NO EFFECTS CASE COMPARING THE FOUR METHODS OF HANDLING DISPROPORTIONALITY TO THE UNIFORM DISTRIBUTION FOR χ^2 = 19.4 UNDER DIFFERENT SEED NUMBERS

Seed Number		Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
		D values			
1st	row	.033	.108*	.033	.031
	col.	.019	.146*	.019	.020
	inter.	.025	.135*	.025	.025
2nd	row	.028	.117*	.028	.019
	col.	.021	.137*	.021	.022
	inter.	.019	.133*	.019	.019
3rd	row	.016	.131*	.016	.011
	col.	.023	.134*	.023	.023
	inter.	.016	.134*	.016	.016
4th	row	.017	.110*	.017	.027
	col.	.030	.120*	.030	.030
	inter.	.014	.137*	.014	.014
5th	row	.017	.131*	.017	.015
	col.	.037	.175*	.037	.039
	inter.	.022	.144*	.022	.022
6th	row	.019	.140*	.019	.019
	col.	.043*	.105*	.043*	.041
	inter.	.018	.136*	.018	.018

*Significant at the .05 level.

Results from Table VI appear to be consistent with Table III across all seed numbers with the exception of the sixth seed number which yielded significant D values on columns for the method of unweighted means and Method 1. The Method 2 value of .041 is close to the critical value. In all cases, the method of expected frequencies yielded significant D values. Table VII contains the frequencies of F values at the .10, .05, and .01 levels of significance for six different seed numbers.

TABLE VII

FREQUENCIES OF ROW, COLUMN, AND INTERACTION F VALUES FOR THE NO EFFECTS CASE AT THE .10, .05, AND .01 LEVELS OF SIGNIFICANCE UNDER DIFFERENT SEED NUMBERS AT χ^2 = 19.4

Seed Number			Equal Cell Anova	Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
	row	.10	83	99	183	99	96
		.05	38	41	124	41	42
		.01	6	7	38	7	4
		.10	109	110	221	110	120
1st	col.	.05	49	63	147	63	64
		.01	7	9	57	9	10
		.10	89	84	179	84	84
	inter.	.05	43	49	115	49	49
		.01	6	8	45	8	8
		.10	88	104	209	104	100
	row	.05	38	41	124	41	42
		.01	6	7	38	7	4
		.10	101	95	202	95	96
2nd	col.	.05	57	51	137	51	53
		.01	13	11	49	11	12

TABLE VII--Continued

Seed Number			Equal Cell Anova	Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
	inter.	.10	87	96	186	96	96
		.05	42	55	116	55	55
		.01	11	11	45	11	11
3rd	row	.10	88	102	194	102	94
		.05	40	48	121	48	48
		.01	8	8	42	8	8
	col.	.10	114	97	205	97	97
		.05	62	43	137	43	48
		.01	9	12	44	12	13
	inter.	.10	105	98	210	98	98
		.05	49	47	121	47	47
		.01	5	9	41	9	9
4th	row	.10	104	97	186	97	96
		.05	45	52	125	52	47
		.01	10	15	46	15	14
	col.	.10	96	97	206	97	95
		.05	42	44	135	44	42
		.01	7	10	46	10	12
	inter.	.10	102	91	202	91	91
		.05	57	41	115	41	41
		.01	9	6	33	6	6
5th	row	.10	96	108	208	108	104
		.05	47	64	128	64	65
		.01	5	11	61	11	11
	col.	.10	93	122	246	122	121
		.05	45	69	168	69	66
		.01	9	7	61	7	8
	inter.	.10	110	108	213	108	108
		.05	49	52	129	52	52
		.01	11	11	46	11	11
	row	.10	99	104	210	104	102
		.05	54	47	136	47	51
		.01	11	11	43	11	15

TABLE VII--Continued

Seed Number			Equal Cell Anova	Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
		.10	96	86	196	86	82
6th	col.	.05	45	37	28	37	34
		.01	8	7	36	7	7
		.10	110	117	210	117	117
	inter.	.05	60	56	136	56	56
		.01	10	11	47	11	11

The six different seed numbers used for Table VII appear to have yielded similar results. In all situations, the method of unweighted means and Method 1 yielded identical frequencies. Method 2 yielded identical frequencies to those two methods for interaction and very close results on rows and columns. The method of expected frequencies in all cases yielded much larger frequencies than all other methods indicating a strong tendency towards Type I errors. Changing seed numbers did cause the method of equal cell analysis to yield different numbers of F values at the .10, .05, and .01 levels. However, the differences were not great; and the relative position of the other four methods appear to be very similar for each seed number.

Row Effects Case

To simulate the row effects case, cell means were established such that row effects would occur with a power of .60

and $\alpha = .05$. Column and interaction effects were not built-in and occurred only by chance. Initially ten numbers were derived for each of the four cells. Disproportionality was established in the same manner as the No Effects Case. Chi-square values were used to measure disproportionality. Each level of disproportionality was run 1000 times. Table VIII contains the D values of rows, columns, and interaction for the row case as disproportionality increases. Table X contains row, column, and interaction F frequencies at the .10, .05, and .01 levels of significance for the four methods of handling disproportionality.

TABLE VIII

D VALUES FOR THE ROW EFFECTS CASE COMPARING THE FOUR METHODS OF HANDLING DISPROPORTIONALITY TO THE EQUAL CELL ANOVA AS χ^2 INCREASES FOR ROWS, COLUMNS, AND INTERACTION

χ^2 value		D values: Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
	row	.000	.000	.000	.000
0.0	col.	.000	.000	.000	.000
	inter.	.000	.000	.000	.000
	row	.024	.013	.024	.024
1.6	col.	.016	.014	.016	.016
	inter.	.026	.021	.026	.026

*Significant at the .05 level.

TABLE VIII--Continued

χ^2 value		Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
		D values			
	row	.027	.017	.027	.027
2.6	col.	.019	.031	.019	.021
	inter.	.026	.021	.026	.026
	row	.045*	.016	.045*	.045*
3.6	col.	.034	.023	.034	.034
	inter.	.036	.017	.036	.036
	row	.084*	.026	.084*	.084*
6.4	col.	.015	.051*	.015	.015
	inter.	.030	.032	.030	.030
	row	.096*	.034	.096*	.095*
8.6	col.	.027	.075*	.025	.027
	inter.	.031	.056*	.031	.031
	row	.130*	.037	.130*	.130*
10.0	col.	.019	.086*	.019	.019
	inter.	.022	.070*	.022	.022
	row	.228*	.072*	.228*	.224*
19.4	col.	.014	.148*	.014	.019
	inter.	.038	.147*	.038	.038
	row	.983*	.970*	.986*	.987*
24.4	col.	.625*	.409*	.670*	.491*

TABLE VIII--Continued

χ^2 value		Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
		D values			
	inter.	.626*	.625*	.670*	.670*
26.6	row	.986*	.953*	.987*	.989*
	col.	.640*	.432*	.692*	.643*
	inter.	.635*	.498*	.684*	.684*
40.6	row	.981*	.942*	.987*	.989*
	col.	.714*	.533*	.741*	.693*
	inter.	.706*	.654*	.732*	.732*

TABLE IX

FREQUENCIES OF ROW, COLUMN, AND INTERACTION F VALUES FOR THE ROW EFFECTS CASE AT THE .10, .05, AND .01 LEVELS OF SIGNIFICANCE AS DISPROPORTIONALITY INCREASES

χ^2 value			Equal Cell Anova	Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
0.0	row	.10	712	712	712	712	712
		.05	579	579	579	579	579
		.01	308	308	308	308	308
	col.	.10	109	109	109	109	109
		.05	49	49	49	49	49
		.01	7	7	7	7	7
	inter.	.10	89	89	89	89	89
		.05	43	43	43	43	43
		.01	6	6	6	6	6

TABLE IX--Continued

χ^2 value			Equal Cell Anova	Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
1.6	row	.10	712	688	705	688	688
		.05	579	556	566	556	556
		.01	308	300	316	300	300
	col.	.10	109	102	108	102	102
		.05	49	51	58	51	51
		.01	7	14	15	14	14
	inter.	.10	89	93	98	93	93
		.05	43	43	45	43	43
		.01	6	13	19	13	13
2.6	row	.10	712	690	713	690	688
		.05	579	552	580	552	553
		.01	308	290	325	290	289
	col.	.10	109	109	122	109	107
		.05	49	60	63	60	58
		.01	7	15	16	15	15
	inter.	.10	89	90	103	90	90
		.05	43	46	52	46	46
		.01	6	10	11	10	10
3.6	row	.10	712	677	701	677	677
		.05	579	534	572	534	534
		.01	308	280	324	280	280
	col.	.10	109	104	111	104	104
		.05	49	55	66	55	55
		.01	7	14	19	14	14
	inter.	.10	89	89	105	89	89
		.05	43	45	51	45	45
		.01	6	9	13	9	9
6.4	row	.10	712	628	690	628	628
		.05	579	509	582	509	509
		.01	308	258	334	258	258
	col.	.10	109	96	134	96	96
		.05	49	47	64	47	47
		.01	7	8	17	8	8

TABLE IX--Continued

χ^2 value			Equal Cell Anova	Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
	inter.	.10	89	88	107	88	88
		.05	43	49	65	49	49
		.01	6	8	13	8	8
8.6	row	.10	712	628	691	628	627
		.05	579	483	575	483	484
		.01	308	235	330	235	234
	col.	.10	109	113	158	113	107
		.05	49	48	83	48	51
		.01	7	14	19	14	12
	inter.	.10	89	86	125	86	86
		.05	43	42	69	42	42
		.01	6	8	16	8	8
10.0	row	.10	712	591	675	591	591
		.05	579	449	553	449	449
		.01	308	222	329	222	222
	col.	.10	109	102	141	102	102
		.05	49	47	87	47	47
		.01	7	9	20	9	9
	inter.	.10	89	86	133	86	86
		.05	43	38	73	38	38
		.01	6	7	16	7	7
19.4	row	.10	712	488	645	488	497
		.05	579	351	536	351	355
		.01	308	143	329	143	155
	col.	.10	109	110	221	110	120
		.05	49	63	147	63	64
		.01	7	9	57	9	10
	inter.	.10	89	84	174	84	84
		.05	43	49	115	49	49
		.01	6	8	45	8	8
	row	.10	712	0	0	0	0
		.05	579	0	0	0	0
		.01	308	0	0	0	0

TABLE IX--Continued

χ^2 value			Equal Cell Anova	Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
		.10	109	0	0	0	0
24.4	col.	.05	49	0	0	0	0
		.01	7	0	0	0	0
		.10	89	0	0	0	0
	inter.	.05	43	0	0	0	0
		.01	6	0	0	0	0
		.10	712	0	0	0	0
	row	.05	579	0	0	0	0
		.01	308	0	0	0	0
		.10	109	0	0	0	0
26.6	col.	.05	49	0	0	0	0
		.01	7	0	0	0	0
		.10	89	0	0	0	0
	inter.	.05	43	0	0	0	0
		.01	6	0	0	0	0
		.10	712	0	0	0	0
	row	.05	579	0	0	0	0
		.01	308	0	0	0	0
		.10	109	0	0	0	0
40.6	col.	.05	49	0	0	0	0
		.01	7	0	0	0	0
		.10	89	0	0	0	0
	inter.	.05	43	0	0	0	0
		.01	6	0	0	0	0

The row values in Table VIII show no significant D values for any method for $\chi^2 \leq 2.6$. The method of unweighted means, Method 1, and Method 2 all gave significant D values for $\chi^2 \geq 3.6$. An examination of Table IX shows that these three methods yielded fewer F values than the equal cell method or the method of expected frequencies; and thus, they were committing Type II errors. The method of expected frequencies did the same thing but not until $\chi^2 \geq 19.4$.

Table VIII shows similar results to the No Effects Case for columns and interactions. Only the method of expected frequencies yielded significant D values for $\chi^2 \leq 19.4$. Table IX shows that these significant values were caused partly by a inordinately high number of significant F values at the .10, .05, and .01 levels. The method of expected frequencies committed Type I errors.

For $\chi^2 \geq 24.4$, all four methods yielded significant D values but no large F values. Thus, no Type I errors were being committed; and the error is in a "safe" direction.

Effect of Changing Power

Row effects were also simulated for power of .80 and .95. For $\chi^2 = 3.6$ and power = .80, the overall results for significant D values was the same as those in Table VIII. For $\chi^2 = 10.0$ and power = .80, the overall results for significant D values were also the same as those in Table VIII, and the D values were quite close to being the same. For $\chi^2 = 19.4$ and power = .80, the D values for columns and interaction were identical to those in Table VIII, and the overall results of the rows were the same.

For power of .95 and $\chi^2 = 10.0$, column and interaction D values were identical to values in Table VIII. Overall, row significant D values were the same. The same values resulted from χ^2 = 19.4 and power equal to .95.

Effect of Changing Seed Numbers

An examination of the effects of four different seed numbers on the results was made for χ^2 = 19.4 and power of .80. While D values and F frequencies varied slightly from seed number to seed number, overall results of significant D values were the same as those in Table VIII.

Column Effects Case

Column effects were created in a similar manner to the Row Effects Case. Cell means were created such that column effects would exist with a power of .60 for α = .05. Rows and interaction had no built-in effects and occurred only by chance. Disproportionality was established in the same manner as the previous two cases.

Table X contains the D values of rows, columns, and interaction for the column case as disproportionaltiy increases. Table XI contains row, column, and interaction F frequencies at the .10, .05, and .01 levels of significance for the four methods of handling disproportionality.

TABLE X

D VALUES FOR THE COLUMN EFFECTS CASE COMPARING THE FOUR METHODS OF HANDLING DISPROPORTIONALITY TO THE EQUAL CELL ANOVA AS χ^2 INCREASES FOR ROWS, COLUMNS, AND INTERACTIONS

χ^2 value		D values: Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
	row	.000	.000	.000	.000
0.0	col.	.000	.000	.000	.000
	inter.	.000	.000	.000	.000
	row	.020	.024	.020	.020
1.6	col.	.021	.017	.022	.021
	inter.	.026	.021	.026	.026
	row	.011	.026	.011	.013
2.6	col.	.036	.011	.036	.035
	inter.	.026	.021	.026	.026
	row	.021	.041	.021	.021
3.6	col.	.044*	.011	.044*	.044*
	inter.	.036	.017	.036	.036
	row	.034	.057*	.034	.034
6.4	col.	.068*	.027	.068*	.068*
	inter.	.030	.032	.030	.030
	row	.023	.076*	.023	.028
8.6	col.	.092*	.036	.092*	.086*
	inter.	.031	.056*	.031	.031

*Significant at the .05 level.

TABLE X--Continued

χ^2 value		Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
		D values			
	row	.030	.089*	.030	.030
10.0	col.	.113*	.033	.113*	.113*
	inter.	.022	.070*	.022	.022
	row	.026	.139*	.026	.035
19.4	col.	.234*	.086*	.234*	.226*
	inter.	.038	.147*	.038	.038
	row	.614*	.506*	.671*	.687*
24.4	col.	.968*	.927*	.970*	.930*
	inter.	.627*	.627*	.671*	.671*
	row	.642*	.421*	.689*	.748*
26.6	col.	.954*	.901*	.962*	.955*
	inter.	.682*	.520*	.684*	.684*
	row	.708*	.474*	.751*	.751*
40.6	col.	.981*	.939*	.981*	.974*
	inter.	.638*	.632*	.730*	.730*

TABLE XI

FREQUENCIES OF ROW, COLUMN, AND INTERACTION F VALUES FOR THE COLUMN EFFECTS CASE AT THE .10, .05, AND .01 LEVELS OF SIGNIFICANCE AS DISPROPORTIONALITY INCREASES

χ^2 value			Equal Cell Anova	Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
		.10	83	83	83	83	83
	row	.05	38	38	38	38	38
		.01	6	6	6	6	6
		.10	707	707	707	707	707
0.0	col.	.05	576	576	576	576	576
		.01	320	320	320	320	320
		.10	89	89	89	89	89
	inter.	.05	43	43	43	43	43
		.01	6	6	6	6	6
		.10	83	72	88	72	72
	row	.05	38	36	37	36	36
		.01	6	6	7	6	6
		.10	707	692	705	692	692
1.6	col.	.05	576	581	593	581	581
		.01	320	299	320	298	299
		.10	89	93	98	93	93
	inter.	.05	43	43	45	43	43
		.01	6	13	19	13	13
		.10	83	81	92	81	80
	row	.05	38	39	45	39	40
		.01	6	6	6	6	6
		.10	707	682	705	682	672
2.6	col.	.05	576	549	576	549	547
		.01	320	284	317	284	287
		.10	89	90	103	90	90
	inter.	.05	43	46	52	46	46
		.01	6	10	11	10	10

TABLE XI--Continued

χ^2 value			Equal Cell Anova	Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
		.10	83	75	93	75	75
	row	.05	38	37	44	37	37
		.01	6	7	7	7	7
		.10	707	674	709	674	674
3.6	col.	.05	576	543	579	543	543
		.01	320	276	317	276	276
		.10	89	89	105	89	89
	inter.	.05	43	45	51	45	45
		.01	6	9	13	9	9
		.10	83	80	102	80	80
	row	.05	38	40	61	40	40
		.01	6	6	12	6	6
		.10	707	647	691	647	641
6.4	col.	.05	576	508	577	508	508
		.01	320	270	347	269	270
		.10	89	88	107	88	88
	inter.	.05	43	49	65	49	49
		.01	6	8	13	8	8
		.10	83	77	111	77	77
	row	.05	38	39	61	39	41
		.01	6	8	15	8	8
		.10	707	615	676	615	621
8.6	col.	.05	576	495	577	495	494
		.01	320	253	346	253	249
		.10	89	86	125	86	86
	inter.	.05	43	42	69	42	42
		.01	6	8	16	8	8
		.10	83	87	127	87	87
	row	.05	38	36	74	36	36
		.01	6	7	17	7	7
		.10	707	606	681	606	606
10.0	col.	.05	576	463	578	463	463
		.01	320	243	352	243	243

TABLE XI--Continued

χ^2 value			Equal Cell Anova	Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
	inter.	.10	89	86	133	86	86
		.05	43	38	73	38	38
		.01	6	7	16	7	7
	row	.10	83	99	183	99	96
		.05	38	41	124	41	42
		.01	6	7	38	7	4
19.4	col.	.10	707	473	645	473	481
		.05	576	361	522	361	357
		.01	320	159	346	158	160
	inter.	.10	89	84	179	84	84
		.05	43	49	115	49	49
		.01	6	8	45	8	8
	row	.10	83	0	0	0	0
		.05	38	0	0	0	0
		.01	6	0	0	0	0
24.4	col.	.10	707	0	0	0	0
		.05	576	0	0	0	0
		.01	320	0	0	0	0
	inter.	.10	89	0	0	0	0
		.05	43	0	0	0	0
		.01	6	0	0	0	0
	row	.10	83	0	0	0	0
		.05	38	0	0	0	0
		.01	6	0	0	0	0
26.6	col.	.10	707	0	0	0	0
		.05	576	0	0	0	0
		.01	320	0	0	0	0
	inter.	.10	89	0	0	0	0
		.05	43	0	0	0	0
		.01	6	0	0	0	0
	row	.10	83	0	0	0	0
		.05	38	0	0	0	0
		.01	6	0	0	0	0

TABLE XI--Continued

χ^2 value			Equal Cell Anova	Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
		.10	707	0	0	0	0
40.6	col.	.05	576	0	0	0	0
		.01	320	0	0	0	0
		.10	89	0	0	0	0
	inter.	.05	43	0	0	0	0
		.01	6	0	0	0	0

For $\chi^2 \leq 2.6$, there were no significant D values. The method of unweighted means, Method 1, and Method 2 produced significant D values for $\chi^2 \geq 3.6$ for columns. An examination of Table XI reveals that for $3.6 \leq \chi^2 \leq 19.4$ these three methods yielded fewer significant F values than the equal cell method; and thus, they were committing Type II errors. The method of expected frequencies had significant D values for $\chi^2 \geq 19.4$. Table XI reveals that at $\chi^2 = 19.4$, this method committed Type II errors also.

For $\chi^2 \geq 24.4$, all four methods produced significant D values and zero significant F values. Thus, all four methods were committing Type II errors for these Chi-square values.

Table X and Table XI show the results of the four methods when no effect was built-in to rows or interaction. All four methods yielded significant D values for $\chi^2 \geq 24.4$. An examination of Table XI shows that the number of significant F values for these methods was zero. However, since no effects were built-in to rows and interactions, these methods were not committing Type I errors. The method of expected frequencies yielded significant D values for $\chi^2 \geq 6.4$ for rows and $\chi^2 \geq 8.6$ for interaction. Table XI shows that these were caused by an exceedingly large number of significant F values. Thus, Type I errors were committed.

The methods of unweighted means, Method 1, and Method 2 yielded no other significant D values than those previously mentioned. Table XI reveals that the method of unweighted means and Method 1 yielded virtually the same number of significant F values for rows and interactions. Method 2 results were extremely close.

Effect of Changing Power

Column effects were also simulated for power of .80 and .95. For power of .80 and $\chi^2 = 3.6$, row and interaction D values were identical to those produced by the power of .60 and $\chi^2 = 3.6$ situation in Table X. The column D values were different but yielded the same overall significance. For $\chi^2 = 10.0$ and power of .80, the D values for columns were close to those from power of .60 and produced the same overall results as Table X. The row and interaction D values were identical to Table X. For $\chi^2 = 19.4$ and power of .80, D values for rows and interaction were again identical to the .60 power values in Table X. The column values were extremely close to the Table X values. Overall significance was the same.

For power of .95 and χ^2 = 10.0 and 19.4, the row and interaction D values were identical to the power of .60 values in Table X. Column D values produced the same overall significance with all four methods significantly different from the equal cell case.

Effect of Changing Seed Numbers

An examination of the effects of four different seed numbers on the results was made for χ^2 = 19.4 and power of .80. Results for three of the four seed numbers were the same in terms of the number of significant D values. One seed number, however, produced all four methods significantly different from the equal cell method on rows while the other three seed numbers resulted in only a significant D value for the method of expected frequencies.

Interaction Effects Case

Interaction effects were created in the same manner that the row case and column case were simulated. Cell means were produced such that interaction effects would occur with a power of .60 at α = .05. There were no built-in row or column effects, and any that occurred did so by chance. Disproportionality was established in the same manner as before.

Table XII contains the D values of rows, columns, and interaction for the interaction case as disproportionality increases. Table XIII contains row, column, and interaction

F frequencies at the .10, .05, and .01 levels of significance for the four methods of handling disproportionality.

TABLE XII

D VALUES FOR THE INTERACTION EFFECTS CASE COMPARING THE FOUR METHODS OF HANDLING DISPROPORTIONALITY TO THE EQUAL CELL ANOVA AS χ^2 INCREASES FOR ROWS, COLUMNS, AND INTERACTION

χ^2 value		D values: Method of Unwtd. Means	D values: Method of Exp. Freq.	D values: Method 1	D values: Method 2
0.0	row	.000	.000	.000	.000
	col.	.000	.000	.000	.000
	inter.	.000	.000	.000	.000
1.6	row	.020	.024	.020	.020
	col.	.016	.014	.016	.016
	inter.	.015	.014	.015	.015
2.6	row	.014	.027	.014	.021
	col.	.020	.038	.020	.031
	inter.	.025	.016	.025	.025
3.6	row	.021	.041	.021	.021
	col.	.034	.023	.034	.034
	inter.	.047*	.014	.047*	.047*
6.4	row	.034	.057*	.034	.034
	col.	.015	.051*	.015	.015
	inter.	.100*	.038	.100*	.100*

*Significant at the .05 level.

TABLE XII--Continued

χ^2 value		Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
		D values			
7.4	row	.029	.070*	.029	.097*
	col.	.020	.194*	.020	.211*
	inter.	.131*	.074*	.131*	.131*
8.6	row	.025	.076*	.025	.024
	col.	.027	.089*	.027	.043*
	inter.	.131*	.054*	.131*	.131*
19.4	row	.027	.189*	.027	.058*
	col.	.017	.177*	.017	.018
	inter.	.270*	.133*	.270*	.270*
26.6	row	.624*	.420*	.688*	.727*
	col.	.640*	.448*	.692*	.643*
	inter.	.977*	.961*	.978*	.976*
40.6	row	.722*	.524*	.751*	.751*
	col.	.707*	.550*	.741*	.693*
	inter.	.983*	.976*	.985*	.985*
59.6	row	.746*	.452*	.787*	.688*
	col.	.738*	.525*	.788*	.846*
	inter.	.978*	.974*	.991*	.991*

TABLE XIII

FREQUENCIES OF ROW, COLUMN, AND INTERACTION F VALUES FOR THE INTERACTION EFFECTS CASE AT THE .10, .05, AND .01 LEVELS OF SIGNIFICANCE AS DISPROPORTIONALITY INCREASES

χ^2 value			Equal Cell Anova	Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
0.0	row	.10	83	83	83	83	83
		.05	38	38	38	38	38
		.01	6	6	6	6	6
	col.	.10	109	109	109	109	109
		.05	49	49	49	49	49
		.01	7	7	7	7	7
	inter.	.10	721	721	721	721	721
		.05	589	589	589	589	589
		.01	317	317	317	317	317
1.6	row	.10	83	72	78	72	72
		.05	38	36	37	36	36
		.01	6	6	7	6	6
	col.	.10	109	102	108	102	102
		.05	49	51	58	51	51
		.01	7	14	15	14	14
	inter.	.10	721	713	723	713	713
		.05	589	585	598	585	585
		.01	317	308	322	308	308
2.6	row	.10	83	80	92	80	84
		.05	38	39	44	39	40
		.01	6	6	6	6	5
	col.	.10	109	113	128	113	118
		.05	49	59	68	59	61
		.01	7	16	20	16	16
	inter.	.10	721	696	712	696	696
		.05	589	564	591	564	564
		.01	317	296	333	296	296

TABLE XIII--Continued

χ^2 value			Equal Cell Anova	Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
3.6	row	.10	83	75	93	75	75
		.05	38	37	44	37	37
		.01	6	7	7	7	7
	col.	.10	109	104	111	104	104
		.05	49	55	66	55	55
		.01	7	14	19	14	14
	inter.	.10	721	701	725	701	701
		.05	589	550	592	550	550
		.01	317	270	303	270	270
6.4	row	.10	83	80	102	80	80
		.05	38	40	61	40	40
		.01	6	7	12	6	6
	col.	.10	109	96	134	96	96
		.05	49	47	64	47	47
		.01	7	8	17	8	8
	inter.	.10	721	633	693	633	633
		.05	589	489	567	489	489
		.01	317	237	309	237	237
7.4	row	.10	83	80	110	80	125
		.05	38	48	64	48	61
		.01	6	8	17	8	20
	col.	.10	109	111	255	111	260
		.05	49	67	153	67	153
		.01	7	14	58	14	50
	inter.	.10	721	597	647	597	597
		.05	589	458	541	458	458
		.01	317	199	290	199	199
8.6	row	.10	83	75	122	75	76
		.05	38	37	65	37	41
		.01	6	8	20	8	8
	col.	.10	109	117	176	116	133
		.05	49	48	101	48	57
		.01	7	12	25	12	11

TABLE XIII--Continued

χ^2 value			Equal Cell Anova	Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
	inter.	.10	721	590	673	590	590
		.05	589	459	541	459	459
		.01	317	221	308	221	221
	row	.10	83	99	243	99	113
		.05	38	46	154	46	56
		.01	6	3	53	3	6
		.10	109	123	250	123	118
19.4	col.	.05	49	61	168	61	63
		.01	7	12	69	12	10
		.10	721	451	590	451	451
	inter.	.05	589	331	493	331	331
		.01	317	136	308	136	136
		.10	83	0	0	0	0
	row	.05	38	0	0	0	0
		.01	6	0	0	0	0
		.10	109	0	0	0	0
26.6	col.	.05	49	0	0	0	0
		.01	7	0	0	0	0
		.10	721	0	0	0	0
	inter.	.05	589	0	0	0	0
		.01	317	0	0	0	0
		.10	83	0	0	0	0
	row	.05	38	0	0	0	0
		.01	6	0	0	0	0
		.10	109	0	0	0	0
40.6	col.	.05	49	0	0	0	0
		.01	7	0	0	0	0
		.10	721	0	0	0	0
	inter.	.05	589	0	0	0	0
		.01	317	0	0	0	0
		.10	83	0	0	0	0
	row	.05	38	0	0	0	0
		.01	6	0	0	0	0

TABLE XIII--Continued

χ^2 value			Equal Cell Anova	Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
59.6	col.	.10	109	0	0	0	0
		.05	49	0	0	0	0
		.01	7	0	0	0	0
	inter.	.10	721	0	0	0	0
		.05	589	0	0	0	0
		.01	317	0	0	0	0

Table XII shows that for $\chi^2 \geq 3.6$ the method of unweighted means, Method 1, and Method 2 all yielded significant D values for interaction. The method of expected frequencies yielded significant D values for $\chi^2 \geq 8.6$. An examination of Table XIII shows that in each of these cases, there were fewer significant F values at the .10, .05, and .01 levels than the equal cell method yielded. This is indicative of the occurrence of Type II errors in each case.

The row and column values in Table XII show that for $\chi^2 \geq 6.4$, the method of expected frequencies yielded significant D values. Table XIII shows that this was due to an exceedingly large number of significant F values for this method except for $\chi^2 \geq 26.6$. Thus, Type I errors are being committed for $6.4 \leq \chi^2 \leq 26.6$. The method of unweighted means and Method 1 produced significant D values for $\chi^2 \geq 26.6$ for both rows and columns. Table XIII reveals that there were zero significant F values for $\chi^2 \geq 26.6$ for all four methods. Thus, no Type I errors were being committed there. Table XII shows that for rows, Method 2 yielded significant D values for $\chi^2 = 7.4$ and $\chi^2 \geq 19.4$. For $\chi^2 =$ 7.4 and 19.4, Table XIII shows that Method 2 yielded too many significant F values at the .05 level. Thus, Type I errors were being committed. For the columns, Table XII shows that Method 2 yielded significant D values at $\chi^2 = 7.4$ and 8.6 and $\chi^2 \geq 26.6$. Table XIII shows that for $\chi^2 = 8.6$, Method 2 yielded too many significant F values. Thus, Type I errors were committed.

Effect of Changing Power

Interaction effects were also simulated for power of .80 and .95. For power of .80 and $\chi^2 = 3.6$, row and column D values were identical to those for power of .60 in Table XII for $\chi^2 = 3.6$. While the D values differed for interactions, the overall results were still the same in terms of which methods were significantly different. For $\chi^2 = 7.4$ and power of .80, all D values were different but overall results of significance were the same as power of .60 in Table XII. For $\chi^2 = 19.4$ and power of .80, the D values were different but overall results were the same.

For $\chi^2 = 7.4$ and power of .95, D values were different from those of power of .60 and $\chi^2 = 7.4$ in Table XII. For rows and interaction, the overall significant D values were the same. However, for columns, all four methods were significantly different for $\chi^2 = 7.4$ and power of .95 whereas in Table XII only the method of expected frequencies and Method 2 are significant. For $\chi^2 = 19.4$ and power of .95, overall significant D values were the same as those in Table XII for $\chi^2 = 19.4$.

Effect of Changing Seed Numbers

An examination of the effects of three different seed numbers on the results was made for $\chi^2 = 7.4$ and power of .80. Two of the three seed numbers yielded the same overall results as those in Table XII. One seed number produced a nonsignificant D value for the method of expected frequencies for rows which Table XII did not. Otherwise, changing seed numbers made no difference in the outcome of significant D values.

Row and Column Effects Case

Row and column effects were created in the same manner that the row, column, and interaction cases were. Cell means were produced such that row effects and column effects would occur with a power of .60 at $\alpha = .05$. There were no built-in interaction effects, and any that occurred did so by chance. Disproportionality was established in the same manner as before.

Table XIV contains the D values of rows, columns, and interaction for the row and column case as disproportionality increases. Table XV contains row, column, and interaction

F frequencies at the .10, .05, and .01 levels of significance for the four methods of handling disproportionality.

TABLE XIV

D VALUES FOR THE ROW AND COLUMN EFFECTS CASE COMPARING THE FOUR METHODS OF HANDLING DISPROPORTIONALITY TO THE EQUAL CELL ANOVA AS χ^2 INCREASES FOR ROWS, COLUMNS, AND INTERACTION

		D values			
χ^2 value		Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
	row	.000	.000	.000	.000
0.0	col.	.000	.000	.000	.000
	inter.	.000	.000	.000	.000
	row	.024	.013	.024	.024
1.6	col.	.021	.017	.022	.021
	inter.	.026	.021	.026	.026
	row	.027	.017	.027	.026
2.6	col.	.036	.011	.036	.035
	inter.	.026	.021	.026	.026
	row	.045*	.016	.045*	.045*
3.6	col.	.044*	.011	.044*	.044*
	inter.	.036	.017	.036	.036
	row	.084*	.026	.084*	.084*
6.4	col.	.068*	.027	.068*	.068*
	inter.	.030	.032	.030	.030

*Significant at the .05 level.

TABLE XIV--Continued

χ^2 value		Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
		D values			
	row	.096*	.033	.096*	.092*
8.6	col.	.092*	.036	.092*	.086*
	inter.	.031	.056*	.031	.031
	row	.131*	.037	.131*	.131*
10.0	col.	.113*	.033	.113*	.113*
	inter.	.022	.070*	.022	.022
	row	.228*	.072*	.228*	.223*
19.4	col.	.225*	.082*	.225*	.226*
	inter.	.038	.147*	.038	.038
	row	.983*	.970*	.986*	.987*
24.4	col.	.966*	.924*	.970*	.930*
	inter.	.620*	.619*	.665*	.665*
	row	.986*	.953*	.987*	.989*
26.6	col.	.952*	.895*	.962*	.955*
	inter.	.636*	.508*	.684*	.684*

TABLE XV

FREQUENCIES OF ROW, COLUMN, AND INTERACTION F VALUES FOR THE ROW AND COLUMN EFFECTS CASE AT THE .10, .05, AND .01 LEVELS OF SIGNIFICANCE AS DISPROPORTIONALITY INCREASES

χ^2 value			Equal Cell Anova	Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
0.0	row	.10	712	712	712	712	712
		.05	579	579	579	579	579
		.01	308	308	308	308	308
	col.	.10	707	707	707	707	707
		.05	576	576	576	576	576
		.01	320	320	320	320	320
	inter.	.10	89	89	89	89	89
		.05	43	43	43	43	43
		.01	6	6	6	6	6
1.6	row.	.10	712	688	705	688	688
		.05	579	555	566	555	555
		.01	308	300	316	300	300
	col.	.10	707	692	705	692	692
		.05	576	581	593	581	581
		.01	320	299	320	298	299
	inter.	.10	89	93	98	93	93
		.05	43	43	45	43	43
		.01	6	13	19	13	13
2.6	row	.10	712	690	713	690	688
		.05	579	552	580	552	553
		.01	308	290	325	290	289
	col.	.10	707	682	705	682	672
		.05	576	549	576	549	547
		.01	320	284	317	284	287
	inter.	.10	89	90	103	90	90
		.05	43	46	52	46	46
		.01	6	10	11	10	10

TABLE XV--Continued

χ^2 value			Equal Cell Anova	Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
3.6	row	.10	712	677	701	677	677
		.05	579	534	572	534	534
		.01	308	280	324	280	280
	col.	.10	707	674	709	674	674
		.05	576	543	579	543	543
		.01	320	276	317	276	276
	inter.	.10	89	89	103	89	89
		.05	43	45	51	45	45
		.01	6	9	13	9	9
6.4	row	.10	712	628	690	628	628
		.05	579	509	582	509	509
		.01	308	258	334	258	258
	col.	.10	707	647	691	647	647
		.05	576	508	577	508	508
		.01	320	269	347	269	269
	inter.	.10	89	88	107	88	88
		.05	43	49	65	49	49
		.01	6	8	13	8	8
8.6	row	.10	712	630	693	630	627
		.05	579	483	575	483	487
		.01	308	235	330	235	235
	col.	.10	707	615	676	615	621
		.05	576	495	577	495	494
		.01	320	253	346	253	249
	inter.	.10	89	86	125	86	86
		.05	43	42	69	42	42
		.01	6	8	16	8	8
10.0	row	.10	712	591	675	591	591
		.05	579	448	553	448	448
		.01	308	222	329	222	222
	col.	.10	707	606	681	606	606
		.05	576	463	578	463	463
		.01	320	243	352	243	243

TABLE XV--Continued

χ^2 value			Equal Cell Anova	Method of Unwtd. Means	Method of Exp. Freq.	Method 1	Method 2
	inter.	.10	89	86	133	86	86
		.05	43	38	73	38	38
		.01	6	7	16	7	7
	row	.10	712	494	648	494	498
		.05	579	351	538	351	356
		.01	308	145	332	145	157
		.10	707	482	648	482	481
19.4	col.	.05	576	369	526	369	361
		.01	320	161	348	161	160
		.10	89	84	179	84	84
	inter.	.05	43	49	115	49	49
		.01	6	8	45	8	8
	row	.10	712	0	0	0	0
		.05	579	0	0	0	0
		.01	308	0	0	0	0
		.10	707	0	0	0	0
24.4	col.	.05	576	0	0	0	0
		.01	320	0	0	0	0
		.10	89	0	0	0	0
	inter.	.05	43	0	0	0	0
		.01	6	0	0	0	0
	row	.10	712	0	0	0	0
		.05	579	0	0	0	0
		.01	308	0	0	0	0
		.10	707	0	0	0	0
26.6	col.	.05	576	0	0	0	0
		.01	320	0	0	0	0
		.10	89	0	0	0	0
	inter.	.05	43	0	0	0	0
		.01	6	0	0	0	0

In analyzing the data from Table XIV, the rows and columns data will be examined separately from interaction due to the built-in row and column effect. In Table XIV, the overall significance of D values followed the same pattern for both rows and columns. For $\chi^2 \leq 2.6$, there are no significant D values. For $\chi^2 \geq 3.6$, the method of unweighted means, Method 1, and Method 2 produced significant D values. An examination of Table XV reveals that for $3.6 \leq \chi^2 \leq 19.4$ there were fewer significant F values for these three methods than the equal cell method for rows and columns. This is an indication of the presence of Type II errors since there were built-in effects for rows and columns. All four methods produced significant D values for $\chi^2 = 24.4$ and 26.6. Table XV shows that the frequencies of significant F values for these levels of disproportionality were zero. Thus, Type II errors were committed. The method of expected frequencies had a significant D value at $\chi^2 = 19.4$. Table XV shows that this method yielded fewer significant F values than the equal cell method. Type II errors were committed.

For interaction, Table XIV shows that all four methods produced significant D values for $\chi^2 \geq 24.4$. Table XV shows that all four methods produced zero significant F values at these levels of disproportionality. Since no effect was built-in for interaction, no Type I error was committed. Otherwise, the only significant D values were for the method of expected frequencies for $\chi^2 \geq 8.6$. Table XV shows that with the exception of $\chi^2 \geq 24.4$, the method of expected frequencies produced many more significant F values at the .10, .05, and .01 levels than the equal cell method for interaction. Type I errors were committed.

Effect of Changing Power

Row and column effects were also simulated for power of .80 and .95. For power of .80 and $\chi^2 = 3.6$, interaction D values were identical to those in Table XIV and row and column overall results of significance were the same. For $\chi^2 = 10.0$ and power of .80, interaction D values were again identical to the power of .60 D values in Table XIV. Row and column results were the same. $\chi^2 = 19.4$ and power of .80 produced similar results. Row and column results were the same as those in Table XIV. Interaction D values were identical to those in Table XIV.

For $\chi^2 = 19.4$ and power of .95, interaction D values were identical to those in Table XIV for $\chi^2 = 19.4$ and power of .60. Row and column overall results were the same as Table XIV.

Effect of Changing Seed Numbers

Four different seed numbers were used to achieve $\chi^2 = 19.4$ and a power of .80. All four seed numbers produced the same overall results. All four methods produced significant D values for rows and columns on all seed numbers. For interaction, all four seed numbers produced significant D values for the method of expected frequencies but not for the other three methods.

CHAPTER V

SUMMARY, FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS

Summary

Traditional methods of computing analysis of variance for two-way designs fail when disproportionate cell frequencies occur. At least eight methods of handling this problem have been mentioned in the literature. These include the method of discarding data, the method of estimating missing data, the method of unweighted means, the method of expected cell frequencies, the method of weighted means, Method 1 (complete linear-model analysis - a multiple regression model), Method 2 (experimental-design analysis - a multiple regression model), and Method 3 (step-down analysis - a multiple regression model). Four of these methods were selected for this study: the method of unweighted means, the method of expected cell frequencies, Method 1, and Method 2.

A Monte Carlo study was conducted to determine the effects of varied disproportionality on these four methods for a two by two factorial design fixed model. Probability distributions of F values for these four methods were compared to an equal cell method by using Kolmogorov-Smirnov tests. Chi-square values were used to measure disproportionality.

Five cases were examined: the no effects case, the row effects case, the column effects case, the interaction effects case, and the row and column effects case. These effects were generated through the use of noncentral F distributions. The cases were used to provide information concerning Type I and Type II errors.

Kolmogorov-Smirnov D values and their significance were presented in tabular form. Frequencies of F values at the .10, .05, and .01 levels of significance were presented for all cases. In each case, an examination of several seed numbers and their effects on results were presented. In cases with built-in effects, power values were changed in order to examine the effect of power on the results.

Findings

It was hypothesized that Method 1 and Method 2 would give diverging results as disproportionality increased. This did not occur in this study. In four of the cases studied (No Effects, Row Effects, Column Effects and Row and Column Effects), Method 1 and Method 2 did not differ in the number of significant D values produced. In the Interaction Effects case, Method 2 produced significant D values three times when Method 1 did not out of thirty-three values. In general, Method 1 and Method 2 produced quite similar results for all levels of disproportionality with a slight divergence as disproportionality increased.

A second hypothesis was that the unweighted means analysis and the method of expected frequencies would give diverging results as disproportionality increased. This was true in this study. As disproportionality increased, these two techniques diverged in every case. Where effects were built-in, the method of expected frequencies maintained much lower D values as Chi-square increased. When no effects were simulated, the unweighted means analysis maintained much lower D values as Chi-square increased.

The third hypothesis was that for moderate levels of disproportionality ($3.6 \leq \chi^2 \leq 19.4$), Method 1 and the unweighted means analysis would give less spurious results than Method 2 and the method of expected frequencies. Method 1 and the unweighted means analysis yielded almost exactly the same D values in all cases for low through moderate disproportionality. The number and location of significant D values was in all cases the same for these two methods. Method 2 also yielded almost the same results. Only three times in 162 levels of disproportionality in five cases did Method 2 yield significant results when Method 1 and the unweighted means analysis did not. The method of expected frequencies yielded much different results than the other three methods. However, this does not mean it yielded more spurious results.

When no effects were simulated in a given case, the method of expected frequencies consistently yielded spurious results for moderate disproportionality. Type I errors were commonly committed. Rarely did Method 1, Method 2, or the unweighted means analysis yield spurious results when no effects were simulated and disproportionality was moderate.

However, when effects were simulated, Method 1, Method 2, and the unweighted means analysis consistently produced spurious results for moderate disproportionality. Type II errors were commonly committed by these three methods. The method of expected frequencies rarely committed Type II errors for moderate disproportionality.

The fourth hypothesis was that for extreme levels of disproportionality ($\chi^2 \geq 19.4$), all four methods would yield results that tend to converge on each other. In all cases, all four methods yielded zero number of F values at the .10 level or lower. All were significantly different from the equal cell case.

The fifth hypothesis that for extreme levels of disproportionality, all four methods would give results that are spurious, seems to be true. Because of the zero number of F values when an effect was simulated, Type II errors were committed in all cases for extreme disproportionality. However, when no effects were present, no Type I errors were committed. A closer examination of the F probability distributions showed that small F values were in abundance, and the F probability distributions were greatly skewed towards the small probabilities (large F values). For extreme disproportionality, all four methods seem to yield an extremely large number of small F values.

The sixth hypothesis was that there would be a point of disproportionality at which one or more of the four methods would give spurious results. In every case where an effect was simulated, at least one and in many instances three methods produced spurious results beginning at $\chi^2 = 3.6$. This value has a probability level of about .06. In all no effects cases, the first spurious results occurred for interaction at $\chi^2 = 8.6$. This value has a probability level of about .006. In all no effects cases, the first spurious results for rows and columns occurred at $\chi^2 = 6.4$ with a probability level of about .01.

Conclusions

Based on this study (within the context of the given parameter) several conclusions were reached.

1. For small levels of disproportionality ($\chi^2 < 3.6$), all four methods will yield similar nonspurious results; and thus, any of the four methods would be appropriate for use.

2. For moderate levels of disproportionality ($3.6 \leq \chi^2 \leq 19.6$), Method 1 and the unweighted means analysis appear to be the best methods to use to control Type I errors. The method of expected frequencies is the best method for control over Type II errors.

3. For extreme levels of disproportionality ($\chi^2 >$ 19.6), none of the four methods is appropriate for use.

4. Method 1 and the unweighted means analysis yield similar enough results that the researcher can use either method with the same success in a given situation.

5. There is little difference between Method 1 and Method 2.

6. At a Chi-square with a probability level of less than or equal to .06, at least one of the four methods yields spurious results in all cases.

Recommendations for Further Research

It is suggested that further research could be done in several areas of this study. Other designs besides a two by two should be investigated to see if these results still hold. Violations to the assumption of equal variance could be examined under these conditions. Mixed and random models could be studied to determine how these methods of handling disproportionality react. Factorial designs other than two-way need to be examined for the effect of disproportionality on these four methods. Other numbers of values per design could be examined. With a larger number of values, a more continuous distribution of potential Chi-square values could be achieved. Thus, there would be more levels of disproportionality to examine.

APPENDIX A

CELL SIZES FOR CHI-SQUARE LEVELS OF DISPROPORTIONALITY

χ^2 values	Number in Cell			
	one	two	three	four
0.0	10	10	10	10
1.6	8	12	12	8
2.6	7	13	12	8
3.6	7	13	13	7
6.4	6	14	14	6
7.4	4	16	11	9
8.6	6	13	16	5
10.0	5	15	15	5
19.4	4	13	20	3
24.4	5	9	3	23
26.6	2	22	13	3
40.6	3	8	27	2
59.6	5	1	3	31

APPENDIX B

THE COMPUTER PROGRAM

```
//BLACK     JOB (2064-2119,25,02),'BLACK',CLASS=K
// EXEC SUBCLG
//FORT.SYSIN DD *
C THE FIRST DATA CARD SHOULD BE THE SEED NUMBER IN COLS. 6 - 13
C THE SECOND DATA CARD SHOULD HAVE THE SEED NUMBER IN COLS. 6 - 13 AND
C THE FOUR CELL SIZE NUMBERS IN COLS. 16 - 23 WITH CELLS IN ROW ONE
C FIRST, ROW TWO SECOND.
C THE THIRD DATA CARD SHOULD HAVE THE NUMBER OF RUNS DESIRED IN COLS. 1
C - 5.
C THIS PROGRAM IS SET UP FOR 2X2 DESIGNS ONLY.  IT MUST BE MODIFIED
C INTERNALLY FOR OTHER DESIGNS.
C TO CHANGE THE NUMBER OF RUNS DESIRED, IN ADDITION TO CHANGING DATA
C THREE, THE 33RD STATEMENT, M = 1,N, MUST BE CHANGED AND 156TH STATE-
C MENT, MTOT = 1,N, MUST BE CHANGED.
C THIS PROGRAM IS SET TO RUN FOR A MEAN = 10 AND A STANDARD DEVIATION OF
C 2 (AM AND S).
C CURRENTLY 'GO TO' STATEMENTS HAVE BEEN INSERTED TO OMIT PRINTING OUT
C INDIVIDUAL METHOD RESULTS.
C THE EQUAL CELL PROGRAM IS IN ONE LOOP AND THE OTHER FOUR METHODS ARE
C IN A SECOND LOOP.
      IMPLICIT REAL*8 (A-H,O-Z)
      DIMENSION SUM2(2,2)
      DIMENSION ANOVAT(2,2,20),CELLN(2,2),SUM(2,2),RSUM(2),CSUM(2)
      DIMENSION ROWN(2),COLN(2)
      DIMENSION X(45),XM(45),XS(45),CNT(45),BW(45)
      DIMENSION RSQ(50),ESS(50),COR(45,45),RW(45,45)
      DIMENSION XRUWKS(21),XREFKS(21),XRM1KS(21),XRM2KS(21)
      DIMENSION XCUWKS(21),XCEFKS(21),XCM1KS(21),XCM2KS(21)
      DIMENSION XIUWKS(21),XIEFKS(21),XIM1KS(21),XIM2KS(21)
      DIMENSION TITLE(20),LABEL(5),FMT(20),MF(99),MFA(14),MFB(14)
      REAL MDFR(21),MDFC(21),MDFI(21),MFR(21),MFC(21),MFI(21)
      REAL MFREF(21),MFCEF(21),MFIEF(21),MF2M1(21),MF3M1(21),MF1M1(21)
      REAL MF5M2(21),MF4M2(21),MF1M2(21)
      READ(5,7)IX
    7 FORMAT(5X,I8)
      DO 90 IDL=1,21
      MFR(IDL)=0.0
      MFC(IDL)=0.0
      MFI(IDL)=0.0
      MFREF(IDL)=0.0
      MFCEF(IDL)=0.0
      MFIEF(IDL)=0.0
      MF2M1(IDL)=0.0
      MF3M1(IDL)=0.0
      MF1M1(IDL)=0.0
      MF5M2(IDL)=0.0
      MF4M2(IDL)=0.0
      MF1M2(IDL)=0.0
      MDFR(IDL)=0.0
      MDFC(IDL)=0.0
      MDFI(IDL)=0.0
   90 CONTINUE
```

```
      DO 100 M=1,1000
      DO 1 I=1,2
      DO 1 J=1,2
      DO 1 K=1,10
      ANOVAT(I,J,K)=0.0
    1 CONTINUE
      S=2.0
      AM=10.0
      V=0.0
C     WRITE(6,5)
C   5 FORMAT(1H1)
      DO 30 I=1,2
      DO 30 J=1,2
      DO 30 L=1,10
C     CALL GAUSS(IX,S,AM,V)
C     SUBROUTINE GAUSS(IX,S,AM,V)
      A=0.0
      DO 50 K=1,12
      CALL RANDME(IX,Y)
   50 A=A+Y
      V=(A-6.0)*S+AM
C     RETURN
C     END
      ANOVAT(I,J,L)=V
   30 CONTINUE
      GO TO 870
  411 WRITE(6,6)
    6 FORMAT(20X,'ANOVA TABLE',////)
      DO 60 K=1,10
      DO 60 J=1,2,2
      I=1
      WRITE(6,2) ANOVAT(I,J,K), ANOVAT(I,J+1,K)
    2 FORMAT(10X,F8.5,12X,F8.5)
   60 CONTINUE
      WRITE(6,3)
    3 FORMAT(//)
      DO 70 K=1,10
      DO 70 J=1,2,2
      I=2
      WRITE(6,4) ANOVAT(I,J,K), ANOVAT(I,J+1,K)
    4 FORMAT(10X,F8.5,12X,F8.5)
   70 CONTINUE
  870 DO 11 I=1,2
      DO 11 J=1,2
      CELLN(I,J)=0.0
      SUM(I,J)=0.0
      RSUM(I)=0.0
      CSUM(J)=0.0
      ROWN(I)=0.0
      COLN(I)=0.0
   11 CONTINUE
      TOTAL=0.0
      TX2=0.0
```

```
      TOTALN=0.0
      DO 8 I=1,2
      DO 8 J=1,2
      DO 8 K=1,10
      Y=ANOVAT(I,J,K)
      CELLN(I,J)=CELLN(I,J)+1.0
      SUM(I,J)=SUM(I,J)+Y
      YY=Y**2
      RSUM(I)=RSUM(I)+Y
      CSUM(J)=CSUM(J)+Y
      TOTAL=TOTAL+Y
      TX2=TX2+YY
      ROWN(I)=ROWN(I)+1.0
      COLN(J)=COLN(J)+1.0
      TOTALN=TOTALN+1.0
    8 CONTINUE
      PSSR1=((1.0/ROWN(1))*(RSUM(1)**2))+((1.0/ROWN(2))*(RSUM(2)**2))
      PSSR2=(TOTAL**2)/TOTALN
      SSR=PSSR1-PSSR2
C     SUM OF SQUARES FOR TWO ROWS
      PSSC1=((1.0/COLN(1))*(CSUM(1)**2))+((1.0/COLN(2))*(CSUM(2)**2))
      SSC=PSSC1-PSSR2
C     SUM OF SQUARES FOR TWO COLUMNS ONLY
      SSI1=((1.0/CELLN(1,1))*(SUM(1,1)**2))+((1.0/CELLN(1,2))*(SUM(1,2)*
     1*2))+((1.0/CELLN(2,1))*(SUM(2,1)**2))+((1.0/CELLN(2,2))*(SUM(2,2)*
     1*2))
C     SUM OF SQUARES FOR FOUR CELLS ONLY
      SSI=SSI1-PSSR1-PSSC1+PSSR2
      SSW=TX2-SSI1
      SST=TX2-PSSR2
      DFR=1.0
      DFC=1.0
      DFI=1.0
      DFW=36.0
      DFT=39.0
      XMSR=SSR/DFR
      XMSC=SSC/DFC
      XMSI=SSI/DFI
      XMSW=SSW/DFW
      FDR=XMSR/XMSW
      FDC=XMSC/XMSW
      FDI=XMSI/XMSW
C     THIS IS THE COMPUTATION OF F DISTRIBUTIONS FOR EQUAL CELL ANALYSIS
      XK=36/(36+FDR)
      QPROB=1.0-(((1.0-XK)**.5)*(1.0+.5*XK+.375*XK**2.0+.3125*XK**3.+.273
     14375*XK**4.0+.2460937*XK**5.0+.2255858*XK**6.0+.2094725*XK**7.0+.1
     1963804*XK**8.0+.1854703*XK**9.0+.1761967*XK**10.0+.1681877*XK**11.
     10+.1611798*XK**12.0+.1549805*XK**13.0+.1494454*XK**14.0+.1444638*X
     1K**15.0+.1399493*XK**16.0+.1358331*XK**17.0)
      IDX=QPROB*20.0+1
      GO TO 871
  412 WRITE(6,840)M,QPROB
  840 FORMAT(/,' QPROB',I3,' = ',F8.6)
```

```
871  MFR(IDX)=MFR(IDX)+1
     IF(QPROB .LT. .01)MFR(21)=MFR(21)+1
     XK=36/(36+FDC)
     QPROB=1.0-((1.0-XK)**.5)*(1.0+.5*XK+.375*XK**2.0+.3125*XK**3.+.273
    14375*XK**4.0+.2460937*XK**5.0+.2255858*XK**6.0+.2094725*XK**7.0+.1
    1963804*XK**8.0+.1854703*XK**9.0+.1761967*XK**10.0+.1681877*XK**11.
    10+.1611798*XK**12.0+.1549805*XK**13.0+.1494454*XK**14.0+.1444638*X
    1K**15.0+.1399493*XK**16.0+.1358331*XK**17.0)
     IDX=QPROB*20.0+1
     MFC(IDX)=MFC(IDX)+1
     IF(QPROB .LT. .01)MFC(21)=MFC(21)+1
     XK=36/(36+FDI)
     QPROB=1.0-((1.0-XK)**.5)*(1.0+.5*XK+.375*XK**2.0+.3125*XK**3.+.273
    14375*XK**4.0+.2460937*XK**5.0+.2255858*XK**6.0+.2094725*XK**7.0+.1
    1963804*XK**8.0+.1854703*XK**9.0+.1761967*XK**10.0+.1681877*XK**11.
    10+.1611798*XK**12.0+.1549805*XK**13.0+.1494454*XK**14.0+.1444638*X
    1K**15.0+.1399493*XK**16.0+.1358331*XK**17.0)
     IDX=QPROB*20.0+1
     MFI(IDX)=MFI(IDX)+1
     IF(QPROB .LT. .01)MFI(21)=MFI(21)+1
     GO TO 872
413  WRITE(6,17)M
 17  FORMAT(///,' ANALYSIS OF VARIANCE FOR A 2X2 DESIGN WITH EQUAL CELL
    1 FREQUENCIES','    NUMBER ',I2,//)
     WRITE(6,18)
 18  FORMAT(' SOURCE OF',11X,'SUM OF',14X,'DEGREES OF',10X,'VARIANCE',/
    1)
     WRITE(6,19)
 19  FORMAT(' VARIATION',11X,'SQUARES',13X,'FREEDOM',13X,'ESTIMATE',/)
     WRITE(6,16)
 16  FORMAT(1X,80('-'),/)
     WRITE(6,15) SSR,DFR,XMSR,FDR
 15  FORMAT(' ROWS',12X,F11.4,9X,7X,F3.0,10X,F11.4,9X,'F =   ',F8.4,/)
     WRITE(6,14) SSC,DFC,XMSC,FDC
 14  FORMAT(' COLUMNS',09X,F11.4,9X,7X,F3.0,10X,F11.4,8X,' F = ',F8.4,/
    1)
     WRITE(6,13) SSI,DFI,XMSI,FDI
 13  FORMAT(' INTERACTION',5X,F11.4,9X,7X,F3.0,10X,F11.4,8X,' F = ',F8.
    14,/)
     WRITE(6,12) SSW,DFW,XMSW
 12  FORMAT(' WITHIN',10X,F11.4,9X,7X,F3.0,10X,F11.4,/,80('-'),/)
     WRITE(6,101) SST,DFT
101  FORMAT(' TOTAL',11X,F11.4,9X,7X,F3.0,/,80('-'),/,80('-'))
872  SSW=SSW
100  CONTINUE
     READ(5,201)IX,N11,N12,N21,N22
201  FORMAT(5X,I8,2X,4I2)
     DO 300 MTOT=1,1000
     DO 301 I=1,2
     DO 301 J=1,2
     IF(I .EQ. 1 .AND. J .EQ. 1)N=N11
     IF(I .EQ. 1 .AND. J .EQ. 2)N=N12
     IF(I .EQ. 2 .AND. J .EQ. 1)N=N21
```

```
      IF(I .EQ. 2 .AND. J .EQ. 2)N=N22
      DO 301 K=1,N
C N = THE NUMBER IN A CELL AND WILL VARY WITH DISPROPORTIONALITY
      ANOVAT(I,J,K)=0.0
  301 CONTINUE
      S=2.0
      AM=10.0
      V=0.0
C     WRITE(6,302)
C 302 FORMAT(1H1)
      DO 303 I=1,2
      DO 303 J=1,2
      IF(I .EQ. 1 .AND. J .EQ. 1)N=N11
      IF(I .EQ. 1 .AND. J .EQ. 2)N=N12
      IF(I .EQ. 2 .AND. J .EQ. 1)N=N21
      IF(I .EQ. 2 .AND. J .EQ. 2)N=N22
      DO 303 L=1,N
C     CALL GAUSS(IX,S,AM,V)
C     SUBROUTINE GAUSS(IX,S,AM,V)
      A=0.0
      DO 350 K=1,12
      CALL RANDME(IX,Y)
  350 A=A+Y
      V=(A-6.0)*S+AM
C     RETURN
C     END
      ANOVAT(I,J,L)=V
  303 CONTINUE
C  UNWEIGHTED MEANS SOLUTION
  334 DO 200 I=1,2
      DO 200 J=1,2
C  I AND J WILL VARY WITH DESIGN
      CELLN(I,J)=0.0
      SUM(I,J)=0.0
      RSUM(I)=0.0
      CSUM(J)=0.0
      ROWN(I)=0.0
      COLN(J)=0.0
  200 CONTINUE
      TOTAL=0.0
      TX2=0.0
      TOTALN=0.0
      DO 208 I=1,2
      DO 208 J=1,2
      IF(I .EQ. 1 .AND. J .EQ. 1)N=N11
      IF(I .EQ. 1 .AND. J .EQ. 2)N=N12
      IF(I .EQ. 2 .AND. J .EQ. 1)N=N21
      IF(I .EQ. 2 .AND. J .EQ. 2)N=N22
      DO 208 K=1,N
      Y=ANOVAT(I,J,K)
      CELLN(I,J)=CELLN(I,J)+1.0
      SUM(I,J)=SUM(I,J)+Y
      ROWN(I)=ROWN(I)+1.0
```

```
      YY=Y**2
      CCLN(J)=CCLN(J)+1.0
      TOTAL=TOTAL+Y
      TX2=TX2+YY
      TOTALN=TOTALN+1.0
  208 CONTINUE
      GO TO 874
  415 WRITE(6,369)CELLN(1,1),CELLN(1,2),CELLN(2,1),CELLN(2,2)
  369 FORMAT(//,5X,'CELLN(1,1) =',F10.5,5X,'CELLN(1,2) =',F10.5,5X,'CELL
     1N(2,1) =',F10.5,5X,'CELLN(2,2) =',F10.5)
  874 SSW=TX2-((1.0/CELLN(1,1))*(SUM(1,1)**2))-((1.0/CELLN(1,2))*(SUM(1,
     12)**2))-((1.0/CELLN(2,1))*(SUM(2,1)**2))-((1.0/CELLN(2,2))*(SUM(2,
     12)**2))
      ACELL1=SUM(1,1)/CELLN(1,1)
      ACELL2=SUM(2,1)/CELLN(2,1)
      ACELL3=SUM(1,2)/CELLN(1,2)
      ACELL4=SUM(2,2)/CELLN(2,2)
C  MORE ACELLS ARE NEEDED FOR DIFFERENT DESIGNS
      ATOTAL=ACELL1+ACELL2+ACELL3+ACELL4
      SSRUNW=((.5*((ACELL1+ACELL3)**2))+(.5*((ACELL2+ACELL4)**2)))-(ATOT
     1AL**2)/4.0
      SSCUNW=((.5*((ACELL1+ACELL2)**2))+(.5*((ACELL3+ACELL4)**2)))-(ATOT
     1AL**2)/4.0
C  THE 1/2=.5 WILL VARY WITH DESIGNS AS WILL THE NO. OF ACELLS AND THE 4
      SSTUNW=((ACELL1**2)+(ACELL2**2)+(ACELL3**2)+(ACELL4**2))-(ATOTAL**
     12)/4.0
      SSIUNW=SSTUNW-SSRUNW-SSCUNW
      XMSRUN=SSRUNW/1.0
      XMSCUN=SSCUNW/1.0
      XMSIUN=SSIUNW/1.0
C   1.0 CAN ONLY BE USED ON A 2X2 DESIGN HERE
      ADSSW=SSW*(.25*(1.0/N11+1.0/N12+1.0/N21+1.0/N22))
      GO TO 875
  416 WRITE(6,370)ADSSW,SSW
  370 FORMAT(//,5X,'ADSSW =',F11.5,5X,'SSW =',F13.5)
  875 XMSWUN=ADSSW/36.0
C  THE .25 AND 36.0 MUST BE CHANGED FOR ANYTHING BESIDES A 2X2
      FR=XMSRUN/XMSWUN
      FC=XMSCUN/XMSWUN
      FI=XMSIUN/XMSWUN
C     COMPUTATION OF F DIST. FOR UNWEIGHTED MEANS ANALYSIS
      XK=36/(36+FR)
      QPROB=1.0-((1.0-XK)**.5)*(1.0+.5*XK+.375*XK**2.0+.3125*XK**3.+.273
     14375*XK**4.0+.2460937*XK**5.0+.2255858*XK**6.0+.2094725*XK**7.0+.1
     1963804*XK**8.0+.1854703*XK**9.0+.1761967*XK**10.0+.1681877*XK**11.
     10+.1611798*XK**12.0+.1549805*XK**13.0+.1494454*XK**14.0+.1444638*X
     1K**15.0+.1399493*XK**16.0+.1358331*XK**17.0)
      IDX=QPROB*20.0+1
      MDFR(IDX)=MDFR(IDX)+1
      IF(QPROB .LT. .01)MDFR(21)=MDFR(21)+1
      XK=36/(36+FC)
      QPROB=1.0-((1.0-XK)**.5)*(1.0+.5*XK+.375*XK**2.0+.3125*XK**3.+.273
```

```
14375*XK**4.0+.2460937*XK**5.0+.2255858*XK**6.0+.2094725*XK**7.0+.1
1963804*XK**8.0+.1854703*XK**9.0+.1761967*XK**10.0+.1681877*XK**11.
10+.1611798*XK**12.0+.1549805*XK**13.0+.1494454*XK**14.0+.1444638*X
1K**15.0+.1399493*XK**16.0+.1358331*XK**17.0)
 IDX=QPROB*20.0+1
 MDFC(IDX)=MDFC(IDX)+1
 IF(QPROB .LT. .01)MDFC(21)=MDFC(21)+1
 XK=36/(36+FI)
 QPROB=1.0-((1.0-XK)**.5)*(1.0+.5*XK+.375*XK**2.0+.3125*XK**3.+.273
14375*XK**4.0+.2460937*XK**5.0+.2255858*XK**6.0+.2094725*XK**7.0+.1
1963804*XK**8.0+.1854703*XK**9.0+.1761967*XK**10.0+.1681877*XK**11.
10+.1611798*XK**12.0+.1549805*XK**13.0+.1494454*XK**14.0+.1444638*X
1K**15.0+.1399493*XK**16.0+.1358331*XK**17.0)
 IDX=QPROB*20.0+1
 MDFI(IDX)=MDFI(IDX)+1
 IF(QPROB .LT. .01)MDFI(21)=MDFI(21)+1
C     METHOD OF EXPECTED FREQUENCIES
 EVCEL1=(ROWN(1)*COLN(1))/TOTALN
 EVCEL2=(ROWN(2)*COLN(1))/TOTALN
 EVCEL3=(ROWN(1)*COLN(2))/TOTALN
 EVCEL4=(ROWN(2)*COLN(2))/TOTALN
 TCEL1=EVCEL1*ACELL1
 TCEL2=EVCEL2*ACELL2
 TCEL3=EVCEL3*ACELL3
 TCEL4=EVCEL4*ACELL4
 TALCEL=TCEL1+TCEL2+TCEL3+TCEL4
 TEV=EVCEL1+EVCEL2+EVCEL3+EVCEL4
 SSREF=((TCEL1+TCEL3)**2.0)/(EVCEL1+EVCEL3)+((TCEL2+TCEL4)**2.0)/(E
1VCEL2+EVCEL4)-(TALCEL**2.0)/TEV
 SSCEF=((TCEL1+TCEL2)**2.0)/(EVCEL1+EVCEL2)+((TCEL3+TCEL4)**2.0)/(E
1VCEL3+EVCEL4)-(TALCEL**2.0)/TEV
 SSIEF=(TCEL1**2.0)/EVCEL1+(TCEL2**2.0)/EVCEL2+(TCEL3**2.0)/EVCEL3+
1(TCEL4**2.0)/EVCEL4-(TALCEL**2.0)/TEV-SSREF-SSCEF
 DFR=1.0
 DFC=1.0
 DFI=1.0
 DFW=36.0
 XMSREF=SSREF/DFR
 XMSCEF=SSCEF/DFC
 XMSIEF=SSIEF/DFI
 XMSWEF=SSW/DFW
 FREF=XMSREF/XMSWEF
 FCEF=XMSCEF/XMSWEF
 FIEF=XMSIEF/XMSWEF
C     COMPUTATION OF F DIST. FOR METHOD OF EXPECTED FREQUENCIES
 XK=36/(36+FREF)
 QPROB=1.0-((1.0-XK)**.5)*(1.0+.5*XK+.375*XK**2.0+.3125*XK**3.+.273
14375*XK**4.0+.2460937*XK**5.0+.2255858*XK**6.0+.2094725*XK**7.0+.1
1963804*XK**8.0+.1854703*XK**9.0+.1761967*XK**10.0+.1681877*XK**11.
10+.1611798*XK**12.0+.1549805*XK**13.0+.1494454*XK**14.0+.1444638*X
1K**15.0+.1399493*XK**16.0+.1358331*XK**17.0)
 IDX=QPROB*20.0+1
```

```
      MFREF(IDX)=MFREF(IDX)+1
      IF(QPROB .LT. .01)MFREF(21)=MFREF(21)+1
      XK=36/(36+FCEF)
      QPROB=1.0-((1.0-XK)**.5)*(1.0+.5*XK+.375*XK**2.0+.3125*XK**3.+.273
     14375*XK**4.0+.2460937*XK**5.0+.2255858*XK**6.0+.2094725*XK**7.0+.1
     1963804*XK**8.0+.1854703*XK**9.0+.1761967*XK**10.0+.1681877*XK**11.
     10+.1611798*XK**12.0+.1549805*XK**13.0+.1494454*XK**14.0+.1444638*X
     1K**15.0+.1399493*XK**16.0+.1358331*XK**17.0)
      IDX=QPROB*20.0+1
      MFCEF(IDX)=MFCEF(IDX)+1
      IF(QPROB .LT. .01)MFCEF(21)=MFCEF(21)+1
      XK=36/(36+FIEF)
      QPROB=1.0-((1.0-XK)**.5)*(1.0+.5*XK+.375*XK**2.0+.3125*XK**3.+.273
     14375*XK**4.0+.2460937*XK**5.0+.2255858*XK**6.0+.2094725*XK**7.0+.1
     1963804*XK**8.0+.1854703*XK**9.0+.1761967*XK**10.0+.1681877*XK**11.
     10+.1611798*XK**12.0+.1549805*XK**13.0+.1494454*XK**14.0+.1444638*X
     1K**15.0+.1399493*XK**16.0+.1358331*XK**17.0)
      IDX=QPROB*20.0+1
      MFIEF(IDX)=MFIEF(IDX)+1
      IF(QPROB .LT. .01)MFIEF(21)=MFIEF(21)+1
      GO TO 877
  418 WRITE(6,400)N11,N12,N21,N22
  400 FORMAT(////,5X,'METHOD OF EXPECTED FREQUENCIES ON A 2X2 DESIGN WIT
     1H CELL SIZES OF',4(2X,I3),/)
      WRITE(6,257)
      WRITE(6,258)
      WRITE(6,257)
      WRITE(6,401)SSREF,XMSREF,FREF
  401 FORMAT(/,9X,'ROWS',21X,'1',7X,F7.4,4X,F7.4,7X,F7.4,/)
      WRITE(6,402)SSCEF,XMSCEF,FCEF
  402 FORMAT(9X,'COLUMNS',18X,'1',7X,F7.4,4X,F7.4,7X,F7.4,/)
      WRITE(6,403)SSIEF,XMSIEF,FIEF
  403 FORMAT(9X,'INTERACTION',14X,'1',5X,F9.4,2X,F9.4,7X,F7.4,/)
C     CHANGE FOR DIFFERENT THAN 2X2
      WRITE(6,404)SSW,XMSWEF
  404 FORMAT(9X,'WITHIN',18X,'36',5X,F10.4,1X,F9.4,/)
      WRITE(6,257)
  257 FORMAT(5X,75('-'))
  258 FORMAT(5X,'SOURCE OF VARIATION',10X,'DF',10X,'SS',10X,'MS',10X,'F'
     1)
  877 ICODEW=1
      REWIND 1
      NSUBJ=40
      NVREAD=3
      NVTOT=6
      XN=NSUBJ
      IF(NVTOT.EQ.0)NVTOT=NVREAD
      NS=1
      DO 456 JKB=1,14
      MFA(JKB)=0
  456 CONTINUE
      DO 458 I=1,14
      MFB(I)=0
```

```
  458 CONTINUE
      DO 468 I=1,NVTOT
      XM(I)=0.0
      DO 468 J=1,NVTOT
  468 COR(I,J)=0.0
      DO 480 I=1,2
      DO 480 J=1,2
      IF(I .EQ. 1 .AND. J .EQ. 1)N=N11
      IF(I .EQ. 1 .AND. J .EQ. 2)N=N12
      IF(I .EQ. 2 .AND. J .EQ. 1)N=N21
      IF(I .EQ. 2 .AND. J .EQ. 2)N=N22
      DO 480 K=1,N
      DO 484 NV=1,NVTOT
  484 X(NV)=0.0
      X(1)=ANOVAT(I,J,K)
      X(2)=I
      X(3)=J
C**** INSERT GENERATING STATEMENTS FOR X( ).
      L=X(2)
      M=X(3)
      X(2)=0.0
      X(3)=0.0
      X(L+1)=1.0
      X(M+3)=1.0
      X(2)=X(2)-X(3)
      X(4)=X(4)-X(5)
      X(6)=X(2)*X(4)
      GO TO (491,492,493), ICODEW
  491 CONTINUE
C**** INSERT  GENERATINT STATEMENTS FOR PROBLEM  1
      GO TO 510
  492 CONTINUE
C**** INSERT  GENERATINT STATEMENTS FOR PROBLEM  2
      GO TO 510
  493 CONTINUE
C**** INSERT  GENERATINT STATEMENTS FOR PROBLEM  3
  510 CONTINUE
C**** INSERT GENERATING STATEMENTS ABOVE THIS CARD.
      NS=NS+1
      DO 480 JA=1,NVTOT
      XM(JA)=XM(JA)+X(JA)
      DO 480 JB=1,NVTOT
  480 COR(JA,JB)=COR(JA,JB)+X(JA)*X(JB)
      DO 483 J=1,NVTOT
      XM(J)=XM(J)/XN
  483 XS(J)=DSQRT((COR(J,J)-XN*XM(J)*XM(J))/(XN-1.0))
      DO 488 JA=1,NVTOT
      DO 488 JB=1,NVTOT
      TEMP=(COR(JA,JB)-XN*XM(JA)*XM(JB))/(XN-1.0)
      COR(JA,JB)=0.0
      IF(XS(JA)*XS(JB).EQ.0.0)GO TO 488
      COR(JA,JB)=TEMP/(XS(JA)*XS(JB))
```

```
  488 CONTINUE
      NLN=60
      IPRINT=10
C**** PRINT MEANS AND STANDARD DEVIATIONS
C**** COMPUTE REGRESSION MODELS.
      MOD=0
      DO 450 IKB=1,6
      MF(1)=01
      IF(IKB .EQ. 1)GO TO 531
      IF(IKB .EQ. 2)GO TO 532
      IF(IKB .EQ. 3)GO TO 533
      IF(IKB .EQ. 4)GO TO 534
      IF(IKB .EQ. 5)GO TO 535
      IF(IKB .EQ. 6)GO TO 536
C     THE ABOVE SET-UP IS ONLY FOR A 2X2 DESIGN
  531 MFA(1)=02
      MFB(1)=02
      MFA(2)=04
      MFB(2)=04
      MFA(3)=06
      MFB(3)=06
      GO TO 529
  532 MFA(1)=02
      MFB(1)=02
      MFA(2)=04
      MFB(2)=04
      GO TO 529
  533 MFA(1)=04
      MFB(1)=04
      MFA(2)=06
      MFB(2)=06
      GO TO 529
  534 MFA(1)=02
      MFB(1)=02
      MFA(2)=06
      MFB(2)=06
      GO TO 529
  535 MFA(1)=02
      MFB(1)=02
      GO TO 529
  536 MFA(1)=04
      MFB(1)=04
      GO TO 529
  529 NC=1
      LY=MF(1)
      DO 530 I=1,14
      IA=MFA(I)
      IB=MFB(I)
      IF(IA.EQ.0)GO TO 530
      DO 545 K=IA,IB
      NC=NC+1
  545 MF(NC)=K
```

```
530 CONTINUE
    MOD=MOD+1
    MODX=MOD+1
    RSQ(1)=0.0
    DO 540 I=2,NC
    IA=MF(I)
    DO 540 J=1,NC
    JA=MF(J)
540 RW(I,J)=COR(IA,JA)
    DO 585 I=2,NC
    CNT(I)=1.0
    TEMP=RW(I,I)
    IF(DABS(TEMP).GT.0.0000001)GO TO 550
    CNT(I)=0.0
    GO TO 585
550 DO 560 J=1,NC
560 RW(I,J)=RW(I,J)/TEMP
    DO 580 IA=2,NC
    TEMP=RW(IA,I)
    IF(I.EQ.IA)TEMP=0.0
    DO 570 J=1,NC
570 RW(IA,J)=RW(IA,J)-TEMP*RW(I,J)
580 CONTINUE
585 CONTINUE
    RSQ(MODX)=0.0
    REGCO=XM(LY)
    DO 590 I=2,NC
    IA=MF(I)
    RW(I,1)=RW(I,1)*CNT(I)
    RSQ(MODX)=RSQ(MODX)+RW(I,1)*COR(IA,LY)
    IF(XS(IA) .EQ. 0.0D0) XS(IA)=.1D-15
    BW(IA)=RW(I,1)*XS(LY)/XS(IA)
590 REGCO=REGCO-BW(IA)*XM(IA)
    DO 599 ILL=1,6
    MFA(ILL)=0.0
    MFB(ILL)=0.0
599 CONTINUE
450 CONTINUE
    ICODEW=ICODEW+1
500 ERRTM=(1.0-RSQ(2))/36.0
    F1=((RSQ(2)-RSQ(3))/1.0)/ERRTM
    F2=((RSQ(2)-RSQ(4))/1.0)/ERRTM
    F3=((RSQ(2)-RSQ(5))/1.0)/ERRTM
    F4=((RSQ(3)-RSQ(6))/1.0)/ERRTM
    F5=((RSQ(3)-RSQ(7))/1.0)/ERRTM
C   COMPUTATION OF F DIST. FOR METHOD 1
    XK=36/(36+F2)
    QPROB=1.0-((1.0-XK)**.5)*(1.0+.5*XK+.375*XK**2.0+.3125*XK**3.+.273
   14375*XK**4.0+.2460937*XK**5.0+.2255858*XK**6.0+.2094725*XK**7.0+.1
   19638O4*XK**8.0+.1854703*XK**9.0+.1761967*XK**10.0+.1681877*XK**11.
   10+.1611798*XK**12.0+.1549805*XK**13.0+.1494454*XK**14.0+.1444638*X
   1K**15.0+.1399493*XK**16.0+.1358331*XK**17.0)
    IDX=QPROB*20.0+1
```

```
      MF2M1(IDX)=MF2M1(IDX)+1
      IF(QPROB .LT. .01)MF2M1(21)=MF2M1(21)+1
      XK=36/(36+F3)
      QPROB=1.0-((1.0-XK)**.5)*(1.0+.5*XK+.375*XK**2.0+.3125*XK**3.+.273
     14375*XK**4.0+.2460937*XK**5.0+.2255858*XK**6.0+.2094725*XK**7.0+.1
     1963804*XK**8.0+.1854703*XK**9.0+.1761967*XK**10.0+.1681877*XK**11.
     10+.1611798*XK**12.0+.1549805*XK**13.0+.1494454*XK**14.0+.1444638*X
     1K**15.0+.1399493*XK**16.0+.1358331*XK**17.0)
      IDX=QPROB*20.0+1
      MF3M1(IDX)=MF3M1(IDX)+1
      IF(QPROB .LT. .01)MF3M1(21)=MF3M1(21)+1
      XK=36/(36+F1)
      QPROB=1.0-((1.0-XK)**.5)*(1.0+.5*XK+.375*XK**2.0+.3125*XK**3.+.273
     14375*XK**4.0+.2460937*XK**5.0+.2255858*XK**6.0+.2094725*XK**7.0+.1
     1963804*XK**8.0+.1854703*XK**9.0+.1761967*XK**10.0+.1681877*XK**11.
     10+.1611798*XK**12.0+.1549805*XK**13.0+.1494454*XK**14.0+.1444638*X
     1K**15.0+.1399493*XK**16.0+.1358331*XK**17.0)
      IDX=QPROB*20.0+1
      MF1M1(IDX)=MF1M1(IDX)+1
      IF(QPROB .LT. .01)MF1M1(21)=MF1M1(21)+1
C     COMPUTATION OF F DIST. FOR  METHOD 2
      XK=36/(36+F5)
      QPROB=1.0-((1.0-XK)**.5)*(1.0+.5*XK+.375*XK**2.0+.3125*XK**3.+.273
     14375*XK**4.0+.2460937*XK**5.0+.2255858*XK**6.0+.2094725*XK**7.0+.1
     1963804*XK**8.0+.1854703*XK**9.0+.1761967*XK**10.0+.1681877*XK**11.
     10+.1611798*XK**12.0+.1549805*XK**13.0+.1494454*XK**14.0+.1444638*X
     1K**15.0+.1399493*XK**16.0+.1358331*XK**17.0)
      IDX=QPROB*20.0+1
      MF5M2(IDX)=MF5M2(IDX)+1
      IF(QPROB .LT. .01)MF5M2(21)=MF5M2(21)+1
      XK=36/(36+F4)
      QPROB=1.0-((1.0-XK)**.5)*(1.0+.5*XK+.375*XK**2.0+.3125*XK**3.+.273
     14375*XK**4.0+.2460937*XK**5.0+.2255858*XK**6.0+.2094725*XK**7.0+.1
     1963804*XK**8.0+.1854703*XK**9.0+.1761967*XK**10.0+.1681877*XK**11.
     10+.1611798*XK**12.0+.1549805*XK**13.0+.1494454*XK**14.0+.1444638*X
     1K**15.0+.1399493*XK**16.0+.1358331*XK**17.0)
      IDX=QPROB*20.0+1
      MF4M2(IDX)=MF4M2(IDX)+1
      IF(QPROB .LT. .01)MF4M2(21)=MF4M2(21)+1
      XK=36/(36+F1)
      QPROB=1.0-((1.0-XK)**.5)*(1.0+.5*XK+.375*XK**2.0+.3125*XK**3.+.273
     14375*XK**4.0+.2460937*XK**5.0+.2255858*XK**6.0+.2094725*XK**7.0+.1
     1963804*XK**8.0+.1854703*XK**9.0+.1761967*XK**10.0+.1681877*XK**11.
     10+.1611798*XK**12.0+.1549805*XK**13.0+.1494454*XK**14.0+.1444638*X
     1K**15.0+.1399493*XK**16.0+.1358331*XK**17.0)
      IDX=QPROB*20.0+1
      MF1M2(IDX)=MF1M2(IDX)+1
      IF(QPROB .LT. .01)MF1M2(21)=MF1M2(21)+1
      GO TO 878
  419 WRITE(6,900)
  900 FORMAT(/)
      WRITE(6,901)
  901 FORMAT(5X,23('-'))
```

```
      WRITE(6,902)
  902 FORMAT(5X,'SOURCE',10X,'F')
      WRITE(6,901)
      WRITE(6,903)
  903 FORMAT(/,11X,'METHOD 1')
      WRITE(6,901)
      WRITE(6,904)F2
  904 FORMAT(5X,'ROWS',12X,F11.8)
      WRITE(6,905)F3
  905 FORMAT(5X,'COLUMNS',9X,F11.8)
      WRITE(6,906)F1
  906 FORMAT(5X,'INTERACTION',5X,F11.8)
      WRITE(6,901)
      WRITE(6,907)
  907 FORMAT(//,11X,'METHOD 2')
      WRITE(6,901)
      WRITE(6,904)F5
      WRITE(6,905)F4
      WRITE(6,906)F1
      WRITE(6,908)
  908 FORMAT(5X,23('-'),//)
      WRITE(6,930)RSQ(1),RSQ(2),RSQ(3),RSQ(4),RSQ(5),RSQ(6),RSQ(7)
  930 FORMAT(//,' RSQ 1,2,3,4,5,6,7',7(4X,F6.4))
  878 SSW=SSW
  300 CONTINUE
      WRITE(6,800)
  800 FORMAT(1H1)
      WRITE(6,801)
  801 FORMAT(1X,'FREQUENCY DISTRIBUTIONS FOR EQUAL CELL METHOD AND FOUR
     1 METHODS OF HANDLING DISPROPORTIONALITY',//)
      WRITE(6,850)
  850 FORMAT(1X,'ROW F DISTRIBUTIONS',/)
      WRITE(6,802)
  802 FORMAT(1X,'1.00',2X,'.95',2X,'.90',2X,'.85',2X,'.80',2X,'.75',2X,'
     1.70',2X,'.65',2X,'.60',2X,'.55',2X,'.50',2X,'.45',2X,'.40',2X,'.35
     1',2X,'.30',2X,'.25',2X,'.20',2X,'.15',2X,'.10',2X,'.05',2X,'.01',/
     1)
      WRITE(6,803)
  803 FORMAT(1X,'EQUAL CELL ANALYSIS OF VARIANCE F DISTRIBUTION',/)
      WRITE(6,804)MFR(20),MFR(19),MFR(18),MFR(17),MFR(16),MFR(15),MFR(14
     1),MFR(13),MFR(12),MFR(11),MFR(10),MFR(9),MFR(8),MFR(7),MFR(6),MFR(
     15),MFR(4),MFR(3),MFR(2),MFR(1),MFR(21)
  804 FORMAT(2X,21(F4.0,1X),/)
      WRITE(6,820)
  820 FORMAT(1X,'F DISTRIBUTION FOR UNWEIGHTED MEANS ANALYSIS',/)
      WRITE(6,805)MDFR(20),MDFR(19),MDFR(18),MDFR(17),MDFR(16),MDFR(15),
     1MDFR(14),MDFR(13),MDFR(12),MDFR(11),MDFR(10),MDFR(9),MDFR(8),MDFR(
     17),MDFR(6),MDFR(5),MDFR(4),MDFR(3),MDFR(2),MDFR(1),MDFR(21)
  805 FORMAT(2X,21(F4.0,1X),/)
      WRITE(6,821)
  821 FORMAT(1X,'F DISTRIBUTION FOR METHOD OF EXPECTED FREQUENCIES',/)
      WRITE(6,806)MFREF(20),MFREF(19),MFREF(18),MFREF(17),MFREF(16),MFRE
```

```
     1F(15),MFREF(14),MFREF(13),MFREF(12),MFREF(11),MFREF(10),MFREF(9),M
     1FREF(8),MFREF(7),MFREF(6),MFREF(5),MFREF(4),MFREF(3),MFREF(2),MFRE
     1F(1),MFREF(21)
 806 FORMAT(2X,21(F4.0,1X),/)
     WRITE(6,822)
 822 FORMAT(1X,'F DISTRIBUTION FOR METHOD 1',/)
     WRITE(6,807)MF2M1(20),MF2M1(19),MF2M1(18),MF2M1(17),MF2M1(16),MF2M
     11(15),MF2M1(14),MF2M1(13),MF2M1(12),MF2M1(11),MF2M1(10),MF2M1(9),M
     1F2M1(8),MF2M1(7),MF2M1(6),MF2M1(5),MF2M1(4),MF2M1(3),MF2M1(2),MF2M
     11(1),MF2M1(21)
 807 FORMAT(2X,21(F4.0,1X),/)
 823 FORMAT(1X,'F DISTRIBUTION FOR METHOD 2',/)
     WRITE(6,823)
     WRITE(6,808)MF5M2(20),MF5M2(19),MF5M2(18),MF5M2(17),MF5M2(16),MF5M
     12(15),MF5M2(14),MF5M2(13),MF5M2(12),MF5M2(11),MF5M2(10),MF5M2(9),M
     1F5M2(8),MF5M2(7),MF5M2(6),MF5M2(5),MF5M2(4),MF5M2(3),MF5M2(2),MF5M
     12(1),MF5M2(21)
 808 FORMAT(2X,21(F4.0,1X),/)
     WRITE(6,824)
 824 FORMAT(//,1X,'COLUMN F DISTRIBUTIONS',/)
     WRITE(6,825)
 825 FORMAT(1X,'EQUAL CELL ANALYSIS OF VARIANCE F DISTRIBUTION',/)
     WRITE(6,809)MFC(20),MFC(19),MFC(18),MFC(17),MFC(16),MFC(15),MFC(14
     1),MFC(13),MFC(12),MFC(11),MFC(10),MFC(9),MFC(8),MFC(7),MFC(6),MFC(
     15),MFC(4),MFC(3),MFC(2),MFC(1),MFC(21)
 809 FORMAT(2X,21(F4.0,1X),/)
     WRITE(6,826)
 826 FORMAT(1X,'F DISTRIBUTION FOR UNWEIGHTED MEANS ANALYSIS',/)
     WRITE(6,810)MDFC(20),MDFC(19),MDFC(18),MDFC(17),MDFC(16),MDFC(15),
     1MDFC(14),MDFC(13),MDFC(12),MDFC(11),MDFC(10),MDFC(9),MDFC(8),MDFC(
     17),MDFC(6),MDFC(5),MDFC(4),MDFC(3),MDFC(2),MDFC(1),MDFC(21)
 810 FORMAT(2X,21(F4.0,1X),/)
     WRITE(6,827)
 827 FORMAT(1X,'F DISTRIBUTION FOR METHOD OF EXPECTED FREQUENCIES',/)
     WRITE(6,811)MFCEF(20),MFCEF(19),MFCEF(18),MFCEF(17),MFCEF(16),MFCE
     1F(15),MFCEF(14),MFCEF(13),MFCEF(12),MFCEF(11),MFCEF(10),MFCEF(9),M
     1FCEF(8),MFCEF(7),MFCEF(6),MFCEF(5),MFCEF(4),MFCEF(3),MFCEF(2),MFCE
     1F(1),MFCEF(21)
 811 FORMAT(2X,21(F4.0,1X),/)
     WRITE(6,828)
 828 FORMAT(1X,'F DISTRIBUTION FOR METHOD 1',/)
     WRITE(6,812)MF3M1(20),MF3M1(19),MF3M1(18),MF3M1(17),MF3M1(16),MF3M
     11(15),MF3M1(14),MF3M1(13),MF3M1(12),MF3M1(11),MF3M1(10),MF3M1(9),M
     1F3M1(8),MF3M1(7),MF3M1(6),MF3M1(5),MF3M1(4),MF3M1(3),MF3M1(2),MF3M
     11(1),MF3M1(21)
 812 FORMAT(2X,21(F4.0,1X),/)
     WRITE(6,829)
 829 FORMAT(1X,'F DISTRIBUTION FOR METHOD 2',/)
     WRITE(6,813)MF4M2(20),MF4M2(19),MF4M2(18),MF4M2(17),MF4M2(16),MF4M
     12(15),MF4M2(14),MF4M2(13),MF4M2(12),MF4M2(11),MF4M2(10),MF4M2(9),M
     1F4M2(8),MF4M2(7),MF4M2(6),MF4M2(5),MF4M2(4),MF4M2(3),MF4M2(2),MF4M
     12(1),MF4M2(21)
 813 FORMAT(2X,21(F4.0,1X),/)
```

```
      WRITE(6,830)
  830 FORMAT(//,1X,'INTERACTION F DISTRIBUTIONS',/)
      WRITE(6,831)
  831 FORMAT(1X,'EQUAL CELL ANALYSIS OF VARIANCE F DISTRIBUTION',/)
      WRITE(6,814)MFI(20),MFI(19),MFI(18),MFI(17),MFI(16),MFI(15),MFI(14
     1),MFI(13),MFI(12),MFI(11),MFI(10),MFI(9),MFI(8),MFI(7),MFI(6),MFI(
     15),MFI(4),MFI(3),MFI(2),MFI(1),MFI(21)
  814 FORMAT(2X,21(F4.0,1X),/)
      WRITE(6,832)
  832 FORMAT(1X,'F DISTRIBUTION FOR UNWEIGHTED MEANS ANALYSIS',/)
      WRITE(6,815)MDFI(20),MDFI(19),MDFI(18),MDFI(17),MDFI(16),MDFI(15),
     1MDFI(14),MDFI(13),MDFI(12),MDFI(11),MDFI(10),MDFI(9),MDFI(8),MDFI(
     17),MDFI(6),MDFI(5),MDFI(4),MDFI(3),MDFI(2),MDFI(1),MDFI(21)
  815 FORMAT(2X,21(F4.0,1X),/)
      WRITE(6,833)
  833 FORMAT(1X,'F DISTRIBUTION FOR METHOD OF EXPECTED FREQUENCIES',/)
      WRITE(6,816)MFIEF(20),MFIEF(19),MFIEF(18),MFIEF(17),MFIEF(16),MFIE
     1F(15),MFIEF(14),MFIEF(13),MFIEF(12),MFIEF(11),MFIEF(10),MFIEF(9),M
     1FIEF(8),MFIEF(7),MFIEF(6),MFIEF(5),MFIEF(4),MFIEF(3),MFIEF(2),MFIE
     1F(1),MFIEF(21)
  816 FORMAT(2X,21(F4.0,1X),/)
      WRITE(6,834)
  834 FORMAT(1X,'F DISTRIBUTION FOR METHOD 1',/)
      WRITE(6,817)MF1M1(20),MF1M1(19),MF1M1(18),MF1M1(17),MF1M1(16),MF1M
     11(15),MF1M1(14),MF1M1(13),MF1M1(12),MF1M1(11),MF1M1(10),MF1M1(9),M
     1F1M1(8),MF1M1(7),MF1M1(6),MF1M1(5),MF1M1(4),MF1M1(3),MF1M1(2),MF1M
     11(1),MF1M1(21)
  817 FORMAT(2X,21(F4.0,1X),/)
      WRITE(6,836)
  836 FORMAT(1X,'F DISTRIBUTION FOR METHOD 2',/)
      WRITE(6,818)MF1M2(20),MF1M2(19),MF1M2(18),MF1M2(17),MF1M2(16),MF1M
     12(15),MF1M2(14),MF1M2(13),MF1M2(12),MF1M2(11),MF1M2(10),MF1M2(9),M
     1F1M2(8),MF1M2(7),MF1M2(6),MF1M2(5),MF1M2(4),MF1M2(3),MF1M2(2),MF1M
     12(1),MF1M2(21)
  818 FORMAT(2X,21(F4.0,1X),/)
      READ(5,999)N1000
  999 FORMAT(I5)
      XN100=N1000
      DO 985 N=1,19
      MFR(N+1)=MFR(N+1)+MFR(N)
      MFC(N+1)=MFC(N+1)+MFC(N)
      MFI(N+1)=MFI(N+1)+MFI(N)
      MDFR(N+1)=MDFR(N+1)+MDFR(N)
      MDFC(N+1)=MDFC(N+1)+MDFC(N)
      MDFI(N+1)=MDFI(N+1)+MDFI(N)
      MFREF(N+1)=MFREF(N+1)+MFREF(N)
      MFCEF(N+1)=MFCEF(N+1)+MFCEF(N)
      MFIEF(N+1)=MFIEF(N+1)+MFIEF(N)
      MF2M1(N+1)=MF2M1(N+1)+MF2M1(N)
      MF3M1(N+1)=MF3M1(N+1)+MF3M1(N)
      MF1M1(N+1)=MF1M1(N+1)+MF1M1(N)
      MF5M2(N+1)=MF5M2(N+1)+MF5M2(N)
      MF4M2(N+1)=MF4M2(N+1)+MF4M2(N)
```

```
      MF1M2(N+1)=MF1M2(N+1)+MF1M2(N)
  985 CONTINUE
      DO 986 IX=1,21
      XRUWKS(IX)=MFR(IX)/XN100-MDFR(IX)/XN100
      XREFKS(IX)=MFR(IX)/XN100-MFREF(IX)/XN100
      XRM1KS(IX)=MFR(IX)/XN100-MF2M1(IX)/XN100
      XRM2KS(IX)=MFR(IX)/XN100-MF5M2(IX)/XN100
      XCUWKS(IX)=MFC(IX)/XN100-MDFC(IX)/XN100
      XIUWKS(IX)=MFI(IX)/XN100-MDFI(IX)/XN100
      XCEFKS(IX)=MFC(IX)/XN100-MFCEF(IX)/XN100
      XIEFKS(IX)=MFI(IX)/XN100-MFIEF(IX)/XN100
      XCM1KS(IX)=MFC(IX)/XN100-MF3M1(IX)/XN100
      XIM1KS(IX)=MFI(IX)/XN100-MF1M1(IX)/XN100
      XCM2KS(IX)=MFC(IX)/XN100-MF4M2(IX)/XN100
      XIM2KS(IX)=MFI(IX)/XN100-MF1M2(IX)/XN100
  986 CONTINUE
      D1=0
      D2=0
      D3=0
      D4=0
      D5=0
      D6=0
      D7=0
      D8=0
      D9=0
      D10=0
      D11=0
      D12=0
      DO 989 N=1,21
      DT=DABS(XRUWKS(N))
      IF(DT .GT. D1) D1=DT
      DT=DABS(XREFKS(N))
      IF(DT .GT. D2) D2=DT
      DT=DABS(XRM1KS(N))
      IF(DT .GT. D3) D3=DT
      DT=DABS(XRM2KS(N))
      IF(DT .GT. D4) D4=DT
      DT=DABS(XCUWKS(N))
      IF(DT .GT. D5) D5=DT
      DT=DABS(XCEFKS(N))
      IF(DT .GT. D6) D6=DT
      DT=DABS(XCM1KS(N))
      IF(DT .GT. D7) D7=DT
      DT=DABS(XCM2KS(N))
      IF(DT .GT. D8) D8=DT
      DT=DABS(XIUWKS(N))
      IF(DT .GT. D9) D9=DT
      DT=DABS(XIEFKS(N))
      IF(DT .GT. D10) D10=DT
      DT=DABS(XIM1KS(N))
      IF(DT .GT. D11) D11=DT
      DT=DABS(XIM2KS(N))
      IF(DT .GT. D12) D12=DT
```

```
989 CONTINUE
    WRITE(6,990) D1,D2
990 FORMAT(//,2X,'D VALUE FOR ROW UN. WT. = ',F5.3,10X,'D VALUE FOR RO
   1W EX. FREQ. = ',F5.3,/)
    WRITE(6,991) D3,D4
991 FORMAT(/,2X,'D VALUE FOR ROW METHOD 1 = ',F5.3,10X,'D VALUE FOR RO
   1W METHOD 2 = ',F5.3,/)
    WRITE(6,992) D5,D6
992 FORMAT(/,2X,'D VALUE FOR COL. UN. WT. = ',F5.3,10X,'D VALUE FOR CO
   1L. EX. FREQ. = ',F5.3,/)
    WRITE(6,993) D7,D8
993 FORMAT(/,2X,'D VALUE FOR COL. METHOD 1 = ',F5.3,10X,'D VALUE FOR C
   1OL. METHOD 2 = ',F5.3,/)
    WRITE(6,994) D9,D10
994 FORMAT(/,2X,'D VALUE FOR INTER. UN. WT. = ',F5.3,10X,'D VALUE FOR
   1INTER. EX. FREQ. = ',F5.3,/)
    WRITE(6,995) D11,D12
995 FORMAT(/,2X,'D VALUE FOR INTER. METHOD 1 = ',F5.3,10X,'D VALUE FOR
   1 INTER. METHOD 2 = ',F5.3,/)
    WRITE(6,996)
996 FORMAT(//,2X,'TABLE VALUE OF D FOR KS TEST, N = 1000 AT .05 LEVEL
   1= .043')
    XNROW1=N11+N12
    XNROW2=N21+N22
    XNCOL1=N11+N21
    XNCOL2=N12+N22
    XNTOTL=XNCOL1+XNCOL2
    XN1EX=(XNROW1*XNCOL1)/XNTOTL
    XN2EX=(XNROW1*XNCOL2)/XNTOTL
    XN3EX=(XNROW2*XNCOL1)/XNTOTL
    XN4EX=(XNROW2*XNCOL2)/XNTOTL
    XN11=N11
    XN12=N12
    XN21=N21
    XN22=N22
    XCHI=(XN11-XN1EX)**2/XN1EX+(XN12-XN2EX)**2/XN2EX+(XN21-XN3EX)**2/X
   1N3EX+(XN22-XN4EX)**2/XN4EX
    WRITE(6,997) XCHI
997 FORMAT(//,2X,'CHISQ FOR 2X2 .30 PROB = 1.07',/,2X,'CHISQ FOR 2X2 .
   120 PROB = 1.64',/,2X,'CHISQ FOR 2X2 .05 PROB = 3.84',/,2X,'CHISQ F
   1OR 2X2 .01 PROB = 6.64',/,2X,'CHISQ FOR THIS RUN = ',F8.3,//)
    DO 898 IX=1,20
    XRUWKS(IX)=.05*IX-MDFR(IX)/XN100
    XREFKS(IX)=.05*IX-MFREF(IX)/XN100
    XRM1KS(IX)=.05*IX-MF2M1(IX)/XN100
    XRM2KS(IX)=.05*IX-MF5M2(IX)/XN100
    XCUWKS(IX)=.05*IX-MDFC(IX)/XN100
    XIUWKS(IX)=.05*IX-MDFI(IX)/XN100
    XCEFKS(IX)=.05*IX-MFCEF(IX)/XN100
    XIEFKS(IX)=.05*IX-MFIEF(IX)/XN100
    XCM1KS(IX)=.05*IX-MF3M1(IX)/XN100
    XIM1KS(IX)=.05*IX-MF1M1(IX)/XN100
    XCM2KS(IX)=.05*IX-MF4M2(IX)/XN100
```

```
      XIM2KS(IX)=.05*IX-MF1M2(IX)/XN100
  898 CONTINUE
      XRUWKS(21)=.01-MDFR(21)/XN100
      XREFKS(21)=.01-MFREF(21)/XN100
      XRM1KS(21)=.01-MF2M1(21)/XN100
      XRM2KS(21)=.01-MF5M2(21)/XN100
      XCUWKS(21)=.01-MDFC(21)/XN100
      XIUWKS(21)=.01-MDFI(21)/XN100
      XCEFKS(21)=.01-MFCEF(21)/XN100
      XIEFKS(21)=.01-MFIEF(21)/XN100
      XCM1KS(21)=.01-MF3M1(21)/XN100
      XIM1KS(21)=.01-MF1M1(21)/XN100
      XCM2KS(21)=.01-MF4M2(21)/XN100
      XIM2KS(21)=.01-MF1M2(21)/XN100
      D1=0
      D2=0
      D3=0
      D4=0
      D5=0
      D6=0
      D7=0
      D8=0
      D9=0
      D10=0
      D11=0
      D12=0
      DO 899 N=1,21
      DT=DABS(XRUWKS(N))
      IF(DT .GT. D1) D1=DT
      DT=DABS(XREFKS(N))
      IF(DT .GT. D2) D2=DT
      DT=DABS(XRM1KS(N))
      IF(DT .GT. D3) D3=DT
      DT=DABS(XRM2KS(N))
      IF(DT .GT. D4) D4=DT
      DT=DABS(XCUWKS(N))
      IF(DT .GT. D5) D5=DT
      DT=DABS(XCEFKS(N))
      IF(DT .GT. D6) D6=DT
      DT=DABS(XCM1KS(N))
      IF(DT .GT. D7) D7=DT
      DT=DABS(XCM2KS(N))
      IF(DT .GT. D8) D8=DT
      DT=DABS(XIUWKS(N))
      IF(DT .GT. D9) D9=DT
      DT=DABS(XIEFKS(N))
      IF(DT .GT. D10) D10=DT
      DT=DABS(XIM1KS(N))
      IF(DT .GT. D11) D11=DT
      DT=DABS(XIM2KS(N))
      IF(DT .GT. D12) D12=DT
  899 CONTINUE
```

```
      WRITE(6,990) D1,D2
      WRITE(6,991) D3,D4
      WRITE(6,992) D5,D6
      WRITE(6,993) D7,D8
      WRITE(6,994) D9,D10
      WRITE(6,995) D11,D12
      WRITE(6,996)
      WRITE(6,20)N11,N12,N21,N22
   20 FORMAT(//,5X,'N11 = ',I2,5X,'N12 = ',I2,5X,'N21 = ',I2,5X,'N22 = '
     1,I2,/)
      STOP
      END
/*
//GO.SYSIN DD *
```

BIBLIOGRAPHY

Books

Anderson, R. L. and T. A. Bancroft, Statistical Theory in Research, New York, McGraw-Hill Book Company, 1952.

Cochran, William G. and Gertrude M. Cox, Experimental Designs, New York, John Wiley and Sons, Inc., 1950.

Dayton, C. M., Design of Educational Experiments, New York McGraw-Hill Book Company, 1970.

Dixon, Wilfrid J. and Frank J. Massey, Jr., Introduction to Statistical Analysis, New York, McGraw-Hill Book Company, Inc., 1951.

Ferguson, George A., Statistical Analysis in Psychology & Education, New York, McGraw-Hill Book Company, 1971.

Glass, Gene V. and Julian C. Stanley, Statistical Methods in Education and Psychology, Englewood Cliffs, New Jersey, Prentice-Hall, 1970.

Johnson, Palmer O. and Robert W. B. Jackson, Modern Statistical Methods: Descriptive and Inductive, Chicago, Rand McNally and Company, 1959.

Kendall, Maurice G., The Advanced Theory of Statistics, London, Charles Griffin and Company Limited, 1948.

Keppel, Geoffrey, Design and Analysis: A Researcher's Handbook, Englewood Cliffs, New Jersey, Prentice-Hall, 1973.

Kirk, Roger E., Experimental Design: Procedures for the Behavioral Sciences, Belmont, California, Brooks/Cole Publishing Company, Inc., 1968.

Mood, Alexander McFarlane, Introduction to the Theory of Statistics, New York, McGraw-Hill Book Company, 1950.

Myers, Jerome L., Fundamentals of Experimental Design, Boston, Allyn and Bacon, Inc., 1972.

Ostle, Bernard, Statistics in Research, Ames, Iowa, The Iowa State College Press, 1954.

Overall, John E. and C. James Klett, Applied Multivariate Analysis, New York, McGraw-Hill Book Company, 1972.

Roscoe, John T., Fundamental Research Statistics for Behavioral Sciences, New York, Holt, Rinehart and Winston, Inc., 1975.

Scheffe', Henry, The Analysis of Variance, New York, John Wiley & Sons, 1959.

Siegel, Sidney, Nonparametric Statistics for the Behavioral Sciences, New York, McGraw-Hill Book Company, 1956.

Snedocer, George W., Statistical Methods, Ames, Iowa, The Iowa State College Press, 1946.

Wert, James E., Charles O. Neidt, and J. Stanley Almann, Statistical Methods in Educational and Psychological Research, New York, Appleton-Century-Crofts, Inc., 1954.

Winer, B. J., Statistical Principals in Experimental Design, New York, McGraw-Hill Book Company, 1971.

Articles

Appelbaum, Mark I. and Elliot M. Cramer, "Some Problems in the Nonorthogonal Analysis of Variance," Psychological Bulletin, LXXXI (1974), 335-343.

Carlson, James E. and Neil H. Timm, "Analysis of Nonorthogonal Fixed-Effects Designs," Psychological Bulletin, LXXXI (September, 1974), 563-570.

Clark, Charles E., "Importance Sampling in Monte Carlo Analyses," Operations Research, IX, no. 5 (September-October, 1961), 603-620.

Cohen, Jacob, "Multiple Regression as a General Data-Analytic System," Psychological Bulletin, LXX (1968), 426-443.

Falzer, Paul Robert, "Representative Design and the General Linear Model," Speech Monographs, XLI (June, 1974), 127-138.

Huck, Schuyler W. and Benjamin H. Layne, "Checking for Proportional n's in Factorial Anova's," Educational and Psychological Measurement, XXXIV (Summer, 1969), 281-287.

Jennings, Earl, "Fixed Effects Analysis of Variance by Regression Analysis," _Multivariate Behavioral Research_ (January, 1967), 95-107.

Keren, Gideon and Charles Lewis, "Nonorthogonal Designs: Sample Versus Population," _Psychological Bulletin_, LXXXIII (1976), 817-826.

Marks, Edmond, "Analysis of Variance Techniques for Unbalanced Data," _Review of Educational Research_, XLIV, no. 3 (Summer, 1974), 351-364.

Newman, Isadore and Michael T. Oravecz, "Solutions to the Problem of Disproportionality: A Discussion of the Models," _Multiple Linear Regression Viewpoints_, VII, no. 3 (1977), 1-25.

Overall, John E. and Douglas K. Spiegal, "Concerning Least Squares Analysis of Experimental Data," _Psychological Bulletin_, LXXII (1969), 311-322.

Overall, John E., Douglas K. Spiegal, and Jacob Cohen, "Equivalence of Orthogonal and Nonorthogonal Analysis of Variance," _Psychological Bulletin_, LXXXII (March, 1975), 182-186.

Overall, John E. and J. Arthur Woodward, "Multivariate Analysis of Variance by Multiple Regression Methods," _Psychological Bulletin_, LXXXII, no. 1 (1975), 21-32.

Steinhorst, R. Kirk and C. Dean Miller, "Disproportionality of Cell Frequencies in Psychological and Educational Experiments Involving Multiple Classifications," _Educational and Psychological Measurement_, XXIX (Winter, 1969), 799-811.

Tsao, Fei, "General Solution of the Analysis of Variance and Covariance in the Case of Unequal or Disproportionate Numbers of Observations in the Subclasses," _Psychometrika_, XI, no. 2 (June, 1946), 107-128.

__________, "Tests of Statistical Hypotheses in the Case of Unequal or Disproportionate Numbers of Observations in the Subclasses," _Psychometrika_, VII, no. 3 (September, 1942), 195-212.

Walberg, Herbert J., "Generalized Regression Models in Educational Research," _American Educational Research Journal_, VIII (January, 1971), 71-91.

Williams, John D., "Two Way Fixed Effects Analysis of Variance With Disproportionate Cell Frequencies," Multivariate Behavioral Research, VII (January, 1972), 67-82.

Williams, John D. and Alfred C. Lindem, "A Computer Program for Two-Way Fixed Effects Analysis of Variance with Disproportionate Cell Frequencies," Educational and Psychological Measurement, XXXII (1972), 169-172.