APRIL 1977

VOLUME 7 NUMBER 3

MULTIPLE LINEAR REGRESSION VIEWPOINTS

A publication of the Special Interest Group
on Multiple Linear Regression

MULTIPLE LINEAR REGRESSION VIEWPOINTS

Chairman .... Michael McShane, American Association of Medical Colleges
One Dupont Circle, Washington, D.C. 20036

Editor .... Isadore Newman, Research and Design Consultant,
The University of Akron, Akron, OH 44325

Assistant .... Diane Vukovich
The University of Akron, Akron, OH 44325

Secretary and Chairman-elect .... Steve Spaner, Behavioral Studies,
University of Missouri, St. Louis, MO 63121

Cover by .... David G. Barr

EDITORIAL BOARD

TABLE OF CONTENTS

Multiple Linear Regression Viewpoints
Vol. 7 No. 3 1977

SOLUTIONS TO THE PROBLEM OF DISPROPORTIONALITY: A DISCUSSION OF THE MODELS*

ISADORE NEWMAN

MICHAEL T. ORAVECZ

The University of Akron

ABSTRACT

There are two major purposes of the paper. The first is to investigate the usefulness of a χ^2 technique in differentiating between varying degrees of disproportionality and their effects on a Type I error. The second purpose is to present and support the position that the major concern for any research model, whether disproportionate or not, is the research question and how well that question is reflected by the model. Three "exact solutions" for disproportional situations, the hierarchial, unadjusted main effects, and fitting constant methods, will also be discussed in terms of the research question that each reflects, and examples will be presented to demonstrate the most appropriate situation for using each solution.

This paper will deal with two main questions. The first will be to identify the effects of different degrees of disproportionality on the nominal level of Type I error and to test the assumption that a χ^2 test can be used to determine when disproportionality is severe enough that corrections are required.

---

*This paper was presented at A.E.R.A., New York, NY, April, 1977.

The second part of this paper will deal with a discussion of some of the more prominent regression approaches ("exact" solutions) used to adjust for disproportionality.

The applied statistician and researcher is plagued with the problem of disproportional cell sizes in factorial experimental designs. This may occur because of mortality in the laboratory animals being used in the experiment; the required number of subjects not available; someone who had agreed to take part in the experiment fails to show up; or the data may represent the proportionality that exists in the "real world." (Disproportionality exists any time the expected values differ from the observed values. Obviously, one can have mild or severe disproportionality. The problem is to determine when the disproportionality is sufficiently severe to require adjustments. Therefore, this paper defines non-significant disproportionality on the bases of a χ^2 test where $\alpha = .25$ and significant disproportionality at $\alpha = .05$.) The effects on factorial designs by significantly disproportional versus non-significantly disproportional cell sizes, as determined by the chi-squared (χ^2) test, have not been studied, or at least reported. Such an investigation was attempted here.

Two studies were made, Study I and Study II. Each study contained four cases with 1,000 experiments per case, with the exception of the fourth case of Study II. In Study I, the main effects in each experiment were not necessarily significantly different. In Study II, the main effects in each experiment were made significantly different.

The following is a discussion of the paradigms used in each study.

The paradigm used in Study I was as follows:

1. Using a computer, 1,000 2 x 2 experiment tables were constructed for each of four cases--equal, proportional, non-significantly disproportional, and significantly disproportional cell sizes. Data points which were uniformly distributed and of equal variance were randomly generated in each of the cells [IBM, 1969].

2. Using computational formulas for sums of squares in the two-factor analysis of variance (ANOVA) with proportional cell frequencies, F-ratios and their corresponding probabilities were computed for each source of variability on each of the four thousand experiments.

3. The result of the analysis on the experiments containing equal frequencies was used as the measure of the actual probability level produced by this research procedure.

4. The probability levels produced by the analysis on the experiments containing significantly disproportional and non-significantly disproportional cell frequencies were compared to each other and to the actual probability level.

The paradigm used in Study II was the same as that for Study I with the exception that a value of five (5) was added to each of the data points of cells one and two so that the row factors were significantly different.

The procedures used in each of the two studies for obtaining the cell sizes, n_i $(i = 1, 2, 3, 4)$, and to determine if the cell sizes were significantly and non-significantly disproportional were as follows:

Case I: Equal n's

A random number between ten and one hundred was generated and this became the size of each cell.

Case II: Proportional n's*

Since the cell sizes are proportional if $\frac{n_1}{n_2} = \frac{n_3}{n_4}$,

n_1, n_2, n_3 were randomly generated and then n_4 was calculated from the equation $n_4 = \frac{n_2 \; n_3}{n_1}$,

where $10 \leq n_i \leq 100$, $i = 1, 2, 3$. See figure below for the labeling of the cells.

n_1	n_2
n_3	n_4

*Determining n_4 in this manner most often resulted in actual values that were non-integer numbers which the computer truncated to the integer value. Hence, n_4 differed from the true proportional value an amount x where $0 \leq x < 1$.

Case III: Non-significant disproportional n's

Cell sizes n_1, n_2, and n_3, were randomly generated, and then the proportional cell size n_4 was computed as in Case II. Then, so that the cell sizes were disproportional, a new n_4, n'_4, was calculated from the equation

$$n'_4 = n_4 + x + 1$$

where x is a randomly generated integer between one and ten. To determine if the cell sizes were non-significantly disproportional, a χ^2 test was made ($p \geq \alpha$ .25) where χ^2_c = 1.32--i.e., if $\chi^2_a \leq \chi^2_c$, the n's were judged to be non-significantly disproportional.

Case IV: Significant disproportional n's

Cell sizes n_1, n_2, n_3, and n_4, were randomly generated. To determine if the cells were significantly disproportional, a χ^2 test was made ($p \geq \alpha$ .05) where χ^2_c = 3.841--i.e., if $\chi^2_a > \chi^2_c$, the n's were judged to be significantly disproportional.

The probability of a computed F-ratio was computed by using formulas 26.6.2, 26.6.10, 26.7.8, and 26.2.18 found in Abramowitz and Stegun [1967]. Algebraically, it can be shown that the above formulas combine so that the probability of a F-value for degrees of freedom in the numerator, v_1, and degrees of freedom in the denominator, v_2, is

$$P(F \mid v_1 = 1, v_2)^* = 1 - (1 + c_1x + c_2x^2 + c_3x^3 + c_4x^4)^{-4}$$

---

*P = .95 means that five times out of one hundred an F-value this large or larger would occur due to chance alone.

where $c_1 = .196854$, $c_2 = .115194$, $c_3 = .000344$, $c_4 = .019527$,

$$x = \frac{t\left(1 - \frac{1}{4v_2}\right)}{\sqrt{1 + \frac{t^2}{2v_2}}}, \text{ and } t = \sqrt{F}.$$

TABLE I

PERCENTAGE OF SIGNIFICANT F-RATIOS FOR EACH OF THE FOUR CASES OF STUDY I

SOURCE	EQUAL n's	PROPORTIONAL n's	NON-SIGNIFICANTLY DISPROPORTIONAL n's	SIGNIFICANTLY DISPROPORTIONAL n's
Row Factor	5.60%	4.10%	4.60%	3.60%
Column Factor	4.60%	4.60%	4.40%	3.40%
Interaction	5.40%	5.40%	6.10%	6.90%

TABLE II

PERCENTAGE OF SIGNIFICANT F-RATIOS FOR EACH OF THE FOUR CASES IN STUDY II

SOURCE	EQUAL n's	PROPORTIONAL n's	NON-SIGNIFICANTLY DISPROPORTIONAL n's	SIGNIFICANTLY DISPROPORTIONAL n's
Row Factor	100.00%	100.00%	100.00%	100.00%
Column Factor	6.40%	3.90%	3.50%	**
Interaction	4.80%	6.50%	11.10%	**

** Interaction sums of squares were negative values. The program was written to reject negative F-values. Hence, the process stayed within a loop and no results were available.

The Monte Carlo data supports the position that mild cases of disproportionality versus severe cases of disproportionality, as operationally defined by the χ^2, may be a useful approach in identifying the limits of when one cannot ignore effects of disproportional cells.

As one can see from Table I, when the null hypothesis is true, it appears that disproportionality does not present a problem. However, when the null hypotheses is not true and there is significant disproportionality, as measured by a χ^2 test ($p \leq .05$), the effects of the disproportionality cannot be ignored. In addition, the non-significant disproportionality case, ($.25 < p < 1$) as can be seen from Table II, has more frequently occurring significant interactions than occurred in the equal and proportional cases. Therefore, even in the non-significant case, the interaction may be more severely effected, and one should be sensitive to this possibility.

Obviously, this Monte Carlo study did not investigate the effects of both row and column being significant, row column and interaction being significant, or either row or column and interaction being significant. The reader should be aware that these situations may or may not produce different results than are presented in Tables I and II. However, the investigators feel that at this point, they have produced some supportive evidence for the potential usefulness for using a χ^2 to aid in decision making for determining when it is necessary to use corrections.

THREE REGRESSION SOLUTIONS FOR DISPROPORTIONALITY

When the researchers feel disproportionality is severe enough to be of concern, there are a variety of procedures that he can utilize to attempt to correct for the potential problems. However, before any corrections are applied, one should be sensitive to the underlying assumption that they are making about the population from which their data is drawn, and the investigator must also be very clear about the research question he is interested in asking.

If one had a research project in which the data and variables came from groups that already exist, such as age, intelligence, socio-economic status, etc., and if one was interested in generalizing and predicting back to the group from which the sample came, there is a good possibility that there would be a correlation that would not be spurious between such variables as I.Q. and socio-economic status. In other words, there may be significantly more above average socio-economic status people who have above average I.Q.'s than one would expect by chance. The data were forced to correspond to a balanced design in which there are an equal number of high and low I.Q. people for an equal number of high and low socio-economic status positions, the result of the study and the statistical analysis may allow one to say something that may only be true for that artificially forced relationship and one could not properly generalize to the population in which this proportionality did not actually exist.

The other side of the coin is if the disproportionality in a research design is an artifact, (it really does not exist in the population) and the disproportionality of vectors is causing a spurious correlation between the variables, then one would have to adjust for this disproportionality and would have to decide which solution of disproportionality would best adjust the data so that it would better reflect the question(s) of interest and the true state of affairs.

To be able to begin to decide upon the correct solution, one has to

a. know something about the theoretical and/or empirical relationship between the variables being studied;

b. know some of the descriptive data about the population one wishes to generalize to in relation to the specific variables being studied;

c. know the specific research question under investigation if one decides an adjustment for disproportionality is needed, then

d. know the underlying assumptions and implications for different adjustment procedures, and

e. know the consequences for using the selected adjustment procedure on the interpretation and generalization of the data.

Many researchers have dealt with items a - c but few have clarified the problems and implications related to d and e on the above list. The following is a brief discussion of the underlying assumptions of some of the frequently used adjustments for disproportionality. It is hoped that this would aid researchers in being more sensitive to the questions they are really asking when writing models that reflect adjustment by different solutions for the problem of disproportionality.

There has been an exciting and thought provoking debate in the literature as to the most desirable (correct or accurate) procedure for correcting the problems of disproportionality. Most notably, Overall and Spiegel, 1969; Timm and Carlson, 1975; Overall, Spiegel, and Cohen, 1975; Werts and Linn, 1971; Rock, Werts, and Linn, 1976; and Applebaum and Cramer, 1973; have created the interest in the literature.

One basic way of organizing discussions on this topic is to categorize suggestions into two broad groups: approximate solutions or least sum of squares solutions. (This paper will discuss only the least sum of squares solutions.)

Three prominent least sum of squares solutions for disproportionality will be defined.

Solution I is the use of the general linear model to simultaneously adjust for the correlations between the main effects and the main effects with interaction. A symbolic example of this procedure is presented below for a two factorial design.

Model 1 $Y_{kab} = \delta + b_1\alpha_a + b_2\beta_b + b_3\alpha\beta_{ab} + \varepsilon_{kab}$

Model 2 $Y_{kab} = \delta + b_4\beta_b + b_5\alpha\beta_{ab} + \varepsilon_{kab}$

Model 3 $Y_{kab} = \delta + b_6\alpha_a + b_7\alpha\beta_{ab} + \varepsilon_{kab}$

Model 4 $Y_{kab} = \delta + b_8\alpha_a + b_9\beta_b + \varepsilon_{kab}$

Y_{kab} = is the score for subject k in row a and column b

δ = is the grand $\overline{X}$

α_a = is the effect for row "a"

β_b = is the effect for column "b"

$\alpha\beta_{ab}$ = is the interaction effect for the row "a" and column "b"

ε_{kab} = is the error term for each subject

$b, \ldots b_n$ are partial regression coefficients

<u>Adjustment for Solution #1</u>

<u>Adjustment for A main effects</u> test Model 1 against Model 2

<u>Adjustment for B main effects</u> test Model 1 against Model 3

<u>Adjustment for A*B effects</u> test Model 1 against Model 4

(see also Figure I, P. 19)

Solution 2 adjusts each main effect in terms of the other main effects. The interaction was adjusted for by all main effects. (This is the same as in Solution 1). The following is a symbolic representation of this solution: (see also Figure II, P. 20)

Adjustment for Solution #2

Model 4 $Y_{kab} = \delta + b_{10}\alpha_a + b_{11}\beta_b + \varepsilon_{kab}$

Model 5 $Y_{kab} = \delta + b_{12}\beta_b + \varepsilon_{kab}$

Model 6 $Y_{kab} = \delta + b_{13}\alpha_a + \varepsilon_{kab}$

Adjustment for A main effects test Model 4 against Model 5

Adjustment for B main effects test Model 4 against Model 6

Adjustment for AB interaction effects test Model 4 against Model 1

Solution 3 assumes an apriori ordering of the importance of the variables under investigation. The apriori ordering decides which variables one will allow to account for as much variance as possible by themselves. The following is a symbolic representation of Solution 3, assuming the researcher considers the A main effects most important, B main effects second, and the interaction least important.

(Many researchers feel that it is unlikely that most investigators will be able to order the importance of their variable. However, we believe this judgment can be made by a competent researcher who is aware of the underlying constructs and theories he is dealing with) (see also Figure III, P. 21)

Admustment for Solution #3

Model 7 $Y_{kab} = \delta + b_{14}\alpha_a + \varepsilon_{kab}$

Model 8 $Y_{kab} = \delta + \varepsilon_{kab}$

Model 9 $Y_{kab} = \delta + b_{15}\alpha_a + b_{16}\beta_b + \varepsilon_{kab}$

Adjustment for A main effects test Model 7 against Model 8

Adjustment for B main effects test Model 9 against Model 7

Adjustment for AB interaction test Model 1 against Model 9

Marks (1974) presents a mathematical proof that one has to have to A main effects and B main effects in the full and restricted models to test for AB interaction. It seems that this is true when dealing with traditional analysis of variance and catagorical variables. That is, the A main effects and B main effects must be fitted first before interaction can be tested. However, this does not seem to be necessarily true when dealing with continuous variables. To the best of our knowledge, this has never been investigated.

Most researchers, specifically the highly mathematically oriented statisticians, tend to evaluate the accuracy of the above three least square solutions for disproportionality in terms of some arbitrary mathematical matrices often based upon unfounded underlying assumptions. A more logical method for deciding upon which solution is most appropriate in a specific case would be to be aware of the underlying assumptions for each of the structural models of each solution and assumptions about the relationship between the variables being studied.

In a 2 x 2 orthogonal design, the correlation between the A main effect, B main effect, and the AB interaction is zero. Therefore, the effects of the A main effect in predicting the criterion can be totally attributed to the A main effect; similarly for the B main effect and the AB interaction. However, when the design is non orthogonal (disproportional n's) correlations between the A main effects, B main effects, and AB interaction are likely. Therefore, the effect of A in predicting the criterion is not necessarily independent of the B main effect and AB interaction.

Each of the three least square solutions make different assumptions about the meaningfulness and "usefullness" of the correlations between the A main effect, B main effect, and AB interaction.

Solution 1, for example, when testing the A main effect, assumes the correlation between A and B and the AB interaction is of an accidental nature, and therefore should not be considered (Rock, et al., 1976). This solution is most likely to be prefered when one can assume that the missing subjects producing disproportionality were random. If one is unable to make this assumption, then it would be inappropriate to use Solution 1, (which may be the case most frequently).

Solution 2 assumes that there is no correlation between the A and B main effects in the population. Therefore, the correlation between A and B in the sample is a function of disproportionality and not representative of the population. Solution 2 then attempts to adjust for this correlation.

However, Solution 2 assumes that the correlations between the main effects and the interaction, which results from the disproportionality, are not spurious and are characteristic of the population. Therefore, it does not attempt to adjust for this correlation.

If one cannot assume that the correlations between the A and B main effects, due to disproportionality, are due to chance, than Solution 2 would be an inappropriate correction.

Solution 3 requires an apriori ordering of the importance of each variable. Let us assume that the apriori ordering are: A main effects, B main effects, AB interaction, respectively.

When testing for the A main effects, Solution 3 assumes that the correlation between the A main effects, B main effects, and A main effects with the AB interaction is representative of the relationship between these variables in the population and therefore relevant to the research question of interest. That is to say, the relationship between these variables are not artifacts of disproportionality.

However, when testing for B main effects, Solution 3 assumes that the correlation between the A main effect and B main effect is not relevant and is therefore an artifact of unequal N's. It also assumes that the correlation between the A main effects and the interaction, and the B main effects and the interaction is not due to just a chance relationship caused by disproportionality.

The key to the most appropriate selection of these three solutions is not as much a statistical concern (mathematical) but rather a comprehensive understanding of the data so that one can more accurately speculate on which correlations between the variables are likely to be meaningful or not meaningful.

FIGURE I

method I

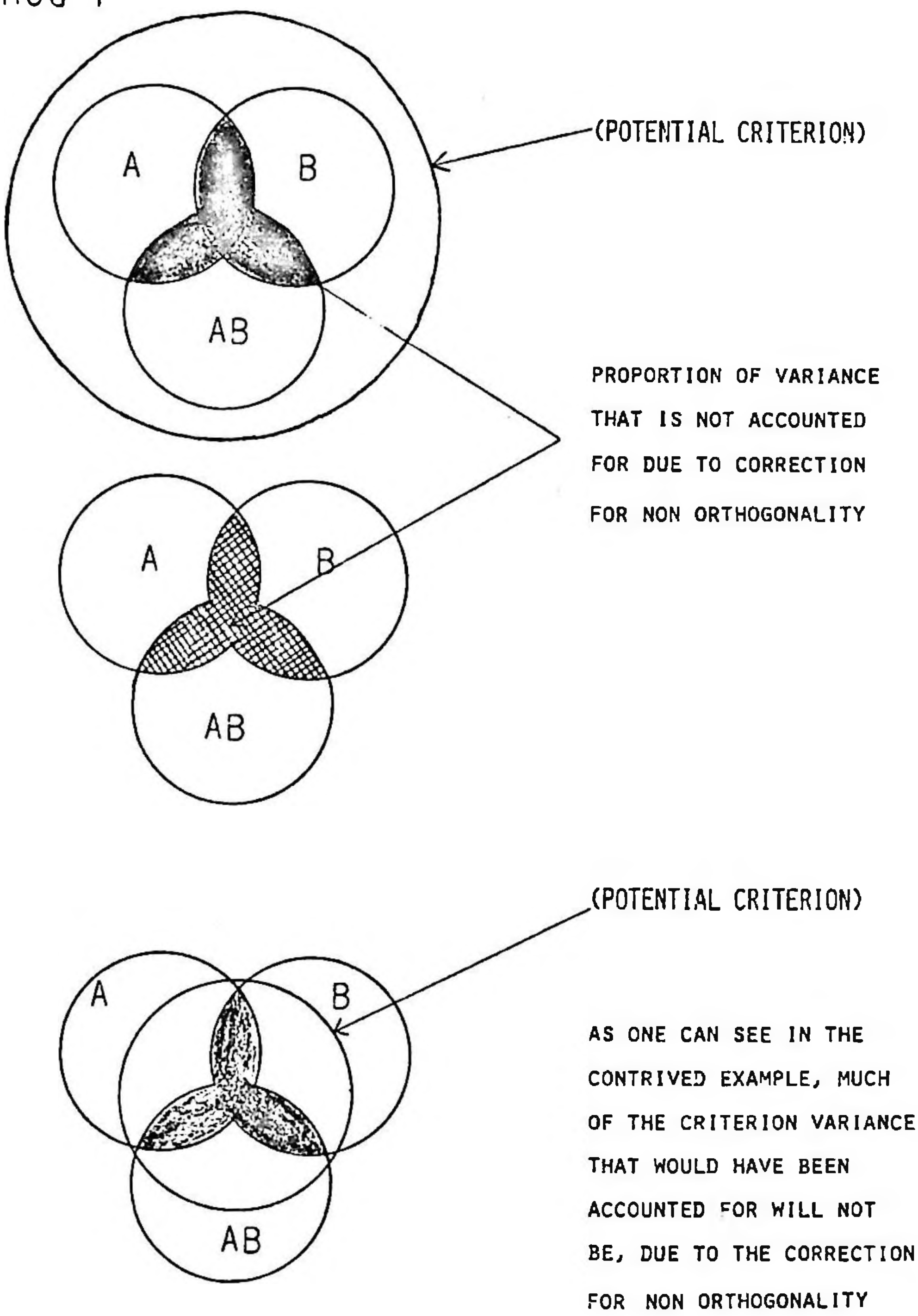
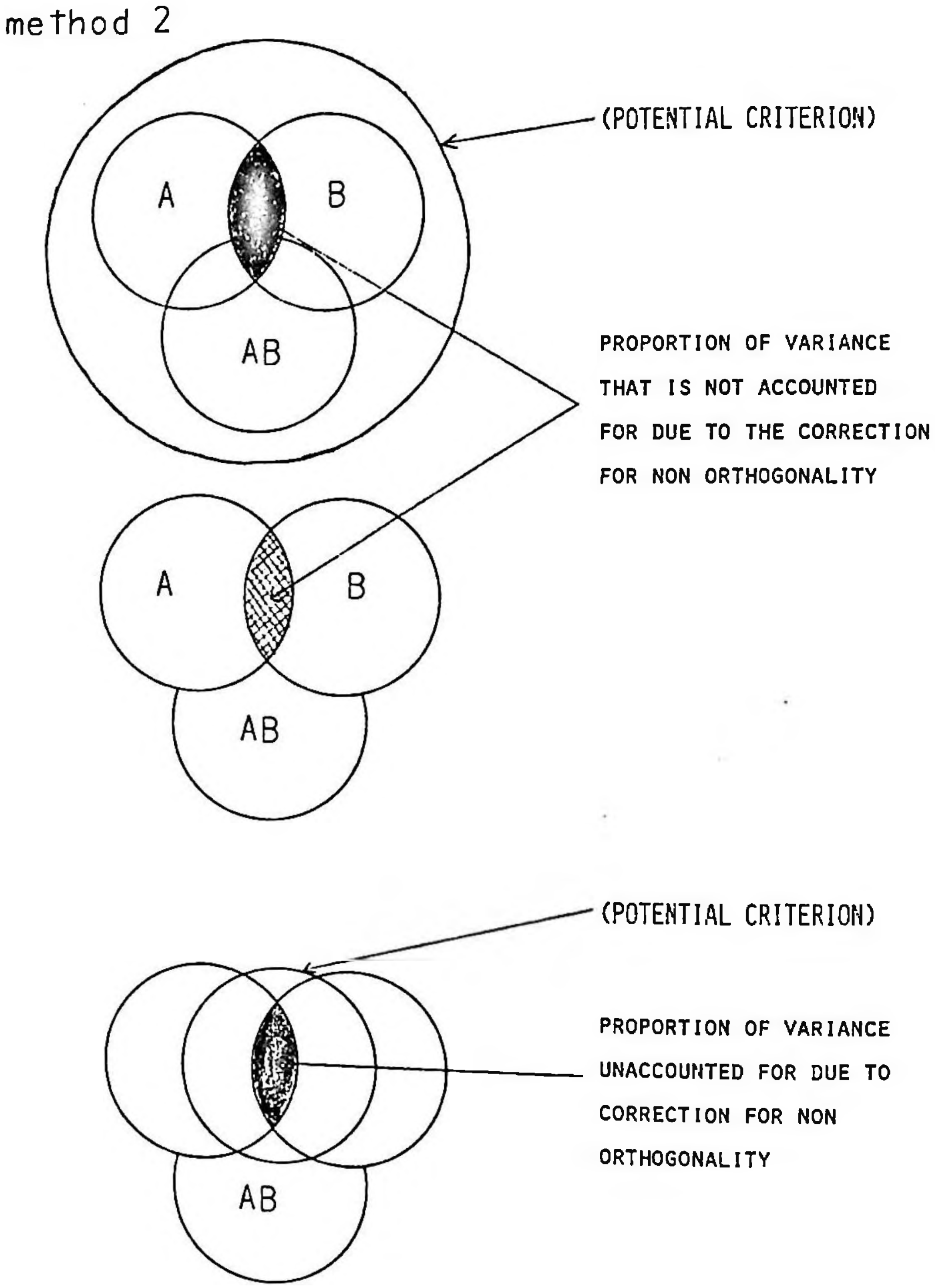


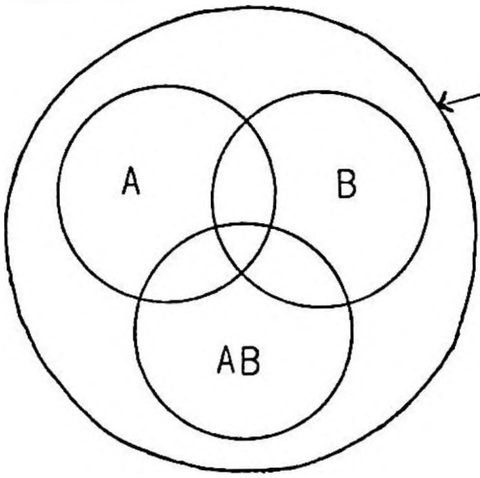
FIGURE II



method 2
(POTENTIAL CRITERION)
A
B
AB
PROPORTION OF VARIANCE
THAT IS NOT ACCOUNTED
FOR DUE TO THE CORRECTION
FOR NON ORTHOGONALITY
A
B
AB
(POTENTIAL CRITERION)
PROPORTION OF VARIANCE
UNACCOUNTED FOR DUE TO
CORRECTION FOR NON
ORTHOGONALITY
AB

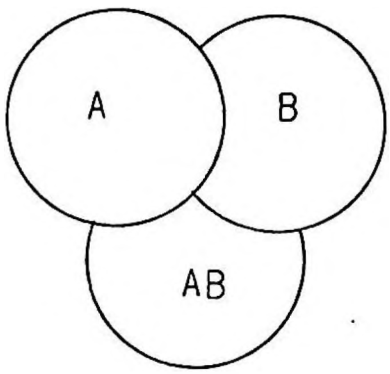
FIGURE III

method 3



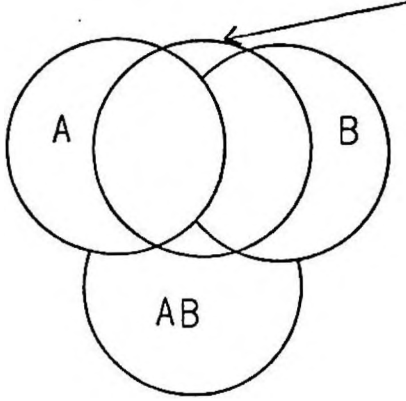
(POTENTIAL CRITERION)

ASSUMING THAT THE ORDERING OF THE VARIABLES OF INTEREST ARE A, B, AB, RESPECTIVELY.



THERE IS NO VARIANCE UNACCOUNTED FOR DUE TO THIS CORRECTION FOR NON ORTHOGONALITY

(POTENTIAL CRITERION)



SOME ADDITIONAL CONSIDERATIONS

Since the three solutions previously discussed are only defined for use with fixed designs, it is initially important to define the basic statistical models of fixed randomized and mixed effects. The fixed effect model is a model in which all the variables have levels which are fixed. That is, the levels of a variable that is to be investigated are determined prior to the investigation. These categorical (high, medium low, etc.,) determinations set the limits to which the investigator can generalize his results.

Randomized effect is a design in which the researcher randomly selects the level of variable that is to be investigated from an infinitely large number of possible levels. This allows the researcher to generalize his results to the entire range of the variable being investigated (that is to the extent that the levels were truly randomly selected and are representative).

The mixed effect model has at least one fixed effect variable and one random effect variable.

Most research conducted is on fixed effect models. This model has a great deal research behind it which indicated that the model is very robust. This means, it is little effected by violations of its assumptions of normality, homogeneity,etc. However, this robustness is not true of the randomized effects model. This model is very sensitive to violations of normality and homogeneity of variance. Most interestingly, in terms of the purpose of this paper, is that the assumptions of equal N's is absolutely essential for deriving correct error terms for randomized effect designs.[1]

Assuming that one is dealing only with a fixed effect design, one alternative is to consider the correlation between the variables, due to the disproportionality as accurately existing in the population one wishes to generalize to. Then, the researcher may choose not to correct for the disproportionality. The problems that arise with this, is that one cannot attribute the variance accounted for to a particular variable. (This is the dilema of much ex post facto research.)

In conclusion, as can be seen from the discussion in Part II of this paper, it appears obvious that it would be inappropriate to attempt to decide on a correction for disproportionality by using Monte Carlo Studies. Since the appropriateness of a solution depends upon the correlations as one can assume to exist between the variables in the population, and this can change from population to population or with varing theoretical positions, it is much more relevant to try to understand these relationships and base your solution on this understanding than to base

---

[1]To our knowledge, there has been no research on methods for correcting unequal N's in randomized designs.

your decision on the outcome of generated data in which the relationships may be totally unrelated to the relationships in the population of interest.

A highly detailed discussion of different solutions to the hypothesis being tested can be found in Searle (1971). There exists problems that go beyond the scope of this paper such as what occurs when interactions are included in the model and one or more cells equal zero; Searle (1971), Marks (197), and Williams (1977), deal with the problems of full versus non-full ranked matrices; Williams (1977) indicates potentially different effects on solutions due to the effects of different coding procedures. All of these questions are of interest. owever, once again the key is to understand your research questions and the assumptions about the relationships that exist among your variables as reflected by your models.

REFERENCES

Abramowitz, M., & Stegun, I. A. (Eds). Handbook of mathematical functions with formulas, graphs, and mathematical tables. (National Bureau of Standars Applied Mathematics Series 55) Washington, D.C.: United States Government Printing Office, 1967.

Applebaum, M. I. and Cramer, E. M. Some problems in the nonorthogonal analysis of variance. Psychological Bulletin, 1974, 81, 335-343.

Carlson, J. E., & Timm, N. H. Analysis of nonorthogonal fixed-effeccts designs. Psychological Bulletin, 1974, 81 (9), 563-570.

Cohen, J., & Cohen, P. "Multiple regression as a general data analytic system," Psychological Bulletin, 1968 70, pp. 426-443.

Dalton, Starrett. Regression approaches and approximate solutions to analysis of variance with disproportionality varied. Multiple Linear Regression Viewpoints. Vol. 7, 2, Jan., 1977.

Darlington, R. B. Multiple regression in research and practice. Psychological Bulletin, 1968, 69, 166-168.

Glass, G. V., & Stanley, J. C. Statistical methods in education and psychology. Englewood Cliffs, N.J.: Prentice-Hall, 1970.

I.B.M. Random number generation and testing. White Plains, N.Y.: I.B.M., Technical Publications Dept., 1969.

Marks, Edmond. Analysis of variance techniques for unbalanced data. Review of Educational Research. Vol. 44, 3, 1974.

Overall, J. E., & Spiegel, D. K. Concerning least squares analysis of experimental data. Psychological Bulletin, 1969, 72, 311-322.

Overall, J. E., Spiegel, D. K., & Cohen, J. Equivalence of orthogonal and nonorthogonal analysis of variance. Psychological Bulletin, 1975, 82, 182-186.

Rock, Werts, & Linn. Structural equations as an aid in the interpretation of the non-orthogonal analysis of variance. Multivariate Behavioral Research, 1976, 443-448.

Speed, F. M. & Hocking R. R. The use of the R( )-notation with unbalanced data. The American Statistican. Feb., 1976, Vol. 30, 1.

Timm, N. H., & Carlson, J. E. Analysis of variance through full rank models. Monograph issued by Multivariate Behavioral Research, 1975.

Williams, J. D. Two way fixed effects analysis of variance with disproportionate cell frequencies. Multivariate Behavioral Research, 1972, 67-83.

Williams, J. D. "Full rank and non full rank models with contrast and binary coding systems for a two-way disproportionate cell frequency analysis." (In press)

APPENDIX

```
C ON110 ORAVECZ-NEWMAN HYPOTHESIS - EQUAL N'S CASE
C MODIFIED FROM ON100   3-2-76  WD WHEATON    ROW FACTOR NOT
                                              NECESSARILY SIGNIFICANT
C DEFINITION - 2X2 CONTINGENCY TABLE
C   CELL(1,1) EQU CELL(1)
C   CELL(1,2) EQU CELL(2)
C   CELL(2,1) EQU CELL(3)
C   CELL(2,2) EQU CELL(4)
      REAL*4 P(4),SUM(4),SUMS(4),F(3),PF(3),PRCT(6)
      INTEGER*4 N(4),NSIG(3),DATE(4)
      EQUIVALENCE (SUM(1),SUM1),(SUM(2),SUM2),(SUM(3),SUM3),
     1(SUM(4),SUM4),(N(1),N1),(N(2),N2),(N(3),N3),(N(4),N4),
     2(F(1),FR),(F(2),FC),(F(3),FRC)
  200 FORMAT('1',35X,'ON110  ORAVECZ & NEWMAN EQUAL N''S CASE   '
     1,4A2,T123,'PAGE',I3//
     2'              CELL SIZE RATIOS                    CELL POPULATION SIZE
     3               F VALUES                              PROBABILITIES'/
     4' -----------------------------------   -----------------------------
     5  -----------------------------------   -----------------------------'/
     6'  (1,1)     (1,2)     (2,1)     (2,2)   (1,1)  (1,2)  (2,1)  (2,2)
     7      ROW          COL          TABLE       ROW        COL     TABLE'/
     8' -------  -------  -------  -------   -----  -----  -----  -----
     9  ----------  ----------  ----------   -------  -------  -------')
  201 FORMAT(F8.5,3F9.5,1X,4I7,1X,3E12.4,1X,3F9.5)
  202 FORMAT('1ON110  ORAVECZ & NEWMAN EQUAL N''S CASE  ',4A2///
     111X,'PERCENT OF CASES'/
     210X,18('-')/
     312X,'SIG       NON-SIG'/
     410X,'-------    -------'/
     5' ROW   ',2F11.2//
     6' COL   ',2F11.2//
     7' TABLE',2F11.2///' TOTAL CASES RUN =',I6)
C GET CURRENT DATE
      CALL TDATE(DATE)
C CONSTANT NEEDED LATER  D1=2./9.=0.222222
      D1=0.222222
C PROBABILITY EXPANSION CONSTANTS
      CC1=0.196854
      CC2=0.115194
      CC3=0.000344
      CC4=0.019527
C NUMBER OF CASES TO RUN
      NC=1000
C NUMBER OF NON-SIGNIFICANT CASES - ROW,COL & TABLE
      NSIG(1)=0
      NSIG(2)=0
      NSIG(3)=0
C PAGE & LINE CNTS
      LINE=100
      IPAGE=1
      IY=9
C LOOP OVER ALL CASES
      DO 306 K=1,NC
C GENERATE CELL SIZE  (.GT. 10)
  300 IX=IY
      CALL RANDU(IX,IY,PP)
      NN=PP*100.
      IF(NN.LT.10) GO TO 300
C SET ALL CELLS EQUAL & SET CELL PROBABILITIES TO ONE
      DO 301 I=1,4
```

```
      N(I)=NN
      P(I)=0.25
  301 CONTINUE
      NSUM=4*NN
C GENERATE MEMBERS OF EACH CELL & CELL SUM
  310 TSUMS=0.0
      DO 303 I=1,4
      SUM(I)=0.0
       SUMS(I)=0.0
      L=N(I)
      DO 304 J=1,L
      IX=IY
      CALL RANDU(IX,IY,A)
      SUM(I)=SUM(I)+A
      SUMS(I)=SUMS(I)+A*A
  304 CONTINUE
      TSUMS=TSUMS+SUMS(I)
  303 CONTINUE
C CALC ROW,COL,& TABLE SUMS
      R1S=SUM1+SUM2
      R2S=SUM3+SUM4
      C1S=SUM1+SUM3
      C2S=SUM2+SUM4
      TABSUM=R1S+R2S
C CALC ROW,COL,$ TABLE SUMS SQUARED
      TSS=TABSUM*TABSUM/FLOAT(NSUM)
      SSR=R1S*R1S/(N1+N2)+R2S*R2S/(N3+N4)-TSS
      SSC=C1S*C1S/(N1+N3)+C2S*C2S/(N2+N4)-TSS
      SST=TSUMS-TSS
      SSCELL=SUM1*SUM1/N1+SUM2*SUM2/N2+SUM3*SUM3/N3+SUM4*SUM4/N4-TSS
      SSRC=SSCELL-SSR-SSC
      SSERR=SST-SSR-SSC-SSRC
C CALC F VALUES FOR ROWS,COLS,& TABLE
      DF=NSUM-4
      SERRM=SSERR/DF
      FR=SSR/SERRM
      FC=SSC/SERRM
      FRC=SSRC/SERRM
C IF F(I) IS NEGATIVE DROP CASE FROM CONSIDERATION
      IF(FR.LT.0.0) GO TO 310
      IF(FC.LT.0.0) GO TO 310
      IF(FRC.LT.0.0) GO TO 310
C CALC PROBABILITIES FOR ROWS,COLS & TABLE
      D2=D1/DF
      DO 305 I=1,3
      X1=SQRT(F(I))*(1.-0.25/DF)/SQRT(1.+0.5*F(I)/DF)
      X2=1.0+X1*(CC1+X1*(CC2+X1*(CC3+X1*CC4)))
      X3=X2*X2*X2*X2
      PF(I)=1.-1./X3
C CNT NON-SIGNIFICANT CASES - ROW,COL & TABLE
      IF(PF(I).LE.0.95) NSIG(I)=NSIG(I)+1
  305 CONTINUE
C PRINT CASE PARAMETERS
      IF(LINE.LE.56) GO TO 307
      WRITE(6,200)DATE,IPAGE
      LINE=6
      IPAGE=IPAGE+1
  307 WRITE(6,201)P,N,F,PF
      LINE=LINE+1
```

```
  306 CONTINUE
C CALC SIGNIFICANT CASES,PERCENTAGES, & PRINT SUMMARY
      ANC=NC
      DO 308 L=1,6,2
      M=L/2+1
      LSIG=NSIG(M)
      MSIG=NC-LSIG
      PRCT(L)=FLOAT(MSIG)*100/ANC
      PRCT(L+1)=FLOAT(LSIG)*100/ANC
  308 CONTINUE
      WRITE(6,202)DATE,PRCT,NC
      CALL EXIT
      END
```

```
C ON110 ORAVECZ-NEWMAN HYPOTHESIS - EQUAL N'S CASE
C MODIFIED FROM ON100 3-2-76 WD WHEATON . Row FACTOR SIGNIFICANT
C DEFINITION - 2X2 CONTINGENCY TABLE
C  CELL(1,1).EQU CELL(1)
C  CELL(1,2) EQU CELL(2)
C  CELL(2,1) EQU CELL(3)
C  CELL(2,2) EQU CELL(4)
      REAL*4 P(4),SUM(4),SUMS(4),F(3),PF(3),PRCT(6)
      INTEGER*4 N(4),NSIG(3),DATE(2)
      EQUIVALENCE (SUM(1),SUM1),(SUM(2),SUM2),(SUM(3),SUM3),
     1(SUM(4),SUM4),(N(1),N1),(N(2),N2),(N(3),N3),(N(4),N4),
     2(F(1),FR),(F(2),FC),(F(3),FRC)
  200 FORMAT('1',35X,'ON110  ORAVECZ & NEWMAN EQUAL N''S CASE  '
     1,4A2,T123,'PAGE',I3//
     2'              CELL SIZE RATIOS                         CELL POPULATION SIZE
     3                  F VALUES                                 PROBABILITIES'/
     4' ----------------------------------------   ------------------------
     5  ----------------------------------------   ------------------------'/
     6'  (1,1)     (1,2)     (2,1)     (2,2)     (1,1)  (1,2)  (2,1)  (2,2)
     7    ROW              COL             TABLE     ROW     COL      TABLE'/
     8' -------   -------   -------   -------    -----  -----  -----  -----
     9  ---------   ---------   ---------    -------  -------  -------')
  201 FORMAT(F8.5,3F9.5,1X,4I7,1X,3E12.4,1X,3F9.5)
  202 FORMAT('1CN110  ORAVECZ & NEWMAN EQUAL N''S CASE  ',2A4///
     111X,'PERCENT OF CASES'/
     210X,18('-')/
     312X,'SIG      NON-SIG'/
     410X,'-------     -------'/
     5' ROW   ',2F11.2//
     6' COL   ',2F11.2//
     7' TABLE',2F11.2///' TOTAL CASES RUN =',I6)
C GET CURRENT DATE
      CALL TDATE(1,DATE)
C CONSTANT NEEDED LATER  D1=2./9.=0.222222 .
      D1=0.222222
C PROBABILITY EXPANSION CONSTANTS
      CC1=0.196854
      CC2=0.115194
      CC3=0.000344
      CC4=0.019527
C NUMBER OF CASES TO RUN
      NC=1000
C NUMBER OF NON-SIGNIFICANT CASES - ROW,COL & TABLE
      NSIG(1)=0
      NSIG(2)=0
      NSIG(3)=0
C PAGE & LINE CNTS
      LINE=100
      IPAGE=1
      IY=9
C LOOP OVER ALL CASES
      DO 306 K=1,NC
C GENERATE CELL SIZE  (.GT. 10)
  300 IX=IY
      CALL RANDU(IX,IY,PP)
      NN=PP*100.
      IF(NN.LT.10) GO TO 300
C SET ALL CELLS EQUAL & SET CELL PROBABILITIES TO ONE
      DO 301 I=1,4
```

```
      N(I)=NN
      P(I)=0.25
 301  CONTINUE
      NSUM=4*NN
C GENERATE MEMBERS OF EACH CELL & CELL SUM
 310  TSUMS=0.0
      DO 303 I=1,4
      SUM(I)=0.0
       SUMS(I)=0.0
      L=N(I)
      DO 304 J=1,L
      IX=IY
      CALL RANDU(IX,IY,A)
      IF(I.LE.2) A=A+5
      SUM(I)=SUM(I)+A
      SUMS(I)=SUMS(I)+A*A
 304  CONTINUE
      TSUMS=TSUMS+SUMS(I)
 303  CONTINUE
C CALC ROW,COL,& TABLE SUMS
C   CALC ROW,COL,& TABLE SUMS SO ROW FACTOR SIGNIF
      R1S=SUM1+SUM2
      R2S=SUM3+SUM4
      C1S=SUM1+SUM3
      C2S=SUM2+SUM4
      TABSUM=R1S+R2S
C CALC ROW,COL,$ TABLE SUMS SQUARED
      TSS=TABSUM*TABSUM/FLOAT(NSUM)
      SSR=R1S*R1S/(N1+N2)+R2S*R2S/(N3+N4)-TSS
      SSC=C1S*C1S/(N1+N3)+C2S*C2S/(N2+N4)-TSS
      SST=TSUMS-TSS
      SSCELL=SUM1*SUM1/N1+SUM2*SUM2/N2+SUM3*SUM3/N3+SUM4*SUM4/N4-TSS
      SSRC=SSCELL-SSR-SSC
      SSERR=SST-SSR-SSC-SSRC
C CALC F VALUES FOR ROWS,COLS,& TABLE
      DF=NSUM-4
      SERRM=SSERR/DF
      FR=SSR/SERRM
      FC=SSC/SERRM
      FRC=SSRC/SERRM
C IF F(I) IS NEGATIVE DROP CASE FROM CONSIDERATION
      IF(FR.LT.0.0) GO TO 310
      IF(FC.LT.0.0) GO TO 310
      IF(FRC.LT.0.0) GO TO 310
C CALC PROBABILITIES FOR ROWS,COLS & TABLE
      D2=D1/DF
      DO 305 I=1,3
      X1=SQRT(F(I))*(1.-0.25/DF)/SQRT(1.+0.5*F(I)/DF)
      X2=1.0+X1*(CC1+X1*(CC2+X1*(CC3+X1*CC4)))
      X3=X2*X2*X2*X2
      PF(I)=1.-1./X3
C CNT NON-SIGNIFICANT CASES - ROW,COL & TABLE
      IF(PF(I).LE.0.95) NSIG(I)=NSIG(I)+1
 305  CONTINUE
C PRINT CASE PARAMETERS
      IF(LINE.LE.56) GO TO 307
      WRITE(6,200)DATE,IPAGE
      LINE=6
      IPAGE=IPAGE+1
```

```
 307  WRITE(6,201)P,N,F,PF
      LINE=LINE+1
 306  CONTINUE
C CALC SIGNIFICANT CASES,PERCENTAGES, & PRINT SUMMARY
      ANC=NC
      DO 308 L=1,6,2
      M=L/2+1
      LSIG=NSIG(M)
      MSIG=NC-LSIG
      PRCT(L)=FLOAT(MSIG)*100/ANC
      PRCT(L+1)=FLOAT(LSIG)*100/ANC
 308  CONTINUE
      WRITE(6,202)DATE,PRCT,NC
      STOP
      END
```

```
C ON100 ORAVECZ-NEWMAN HYPOTHESIS - PROPORTIONAL CASE
C WRITTEN 2-11-76  WD WHEATON
C DEFINITION - 2X2 CONTINGENCY TABLE
C  CELL(1,1) EQU CELL(1)
C  CELL(1,2) EQU CELL(2)
C  CELL(2,1) EQU CELL(3)
C  CELL(2,2) EQU CELL(4)
      REAL*4 P(4),SUM(4),SUMS(4),F(3),PF(3),PRCT(6)
      INTEGER*4 N(4),NSIG(3),DATE(4)
      EQUIVALENCE (SUM(1),SUM1),(SUM(2),SUM2),(SUM(3),SUM3),
     1(SUM(4),SUM4),(N(1),N1),(N(2),N2),(N(3),N3),(N(4),N4),
     2(F(1),FR),(F(2),FC),(F(3),FRC)
  200 FORMAT('1',35X,'ON100  ORAVECZ & NEWMAN PROPORTIONAL CELL SIZES  '
     1,4A2,T123,'PAGE',I3//
     2'            CELL SIZE RATIOS                           CELL POPULATION SIZE
     3              F VALUES                                  PROBABILITIES'/
     4' ---------------------------------   -------------------------
     5  ---------------------------------   -------------------------'/
     6'  (1,1)     (1,2)     (2,1)     (2,2)    (1,1)  (1,2)  (2,1)  (2,2)
     7       ROW          COL          TABLE        ROW      COL      TABLE'/
     8' -------  -------  -------  -------   -----  -----  -----  -----
     9  ----------  ----------  ----------   -------  -------  -------')
  201 FORMAT(F8.5,3F9.5,1X,4I7,1X,3F12.4,1X,3F9.5)
  202 FORMAT('1ON100  ORAVECZ & NEWMAN PROPORTIONAL CELL SIZES  ',4A2///
     111X,'PERCENT OF CASES'/
     210X,18('-')/
     312X,'SIG      NON-SIG'/
     410X,'-------    -------'/
     5' ROW   ',2F11.2//
     6' COL   ',2F11.2//
     7' TABLE',2F11.2///' TOTAL CASES RUN =',I6)
C GET CURRENT DATE
      CALL TDATE(DATE)
C CONSTANT NEEDED LATER  D1=2./9.=0.222222
      D1=0.222222
C PROBABILITY EXPANSION CONSTANTS
      CC1=0.196854
      CC2=0.115194
      CC3=0.000344
      CC4=0.019527
C NUMBER OF CASES TO RUN
      NC=1000
C NUMBER OF NON-SIGNIFICANT CASES - ROW,COL & TABLE
      NSIG(1)=0
      NSIG(2)=0
      NSIG(3)=0
C PAGE & LINE CNTS
      LINE=100
      IPAGE=1
      IY=9
C LOOP OVER ALL CASES
      DO 306 K=1,NC
C GENERATE CELL SIZE  (.GT. 10)
  300 IX=IY
      CALL RANDU(IX,IY,PP)
      NN=PP*100.
      IF(NN.LT.10) GO TO 300
C CALC CELL SIZES
      NSUM=0
```

ROW FACTOR NOT NECESSARILY SIGNIFICANT

```
      DO 301 I=1,3
 302  IY=IY
      CALL RANDU(IX,IY,A)
      NN=A*100.
      IF(NN.LT.10) GO TO 302
      N(I)=NN
      NSUM=NSUM+NN
 301  CONTINUE
      N4=N2*N3/N1
C IF N4 IS LESS THAN 6 DROP CASE
      IF(N4.LT.6) GO TO 300
      NSUM=NSUM+N4
      SUMN=NSUM
      DO 309 I=1,4
      AN=N(I)
      P(I)=AN/SUMN
 309  CONTINUE
C GENERATE MEMBERS OF EACH CELL &.CELL SUM
 310  TSUMS=0.0
      DO 303 I=1,4
      SUM(I)=0.0
       SUMS(I)=0.0
      L=N(I)
      DO 304 J=1,L
      IX=IY
      CALL RANDU(IX,IY,A)
      SUM(I)=SUM(I)+A
      SUMS(I)=SUMS(I)+A*A
 304  CONTINUE
      TSUMS=TSUMS+SUMS(I)
 303  CONTINUE
C CALC ROW,COL,& TABLE SUMS
      R1S=SUM1+SUM2
      R2S=SUM3+SUM4
      C1S=SUM1+SUM3
      C2S=SUM2+SUM4
      TABSUM=R1S+R2S
C CALC ROW,COL,$ TABLE SUMS SQUARED
      TSS=TABSUM*TABSUM/FLOAT(NSUM)
      SSR=R1S*R1S/(N1+N2)+R2S*R2S/(N3+N4)-TSS
      SSC=C1S*C1S/(N1+N3)+C2S*C2S/(N2+N4)-TSS
      SST=TSUMS-TSS
      SSCELL=SUM1*SUM1/N1+SUM2*SUM2/N2+SUM3*SUM3/N3+SUM4*SUM4/N4-TSS
      SSRC=SSCELL-SSR-SSC
      SSERR=SST-SSR-SSC-SSRC
C CALC F VALUES FOR ROWS,COLS,& TABLE
      DF=NSUM-4
      SERRM=SSERR/DF
      FR=SSR/SERRM
      FC=SSC/SERRM
      FRC=SSRC/SERRM
C IF F(I) IS NEGATIVE DROP CASE FROM CONSIDERATION
      IF(FR.LT.0.0) GO TO 310
      IF(FC.LT.0.0) GO TO 310
      IF(FRC.LT.0.0) GO TO 310
C CALC PROBABILITIES FOR ROWS,COLS & TABLE
      D2=D1/DF
      DO 305 I=1,3
      X1=SQRT(F(I))*(1.-0.25/DF)/SQRT(1.+0.5*F(I)/DF)
```

```
      X2=1.0+X1*(CC1+X1*(CC2+X1*(CC3+X1*CC4)))
      X3=X2*X2*X2*X2
      PF(I)=1.-1./X3
C CNT NON-SIGNIFICANT CASES - ROW,COL & TABLE
      IF(PF(I).LE.0.95) NSIG(I)=NSIG(I)+1
  305 CONTINUE
C PRINT CASE PARAMETERS
      IF(LINE.LT.55) GO TO 307
      WRITE(6,200)DATE,IPAGE
      LINE=6
      IPAGE=IPAGE+1
  307 WRITE(6,201)P,N,F,PF
      LINE=LINE+1
  306 CONTINUE
C CALC SIGNIFICANT CASES,PERCENTAGES, & PRINT SUMMARY
      ANC=NC
      DO 308 L=1,6,2
      M=L/2+1
      LSIG=NSIG(M)
      MSIG=NC-LSIG
      PRCT(L)=FLOAT(MSIG)*100/ANC
      PRCT(L+1)=FLOAT(LSIG)*100/ANC
  308 CONTINUE
      WRITE(6,202)DATE,PRCT,NC
      CALL EXIT
      END
```

```
C ON100 ORAVECZ-NEWMAN HYPOTHESIS - PROPORTIONAL CASE
C WRITTEN 2-11-76  WD WHEATON          ROW FACTOR SIGNIFICANT
C DEFINITION - 2X2 CONTINGENCY TABLE
C  CELL(1,1) EQU CELL(1)
C  CELL(1,2) EQU CELL(2)
C  CELL(2,1) EQU CELL(3)
C  CELL(2,2) EQU CELL(4)
      REAL*4 P(4),SUM(4),SUMS(4),F(3),PF(3),PRCT(6)
      INTEGER*4 N(4),NSIG(3),DATE(2)
      EQUIVALENCE (SUM(1),SUM1),(SUM(2),SUM2),(SUM(3),SUM3),
     1(SUM(4),SUM4),(N(1),N1),(N(2),N2),(N(3),N3),(N(4),N4),
     2(F(1),FR),(F(2),FC),(F(3),FRC)
  200 FORMAT('1',35X,'ON100  ORAVECZ & NEWMAN PROPORTIONAL CELL SIZES  '
     1,4A2,T123,'PAGE',I3//
     2'           CELL SIZE RATIOS                        CELL POPULATION SIZE
     3                  F VALUES                              PROBABILITIES'/
     4' ------------------------------------------   --------------------------
     5  ------------------------------------------   --------------------------'/
     6'  (1,1)     (1,2)      (2,1)      (2,2)     (1,1)   (1,2)   (2,1)   (2,2)
     7      ROW           COL           TABLE       ROW      COL        TABLE'/
     8' -------   -------   -------   -------      -----   -----   -----   -----
     9  -----------   -----------   -----------    -------   -------   --------')
  201 FORMAT(F8.5,3F9.5,1X,4I7,1X,3E12.4,1X,3F9.5)
  202 FORMAT('1ON100  ORAVECZ & NEWMAN PROPORTIONAL CELL SIZES   ',2A4///
     111X,'PERCENT OF CASES'/
     210X,18('-')/
     312X,'SIG      NON-SIG'/
     410X,'-------   -------'/
     5' ROW   ',2F11.2//
     6' COL   ',2F11.2//
     7' TABLE',2F11.2///' TOTAL CASES RUN =',I6)
C GET CURRENT DATE
      CALL TDATE(1,DATE)
C CONSTANT NEEDED LATER  D1=2./9.=0.222222
      D1=0.222222
C PROBABILITY EXPANSION CONSTANTS
      CC1=0.196854
      CC2=0.115194
      CC3=0.000344
      CC4=0.019527
C NUMBER OF CASES TO RUN
      NC=1000
C NUMBER OF NON-SIGNIFICANT CASES - ROW,COL & TABLE
      NSIG(1)=0
      NSIG(2)=0
      NSIG(3)=0
C PAGE & LINE CNTS
      LINE=100
      IPAGE=1
      IY=9
C LOOP OVER ALL CASES
      DO 306 K=1,NC
C GENERATE CELL SIZE  (.GT. 10)
  300 IX=IY
      CALL RANDU(IX,IY,PP)
      NN=PP*100.
      IF(NN.LT.10) GO TO 300
C CALC CELL SIZES
      NSUM=0
```

```
      DO 301 I=1,3
  302 IX=IY
      CALL RANDU(IX,IY,A)
      NN=A*100.
      IF(NN.LT.10) GO TO 302
      N(I)=NN
      NSUM=NSUM+NN
  301 CONTINUE
      N4=N2*N3/N1
C IF N4 IS LESS THAN 6 DROP CASE
      IF(N4.LT.6) GO TO 300
      NSUM=NSUM+N4
      SUMN=NSUM
      DO 309 I=1,4
      AN=N(I)
      P(I)=AN/SUMN
  309 CONTINUE
C GENERATE MEMBERS OF EACH CELL & CELL SUM
  310 TSUMS=0.0
      DO 303 I=1,4
      SUM(I)=0.0
       SUMS(I)=0.0
      L=N(I)
      DO 304 J=1,L
      IX=IY
      CALL RANDU(IX,IY,A)
      IF(I.LE.2) A=A+5
      SUM(I)=SUM(I)+A
      SUMS(I)=SUMS(I)+A*A
  304 CONTINUE
      TSUMS=TSUMS+SUMS(I)
  303 CONTINUE
C CALC ROW,COL,& TABLE SUMS
C   CALC ROW,COL,& TABLE SUMS SO ROW FACTOR SIGNIF
      R1S=SUM1+SUM2
      R2S=SUM3+SUM4
      C1S=SUM1+SUM3
      C2S=SUM2+SUM4
      TABSUM=R1S+R2S
C CALC ROW,COL,$ TABLE SUMS SQUARED
      TSS=TABSUM*TABSUM/FLOAT(NSUM)
      SSR=R1S*R1S/(N1+N2)+R2S*R2S/(N3+N4)-TSS
      SSC=C1S*C1S/(N1+N3)+C2S*C2S/(N2+N4)-TSS
      SST=TSUMS-TSS
      SSCELL=SUM1*SUM1/N1+SUM2*SUM2/N2+SUM3*SUM3/N3+SUM4*SUM4/N4-TSS
      SSRC=SSCELL-SSR-SSC
      SSERR=SST-SSR-SSC-SSRC
C CALC F VALUES FOR ROWS,COLS,& TABLE
      DF=NSUM-4
      SERRM=SSERR/DF
      FR=SSR/SERRM
      FC=SSC/SERRM
      FRC=SSRC/SERRM
C IF F(I) IS NEGATIVE DROP CASE FRCM CONSIDERATION
      IF(FR.LT.0.0) GO TO 310
      IF(FC.LT.0.0) GO TO 310
      IF(FRC.LT.0.0) GO TO 310
C CALC PROBABILITIES FOR ROWS,COLS & TABLE
      D2=D1/DF
```

```
      DO 305 I=1,3
      X1=SQRT(F(I))*(1.-0.25/DF)/SQRT(1.+0.5*F(I)/DF)
      X2=1.0+X1*(CC1+X1*(CC2+X1*(CC3+X1*CC4)))
      X3=X2*X2*X2*X2
      PF(I)=1.-1./X3
C CNT NON-SIGNIFICANT CASES - ROW,COL & TABLE
      IF(PF(I).LE.0.95) NSIG(I)=NSIG(I)+1
  305 CONTINUE
C PRINT CASE PARAMETERS
      IF(LINE.LE.56) GO TO 307
      WRITE(6,200)DATE,IPAGE
      LINE=6
      IPAGE=IPAGE+1
  307 WRITE(6,201)P,N,F,PF
      LINE=LINE+1
  306 CONTINUE
C CALC SIGNIFICANT CASES,PERCENTAGES, & PRINT SUMMARY
      ANC=NC
      DO 308 L=1,6,2
      M=L/2+1
      LSIG=NSIG(M)
      MSIG=NC-LSIG
      PRCT(L)=FLOAT(MSIG)*100/ANC
      PRCT(L+1)=FLOAT(LSIG)*100/ANC
  308 CONTINUE
      WRITE(6,202)DATE,PRCT,NC
      STOP
      END
```

```
C ON120 ORAVECZ-NEWMAN HYPOTHESIS - NONSIGNIF DISPROP CASE
C MODIFIED FROM ON100  03-09-76  WD WHEATON    ROW FACTOR NOT
C DEFINITION - 2X2 CONTINGENCY TABLE           NECESSARILY SIGNIFICANT
C  CELL(1,1) EQU CELL(1)
C  CELL(1,2) EQU CELL(2)
C  CELL(2,1) EQU CELL(3)
C  CELL(2,2) EQU CELL(4)
      REAL*4 P(4),SUM(4),SUMS(4),F(3),PF(3),FSIG(2,6)
C DEFN - MATRIX NSIG(2,3) CONTAINS COUNT OF NON-SIG F'S
C    COL 1 CHI SQ'D SIG
C    COL 2 CHI SQ'D NON-SIG
C    ROW 1  ROW OF TABLE
C    ROW 2  COL OF TABLE
C    ROW 3  TABLE
      INTEGER*4 N(4),SIG,NSIG(2,3),SCHI,DATE(2)
      EQUIVALENCE (SUM(1),SUM1),(SUM(2),SUM2),(SUM(3),SUM3),
     1(SUM(4),SUM4),(N(1),N1),(N(2),N2),(N(3),N3),(N(4),N4),
     2(F(1),FR),(F(2),FC),(F(3),FRC)
      DATA NSIG/6*0/,LTRN,LTRS/'N','S'/
  200 FORMAT('1',35X,'ON120  ORAVECZ & NEWMAN DISPROPORTIONAL CELL SIZES
     1 ',4A2,T123,'PAGE',I3//
     2'          CELL SIZE RATIOS                      CELL POPULATION SIZE
     3            F VALUES                                 PROBABILITIES'/
     4' -----------------------------------   -------------------------
     5 -----------------------------------   -------------------------'/
     6'  (1,1)      (1,2)      (2,1)      (2,2)     (1,1)  (1,2)  (2,1)  (2,2)
     7    ROW         COL           TABLE         ROW     COL      TABLE'/
     8' -------   -------   -------   -------   -----  -----  -----  -----
     9 ---------   ---------   ---------   -------   -------   -------')
  201 FORMAT(F8.5,3F9.5,1X,4I7,1X,3E12.4,1X,3F9.5,1X,A1)
  202 FORMAT('1ON120  ORAVECZ & NEWMAN DISPROPORTIONAL CELL SIZES  ',2A4
     1///' CHI SQ''D NON-SIG   F TEST RESULTS (PERCENTS)'//
     210X,'NON-SIG        SIG  '/
     310X,'-------      -------'/
     4' ROW   ',2F11.2//
     5' COL   ',2F11.2//
     6' TABLE',2F11.2//
     7///' CHI SQ''D SIG  F TEST RESULTS (PERCENTS)'//
     810X,'NON-SIG        SIG  '/
     910X,'-------      -------'/
     A' ROW   ',2F11.2//
     B' COL   ',2F11.2//
     C' TABLE',2F11.2//
     D///' CASES CONSIDERED FOR EACH STATE CHI SQ''D =',I4)
C FIX TOTAL NUMBER OF CASES
      NC=1000
C GET CURRENT DATE
      CALL TDATE(1,DATE)
C CONSTANT NEEDED LATER  D1=2./9.=0.222222
      D1=0.222222
C PROBABILITY EXPANSION CONSTANTS
      CC1=0.196854
      CC2=0.115194
      CC3=0.000344
      CC4=0.019527
C NUMBER SIG & NON-SIG CASES CHI SQ'D
      NSCHI=0
      SCHI=0
C PAGE & LINE CNTS
```

```
      LINE=100
      IPAGE=1
      IY=9
C GENERATE CELL SIZES (BETWEEN 10 AND 1000)
  306 NSUM=0
      DO 300 I=1,3
  301 IX=IY
      CALL RANDU(IX,IY,A)
      IF(A.LT.0.10) GO TO 301
      N(I)=A*100.
      NSUM=NSUM+N(I)
  300 CONTINUE
C GENERATE N4 SO DISPROPORTIONAL
  302 IX=IY
      CALL RANDU(IX,IY,A)
      NA=A*10.+1.
      N4=N2*N3/N1+NA
      NSUM=NSUM+N4
      SUMN=NSUM
C CALC CHI SQUARED
      CHI2A=N1*N4-N2*N3
      CHI2=SUMN*CHI2A*CHI2A/((N1+N2)*(N3+N4)*(N1+N3)*(N2+N4))
C TEST SIGNIFICANCE OF CHI SQ'D
      IF(CHI2.GT.1.320) GO TO 306
C CHI SQ NON-SIG TEST FOR NC CASES
      IF(NSCHI.EQ.NC) GO TO 312
C NOT ENOUGH  ADD THIS ONE
      LS=2
      NSCHI=NSCHI+1
      SIG=LTRN
C GENERATE MEMBERS OF EACH CELL & CELL SUM
  313 TSUMS=0.0
      DO 303 I=1,4
      SUM(I)=0.0
      SUMS(I)=0.0
      P(I)=FLOAT(N(I))/SUMN
      L=N(I)
      DO 304 J=1,L
      IX=IY
      CALL RANDU(IX,IY,A)
      SUM(I)=SUM(I)+A
      SUMS(I)=SUMS(I)+A*A
  304 CONTINUE
      TSUMS=TSUMS+SUMS(I)
  303 CONTINUE
C CALC F STATISTICS
C CALC ROW,COL,& TABLE SUMS
C    CALC ROW,COL,& TABLE SUMS SO ROW FACTOR SIGNIF
      R1S=SUM1+SUM2
      R2S=SUM3+SUM4
      C1S=SUM1+SUM3
      C2S=SUM2+SUM4
      TABSUM=R1S+R2S
C CALC ROW,COL,$ TABLE SUMS SQUARED
      TSS=TABSUM*TABSUM/FLOAT(NSUM)
      SSR=R1S*R1S/(N1+N2)+R2S*R2S/(N3+N4)-TSS
      SSC=C1S*C1S/(N1+N3)+C2S*C2S/(N2+N4)-TSS
      SST=TSUMS-TSS
      SSCELL=SUM1*SUM1/N1+SUM2*SUM2/N2+SUM3*SUM3/N3+SUM4*SUM4/N4-TSS
```

```
      SSRC=SSCELL-SSR-SSC
      SSERR=SST-SSR-SSC-SSRC
C CALC F VALUES FOR ROWS,COLS,& TABLE
      DF=NSUM-4
      SERRM=SSERR/DF
      FR=SSR/SERRM
      FC=SSC/SERRM
      FRC=SSRC/SERRM
  500 CONTINUE
C IF F(I) IS NEGATIVE DROP CASE FROM CONSIDERATION
      IF(FR.LT.0.0) GO TO 316
      IF(FC.LT.0.0) GO TO 316
      IF(FRC.LT.0.0) GO TO 316
C CALC PROBABILITIES FOR ROWS,COLS & TABLE
      D2=D1/DF
      DO 305 I=1,3
      X1=SQRT(F(I))*(1.-0.25/DF)/SQRT(1.+0.5*F(I)/DF)
      X2=1.0+X1*(CC1+X1*(CC2+X1*(CC3+X1*CC4)))
      X3=X2*X2*X2*X2
      PF(I)=1.-1./X3
C CNT NON-SIG CASES OF F
      IF(PF(I).LE.0.95) NSIG(LS,I)=NSIG(LS,I)+1
  305 CONTINUE
C PRINT CASE PARAMETERS
      IF(LINE.LE.56) GO TO 307
      WRITE(3,200)DATE,IPAGE
      LINE=6
      IPAGE=IPAGE+1
  307 WRITE(3,201)P,N,F,PF,SIG
      LINE=LINE+1
C DO WE HAVE ENOUGH CASES OF SIG AND NON-SIG CHI YET
X 312 IF(NSCHI.LT.NC) GO TO 306
C CALC & PRINT SUMMARY
C MATRIX FSIG(2,6) CONTAINS THE PERCENTS AS FOLLOWS
C     COL 1   PERCENT NON-SIG F'S
C     COL 2   PERCENT SIG F'S
C     ROW 1   ROW F NON-SIG CHI SQ'D
C     ROW 2   COL F NON-SIG CHI SQ'D
C     ROW 3 TABL F NON-SIG CHI SQ'D
C     ROW 4   ROW SIG CHI SQ'D
C     ROW 5   COL SIG CHI SQ'D
C     ROW 6   TAB SIG CHI SQ'D
  315 CN=NC
      K=2
      DO 314 I=1,6
      L=I-(I-1)/3*3
      IF(I.EQ.4) K=1
      TMP=FLOAT(NSIG(K,L))
      FSIG(1,I)=TMP/CN*100
      FSIG(2,I)=(CN-TMP)/CN*100.
  314 CONTINUE
      WRITE(3,202)DATE,FSIG,NC
      STOP
C F(I) IS NEG DROP CASE
  316 IF(SIG.EQ.LTRS) SCHI=SCHI-1
      IF(SIG.EQ.LTRN) NSCHI=NSCHI-1
      GO TO 306
      END
```

```
      C ON120 ORAVECZ-NEWMAN HYPOTHESIS - NONSIGNIF DISPROP CASE
      C MODIFIED FROM ON100  03-09-76  WD WHEATON   Row FACTOR SIGNIFICANT
      C DEFINITION - 2X2 CONTINGENCY TABLE
      C  CELL(1,1) EQU CELL(1)
      C  CELL(1,2) EQU CELL(2)
      C  CELL(2,1) EQU CELL(3)
      C  CELL(2,2) EQU CELL(4)
            REAL*4 P(4),SUM(4),SUMS(4),F(3),PF(3),FSIG(2,6)
      C DEFN - MATRIX NSIG(2,3) CONTAINS COUNT OF NON-SIG F'S
      C    COL 1 CHI SQ'D SIG
      C    COL 2 CHI SQ'D NON-SIG
      C    ROW 1  ROW OF TABLE
      C    ROW 2  COL OF TABLE
      C    ROW 3  TABLE
            INTEGER*4 N(4),SIG,NSIG(2,3),SCHI,DATE(2)
            EQUIVALENCE (SUM(1),SUM1),(SUM(2),SUM2),(SUM(3),SUM3),
           1(SUM(4),SUM4),(N(1),N1),(N(2),N2),(N(3),N3),(N(4),N4),
           2(F(1),FR),(F(2),FC),(F(3),FRC)
            DATA NSIG/6*0/,LTRN,LTRS/'N','S'/
        200 FORMAT('1',35X,'ON120  ORAVECZ & NEWMAN DISPROPORTIONAL CELL SIZES
           1 ',4A2,T123,'PAGE',I3//
           2'               CELL SIZE RATIOS                     CELL POPULATION SIZE
           3                    F VALUES                          PROBABILITIES'/
           4' -----------------------------------   ------------------------
           5 -----------------------------------   ------------------------'/
           6'  (1,1)     (1,2)     (2,1)     (2,2)     (1,1)  (1,2)  (2,1)  (2,2)
           7     ROW             COL           TABLE      ROW      COL.     TABLE'/
           8' -------  -------  -------  -------   -----  -----  -----  -----
           9 -------  -------  -------  -------   -------  -------  -------')
        201 FORMAT(F8.5,3F9.5,1X,4I7,1X,3E12.4,1X,3F9.5,1X,A1)
        202 FORMAT('1ON120  ORAVECZ & NEWMAN DISPROPORTIONAL CELL SIZES  ',2A4
           1///' CHI SQ''D NON-SIG  F TEST RESULTS (PERCENTS)'//
           210X,'NON-SIG       SIG  '/
           310X,'-------      -------'/
           4' ROW   ',2F11.2//
           5' COL   ',2F11.2//
           6' TABLE',2F11.2//
           7///' CHI SQ''D SIG  F TEST RESULTS (PERCENTS)'//
           810X,'NON-SIG       SIG  '/
           910X,'-------      -------'/
           A' ROW   ',2F11.2//
           B' COL   ',2F11.2//
           C' TABLE',2F11.2//
           D///' CASES CONSIDERED FOR EACH STATE CHI SQ''D =',I4)
      C FIX TOTAL NUMBER OF CASES
            NC=1000
      C GET CURRENT DATE
            CALL TDATE(1,DATE)
      C CONSTANT NEEDED LATER  D1=2./9.=0.222222
            D1=0.222222
      C PROBABILITY EXPANSION CONSTANTS
            CC1=0.196854
            CC2=0.115194
            CC3=0.000344
            CC4=0.019527
      C NUMBER SIG & NON-SIG CASES CHI SQ'D
            NSCHI=0
            SCHI=0
      C PAGE & LINE CNTS
```

```
          LINE=100
          IPAGE=1
          IY=9
    C GENERATE CELL SIZES (BETWEEN 10 AND 1000)
      306 NSUM=0
          DO 300 I=1,3
      301 IX=IY
          CALL RANDU(IX,IY,A)
          IF(A.LT.0.10) GO TO 301
          N(I)=A*100.
          NSUM=NSUM+N(I)
      300 CONTINUE
  X C GENERATE N4 SO DISPROPORTIONAL
  X   302 IX=IY
        X CALL RANDU(IX,IY,A)
        X NA=A*10.+1.
        X N4=N2*N3/N1+NA
        X NSUM=NSUM+N4
          SUMN=NSUM
    C CALC CHI SQUARED
          CHI2A=N1*N4-N2*N3
          CHI2=SUMN*CHI2A*CHI2A/((N1+N2)*(N3+N4)*(N1+N3)*(N2+N4))
    C TEST SIGNIFICANCE OF CHI SQ'D
          IF(CHI2.GT.1.320) GO TO 306
    C CHI SQ NON-SIG TEST FOR NC CASES
          IF(NSCHI.EQ.NC) GO TO 312
    C NOT ENOUGH  ADD THIS ONE
          LS=2
          NSCHI=NSCHI+1
          SIG=LTRN
    C GENERATE MEMBERS OF EACH CELL & CELL SUM
      313 TSUMS=0.0
          DO 303 I=1,4
          SUM(I)=0.0
           SUMS(I)=0.0
          P(I)=FLOAT(N(I))/SUMN
          L=N(I)
          DO 304 J=1,L
          IX=IY
          CALL RANDU(IX,IY,A)
    C FOR CELLS 1 AND 2 ADD FIVE TO EACH DATA POINT
          IF(I.LE.2) A=A+5
          SUM(I)=SUM(I)+A
          SUMS(I)=SUMS(I)+A*A
      304 CONTINUE
          TSUMS=TSUMS+SUMS(I)
      303 CONTINUE
    C CALC F STATISTICS
    C CALC ROW,COL,& TABLE SUMS
    C    CALC ROW,COL,& TABLE SUMS SO ROW FACTOR SIGNIF
          R1S=SUM1+SUM2
          R2S=SUM3+SUM4
          C1S=SUM1+SUM3
          C2S=SUM2+SUM4
          TABSUM=R1S+R2S
    C CALC ROW,COL,$ TABLE SUMS SQUARED
          TSS=TABSUM*TABSUM/FLOAT(NSUM)
          SSR=R1S*R1S/(N1+N2)+R2S*R2S/(N3+N4)-TSS
          SSC=C1S*C1S/(N1+N3)+C2S*C2S/(N2+N4)-TSS
```

```
      SST=TSUMS-TSS
      SSCELL=SUM1*SUM1/N1+SUM2*SUM2/N2+SUM3*SUM3/N3+SUM4*SUM4/N4-TSS
      SSRC=SSCELL-SSR-SSC
      SSERR=SST-SSR-SSC-SSRC
C CALC F VALUES FOR ROWS,COLS,& TABLE
      DF=NSUM-4
      SERRM=SSERR/DF
      FR=SSR/SERRM
      FC=SSC/SERRM
      FRC=SSRC/SERRM
  500 CONTINUE
C IF F(I) IS NEGATIVE DROP CASE FROM CONSIDERATION
      IF(FR.LT.0.0) GO TO 316
      IF(FC.LT.0.0) GO TO 316
      IF(FRC.LT.0.0) GO TO 316
C CALC PROBABILITIES FOR ROWS,COLS & TABLE
      D2=D1/DF
      DO 305 I=1,3
      X1=SQRT(F(I))*(1.-0.25/DF)/SQRT(1.+0.5*F(I)/DF)
      X2=1.0+X1*(CC1+X1*(CC2+X1*(CC3+X1*CC4)))
      X3=X2*X2*X2*X2
      PF(I)=1.-1./X3
C CNT NON-SIG CASES OF F
      IF(PF(I).LE.0.95) NSIG(LS,I)=NSIG(LS,I)+1
  305 CONTINUE
C PRINT CASE PARAMETERS
      IF(LINE.LE.56) GO TO 307
      WRITE(3,200)DATE,IPAGE
      LINE=6
      IPAGE=IPAGE+1
  307 WRITE(3,201)P,N,F,PF,SIG
      LINE=LINE+1
C DO WE HAVE ENOUGH CASES OF SIG AND NON-SIG CHI YET
  312 IF(NSCHI.LT.NC) GO TO 306
C CALC & PRINT SUMMARY
C MATRIX FSIG(2,6) CONTAINS THE PERCENTS AS FOLLOWS
C    COL 1  PERCENT NON-SIG F'S
C    COL 2  PERCENT SIG F'S
C    ROW 1  ROW F NON-SIG CHI SQ'D
C    ROW 2  COL F NON-SIG CHI SQ'D
C    ROW 3 TABL F NON-SIG CHI SQ'D
C    ROW 4  ROW SIG CHI SQ'D
C    ROW 5  COL SIG CHI SQ'D
C    ROW 6  TAB SIG CHI SQ'D
  315 CN=NC
      K=2
      DO 314 I=1,6
      L=I-(I-1)/3*3
      IF(I.EQ.4) K=1
      TMP=FLOAT(NSIG(K,L))
      FSIG(1,I)=TMP/CN*100.
      FSIG(2,I)=(CN-TMP)/CN*100.
  314 CONTINUE
      WRITE(3,202)DATE,FSIG,NC
      STOP
C F(I) IS NEG DROP CASE
  316 IF(SIG.EQ.LTRS) SCHI=SCHI-1
      IF(SIG.EQ.LTRN) NSCHI=NSCHI-1
      GO TO 306
      END
```

ROW FACTOR NOT NECCESSARILY SIGNIFICANT

```
X C ON120 ORAVECZ-NEWMAN HYPOTHESIS - SIGNIF DISPROPORTIONAL CASE
  C MODIFIED FROM ON100  03-09-76  WD WHEATON
  C DEFINITION - 2X2 CONTINGENCY TABLE
  C  CELL(1,1) EQU CELL(1)
  C  CELL(1,2) EQU CELL(2)
  C  CELL(2,1) EQU CELL(3)
  C  CELL(2,2) EQU CELL(4)
        REAL*4 P(4),SUM(4),SUMS(4),F(3),PF(3),FSIG(2,6)
  C DEFN - MATRIX NSIG(2,3) CONTAINS COUNT OF NON-SIG F'S
  C   COL 1 CHI SQ'D SIG
  C   COL 2 CHI SQ'D NON-SIG
  C   ROW 1  ROW OF TABLE
  C   ROW 2  COL OF TABLE
  C   ROW 3  TABLE
        INTEGER*4 N(4),SIG,NSIG(2,3),SCHI,DATE(2)
        EQUIVALENCE (SUM(1),SUM1),(SUM(2),SUM2),(SUM(3),SUM3),
       1(SUM(4),SUM4),(N(1),N1),(N(2),N2),(N(3),N3),(N(4),N4),
       2(F(1),FR),(F(2),FC),(F(3),FRC)
        DATA NSIG/6*0/,LTRN,LTRS/'N','S'/
  200   FORMAT('1',35X,'ON120  ORAVECZ & NEWMAN DISPROPORTIONAL CELL SIZES
       1  ',4A2,T123,'PAGE',I3//
       2'            CELL SIZE RATIOS                      CELL POPULATION SIZE
       3                 F VALUES                            PROBABILITIES'/
       4' -------------------------------------    ------------------------------
       5  -------------------------------------    ------------------------------'/
       6'  (1,1)     (1,2)     (2,1)      (2,2)     (1,1)   (1,2)    (2,1)   (2,2)
       7      ROW            COL          TABLE       ROW      COL       TABLE'/
       8'  -------  -------  -------    ---------    ------  -------  ------  ------
       9  ---------  ---------   -------  ---------   ------  --------  ----------')
  201   FORMAT(F8.5,3F9.5,1X,4I7,1X,3E12.4,1X,3F9.5,1X,A1)
  202   FORMAT('1ON120  ORAVECZ & NEWMAN DISPROPORTIONAL CELL SIZES  ',2A4
       1///' CHI SQ''D NON-SIG  F TEST RESULTS (PERCENTS)'//
       210X,'NON-SIG        SIG  '/
       310X,'-------      -------'/
       4' ROW   ',2F11.2//
       5' COL   ',2F11.2//
       6' TABLE',2F11.2//
       7///' CHI SQ''D SIG  F TEST RESULTS (PERCENTS)'//
       810X,'NON-SIG        SIG  '/
       910X,'-------      -------'/
       A' ROW   ',2F11.2//
       B' COL   ',2F11.2//
       C' TABLE',2F11.2//
       D///' CASES CONSIDERED FOR EACH STATE CHI SQ''D =',I4)
  C FIX TOTAL NUMBER OF CASES
        NC=1000
  C GET CURRENT DATE
        CALL TDATE(1,DATE)
  C CONSTANT NEEDED LATER  D1=2./9.=0.222222
        D1=0.222222
  C PROBABILITY EXPANSION CONSTANTS
        CC1=0.196854
        CC2=0.115194
        CC3=0.000344
        CC4=0.019527
  C NUMBER SIG & NON-SIG CASES CHI SQ'D
        NSCHI=0
        SCHI=0
  C PAGE & LINE CNTS
```

```
      LINE=100
      IPAGE=1
      IY=9
C GENERATE CELL SIZES (BETWEEN 10 AND 1000)
  306 NSUM=0
X     DO 300 I=1,4
  301 IX=IY
      CALL RANDU(IX,IY,A)
      IF(A.LT.0.10) GO TO 301
      N(I)=A*100.
      NSUM=NSUM+N(I)
  300 CONTINUE
      SUMN=NSUM
C CALC CHI SQUARED
      CHI2A=N1*N4-N2*N3
      CHI2=SUMN*CHI2A*CHI2A/((N1+N2)*(N3+N4)*(N1+N3)*(N2+N4))
C TEST SIGNIFICANCE OF CHI SQ'D
X     IF(CHI2.LE.3.841) GO TO 306
XC CHI SQ'D SIG  TEST FOR NC CASES
X     IF(SCHI.EQ.NC) GO TO 312
XC NOT ENOUGH  ADD THIS ONE
X     LS=1
X     SCHI=SCHI+1
X     SIG=LTRS
C GENERATE MEMBERS OF EACH CELL & CELL SUM
  313 TSUMS=0.0
      DO 303 I=1,4
      SUM(I)=0.0
       SUMS(I)=0.0
      P(I)=FLOAT(N(I))/SUMN
      L=N(I)
      DO 304 J=1,L
      IX=IY
      CALL RANDU(IX,IY,A)
      SUM(I)=SUM(I)+A
      SUMS(I)=SUMS(I)+A*A
  304 CONTINUE
      TSUMS=TSUMS+SUMS(I)
  303 CONTINUE
C CALC F STATISTICS
C CALC ROW,COL,& TABLE SUMS
C    CALC ROW,COL,& TABLE SUMS SO ROW FACTOR SIGNIF
      R1S=SUM1+SUM2
      R2S=SUM3+SUM4
      C1S=SUM1+SUM3
      C2S=SUM2+SUM4
      TABSUM=R1S+R2S
C CALC ROW,COL,$ TABLE SUMS SQUARED
      TSS=TABSUM*TABSUM/FLOAT(NSUM)
      SSR=R1S*R1S/(N1+N2)+R2S*R2S/(N3+N4)-TSS
      SSC=C1S*C1S/(N1+N3)+C2S*C2S/(N2+N4)-TSS
      SST=TSUMS-TSS
      SSCELL=SUM1*SUM1/N1+SUM2*SUM2/N2+SUM3*SUM3/N3+SUM4*SUM4/N4-TSS
      SSRC=SSCELL-SSR-SSC
      SSERR=SST-SSR-SSC-SSRC
C CALC F VALUES FOR ROWS,COLS,& TABLE
      DF=NSUM-4
      SERRM=SSERR/DF
      FR=SSR/SERRM
```

```
      FC=SSC/SERRM
      FRC=SSRC/SERRM
  500 CONTINUE
C IF F(I) IS NEGATIVE DROP CASE FROM CONSIDERATION
      IF(FR.LT.0.0) GO TO 316
      IF(FC.LT.0.0) GO TO 316
      IF(FRC.LT.0.0) GO TO 316
C CALC PROBABILITIES FOR ROWS,COLS & TABLE
      D2=D1/DF
      DO 305 I=1,3
      X1=SQRT(F(I))*(1.-0.25/DF)/SQRT(1.+0.5*F(I)/DF)
      X2=1.0+X1*(CC1+X1*(CC2+X1*(CC3+X1*CC4)))
      X3=X2*X2*X2*X2
      PF(I)=1.-1./X3
C CNT NON-SIG CASES OF F
      IF(PF(I).LE.0.95) NSIG(LS,I)=NSIG(LS,I)+1
  305 CONTINUE
C PRINT CASE PARAMETERS
      IF(LINE.LE.56) GO TO 307
      WRITE(3,200)DATE,IPAGE
      LINE=6
      IPAGE=IPAGE+1
  307 WRITE(3,201)P,N,F,PF,SIG
      LINE=LINE+1
C DO WE HAVE ENOUGH CASES OF SIG AND NON-SIG CHI YET
X312  IF(SCHI.LT.NC) GO TO 306
C CALC & PRINT SUMMARY
C MATRIX FSIG(2,6) CONTAINS THE PERCENTS AS FOLLOWS
C    COL 1  PERCENT NON-SIG F'S
C    COL 2  PERCENT SIG F'S
C    ROW 1  ROW F NON-SIG CHI SQ'D
C    ROW 2  COL F NON-SIG CHI SQ'D
C    ROW 3 TABL F NON-SIG CHI SQ'D
C    ROW 4  ROW SIG CHI SQ'D
C    ROW 5  COL SIG CHI SQ'D
C    ROW 6  TAB SIG CHI SQ'D
  315 CN=NC
      K=2
      DO 314 I=1,6
      L=I-(I-1)/3*3
      IF(I.EQ.4) K=1
      TMP=FLOAT(NSIG(K,L))
      FSIG(1,I)=TMP/CN*100.
      FSIG(2,I)=(CN-TMP)/CN*100.
  314 CONTINUE
      WRITE(3,202)DATE,FSIG,NC
      STOP
C F(I) IS NEG DROP CASE
  316 IF(SIG.EQ.LTRS) SCHI=SCHI-1
      IF(SIG.EQ.LTRN) NSCHI=NSCHI-1
      GO TO 306
      END
```

```
C CN120 ORAVECZ-NEWMAN HYPOTHESIS - SIGNIF DISPROPORTIONAL CASE
C MODIFIED FROM CN100  03-09-76  WD WHEATON     ROW FACTOR SIGNIFICANT
C DEFINITION - 2X2 CONTINGENCY TABLE
C  CELL(1,1) EQU CELL(1)
C  CELL(1,2) EQU CELL(2)
C  CELL(2,1) EQU CELL(3)
C  CELL(2,2) EQU CELL(4)
      REAL*4 P(4),SUM(4),SUMS(4),F(3),PF(3),FSIG(2,6)
C DEFN - MATRIX NSIG(2,3) CONTAINS COUNT OF NON-SIG F'S
C   COL 1 CHI SQ'D SIG
C   COL 2 CHI SQ'D NON-SIG
C   ROW 1  ROW OF TABLE
C   ROW 2  COL OF TABLE
C   ROW 3  TABLE
      INTEGER*4 N(4),SIG,NSIG(2,3),SCHI,DATE(2)
      EQUIVALENCE (SUM(1),SUM1),(SUM(2),SUM2),(SUM(3),SUM3),
     1(SUM(4),SUM4),(N(1),N1),(N(2),N2),(N(3),N3),(N(4),N4),
     2(F(1),FR),(F(2),FC),(F(3),FRC)
      DATA NSIG/6*0/,LTRN,LTRS/'N','S'/
  200 FORMAT('1',35X,'CN120  ORAVECZ & NEWMAN DISPROPORTIONAL CELL SIZES
     1 ',4A2,T123,'PAGE',I3//
     2'              CELL SIZE RATIOS                  CELL POPULATION SIZE
     3                  F VALUES                          PROBABILITIES'/
     4' -------------------------------------   ---------------------------
     5 -------------------------------------   ---------------------------'/
     6'  (1,1)     (1,2)     (2,1)     (2,2)    (1,1)  (1,2)  (2,1)  (2,2)
     7     ROW           COL          TABLE       ROW     COL      TABLE'/
     8' -------  -------  -------  -------   -----  -----  -----  -----
     9 ----------  ----------  ----------   -------  -------  -------')
  201 FORMAT(F8.5,3F9.5,1X,4I7,1X,3E12.4,1X,3F9.5,1X,A1)
  202 FORMAT('1CN120  ORAVECZ & NEWMAN DISPROPORTIONAL CELL SIZES  ',2A4
     1///' CHI SQ''D NON-SIG   F TEST RESULTS (PERCENTS)'//
     210X,'NON-SIG        SIG  '/
     310X,'-------     -------'/
     4' ROW   ',2F11.2//
     5' COL   ',2F11.2//
     6' TABLE',2F11.2//
     7///' CHI SQ''D SIG   F TEST RESULTS (PERCENTS)'//
     810X,'NON-SIG        SIG  '/
     910X,'-------     -------'/
     A' ROW   ',2F11.2//
     B' COL   ',2F11.2//
     C' TABLE',2F11.2//
     D///' CASES CONSIDERED FOR EACH STATE CHI SQ''D =',I4)
C FIX TOTAL NUMBER OF CASES
      NC=2
C FIX NUMBER OF CASES IN DEPLC
      ICNT=0
C GET CURRENT DATE
      CALL TDATE(1,DATE)
C CONSTANT NEEDED LATER  D1=2./9.=0.222222
      D1=0.222222
C PROBABILITY EXPANSION CONSTANTS
      CC1=0.196854
      CC2=0.115194
      CC3=0.000344
      CC4=0.019527
C NUMBER SIG & NON-SIG CASES CHI SQ'D
      NSCHI=0
```

```
      SCHI=0
C PAGE & LINE CNTS
      LINE=100
      IPAGE=1
      IY=9
C GENERATE CELL SIZES (BETWEEN 10 AND 1000)
  306 NSUM=0
      DO 300 I=1,4
  301 IX=IY
      CALL RANDU(IX,IY,A)
      IF(A.LT.0.10) GO TO 301
      N(I)=A*100.
      NSUM=NSUM+N(I)
  300 CONTINUE
      SUMN=NSUM
C CALC CHI SQUARED
      CHI2A=N1*N4-N2*N3
      CHI2=SUMN*CHI2A*CHI2A/((N1+N2)*(N3+N4)*(N1+N3)*(N2+N4))
C TEST SIGNIFICANCE OF CHI SQ'D
      IF(CHI2.LE.3.84)) GO TO 306
C CHI SQ'D SIG  TEST FOR NC CASES
      IF(SCHI.EQ.NC) GO TO 312
C NOT ENOUGH  ADD THIS ONE
      LS=1
      SCHI=SCHI+1
      SIG=LTRS
C GENERATE MEMBERS OF EACH CELL & CELL SUM
  313 TSUMS=0.0
      DO 303 I=1,4
      SUM(I)=0.0
       SUMS(I)=0.0
      P(I)=FLOAT(N(I))/SUMN
      L=N(I)
      DO 304 J=1,L
      IX=IY
      CALL RANDU(IX,IY,A)
C FOR CELLS 1 AND 2 ADD FIVE TO EACH DATA POINT
      IF(I.LE.2) A=A+5
      SUM(I)=SUM(I)+A
      SUMS(I)=SUMS(I)+A*A
  304 CONTINUE
      TSUMS=TSUMS+SUMS(I)
  303 CONTINUE
C CALC F STATISTICS
C CALC ROW,COL,& TABLE SUMS
C    CALC ROW,COL,& TABLE SUMS SO ROW FACTOR SIGNIF
      R1S=SUM1+SUM2
      R2S=SUM3+SUM4
      C1S=SUM1+SUM3
      C2S=SUM2+SUM4
      TABSUM=R1S+R2S
C CALC ROW,COL,$ TABLE SUMS SQUARED
      TSS=TABSUM*TABSUM/FLOAT(NSUM)
      SSR=R1S*R1S/(N1+N2)+R2S*R2S/(N3+N4)-TSS
      SSC=C1S*C1S/(N1+N3)+C2S*C2S/(N2+N4)-TSS
      SST=TSUMS-TSS
      SSCELL=SUM1*SUM1/N1+SUM2*SUM2/N2+SUM3*SUM3/N3+SUM4*SUM4/N4-TSS
      SSRC=SSCELL-SSR-SSC
      SSERR=SST-SSR-SSC-SSRC
```



```
C CALC F VALUES FOR ROWS,COLS,& TABLE
      DF=NSUM-4
      SERRM=SSERR/DF
```

```
      FR=SSR/SERRM
      FC=SSC/SERRM
      FRC=SSRC/SERRM
  500 CONTINUE
C IF F(I) IS NEGATIVE DROP CASE FROM CONSIDERATION
      IF(FR.LT.0.0) GO TO 316
      IF(FC.LT.0.0) GO TO 316
      IF(FRC.LT.0.0) GO TO 316
C CALC PROBABILITIES FOR ROWS,COLS & TABLE
      C2=C1/DF
      DO 305 I=1,3
      X1=SQRT(F(I))*(1.-0.25/DF)/SQRT(1.+C.5*F(I)/DF)
      X2=1.0+X1*(CC1+X1*(CC2+X1*(CC3+X1*CC4)))
      X3=X2*X2*X2*X2
      PF(I)=1.-1./X3
C CNT NON-SIG CASES OF F
      IF(PF(I).LE.0.95) NSIG(LS,I)=NSIG(LS,I)+1
  305 CONTINUE
C PRINT CASE PARAMETERS
      IF(LINE.LE.56) GO TO 307
      WRITE(3,200)DATE,IPAGE
      LINE=6
      IPAGE=IPAGE+1
  307 WRITE(3,201)P,N,F,PF,SIG
      LINE=LINE+1
C DO WE HAVE ENOUGH CASES OF SIG AND NON-SIG CHI YET
  312 IF(SCHI.LT.NC) GO TO 306
C CALC & PRINT SUMMARY
C MATRIX FSIG(2,6) CONTAINS THE PERCENTS AS FOLLOWS
C   COL 1  PERCENT NON-SIG F'S
C   COL 2  PERCENT SIG F'S
C   ROW 1  ROW F NON-SIG CHI SQ'D
C   ROW 2  COL F NON-SIG CHI SQ'D
C   ROW 3 TABL F NON-SIG CHI SQ'D
C   ROW 4  ROW SIG CHI SQ'D
C   ROW 5  COL SIG CHI SQ'D
C   ROW 6  TAB SIG CHI SQ'D
  315 CN=NC
      K=2
      DO 314 I=1,6
      L=I-(I-1)/3*3
      IF(I.EQ.4) K=1
      TMP=FLOAT(NSIG(K,L))
      FSIG(1,I)=TMP/CN*100.
      FSIG(2,I)=(CN-TMP)/CN*100.
  314 CONTINUE
      WRITE(3,202)DATE,FSIG,NC
      STOP
C F(I) IS NEG DROP CASE
  316 IF(SIG.EQ.LTRS) SCHI=SCHI-1
      IF(SIG.EQ.LTRN) NSCHI=NSCHI-1
      GO TO 306
      DEBUG UNIT(3)
      AT 304
      DISPLAY I,A
      AT 500
```

G LEVEL 21 MAIN DATE = 77056 11/40/57

```
      DISPLAY TSUMS,R1S,R2S,C1S,C2S,TABSUM,TSS,SSR,SSC,SST,SSCELL,SSRC,
     1SSERR,DF,SERRM,FR,FC,FRC
      ICNT=ICNT+1
      IF(ICNT.GT.2) STOP
      END
```

Multiple Linear Regression Viewpoints
Vol. 7 No. 3 1977

SHRINKAGE IN R^2 AND UNBIASED ESTIMATES OF TREATMENT EFFECTS USING $\hat{\omega}$

STARRETT DALTON
UNIVERSITY OF CALIFORNIA, RIVERSIDE

Abstract

The amount of variance accounted for by treatment can be estimated with $\hat{\omega}^2$or with R^2 (symbolized as R^2_c after a shrinkage formula has been applied). Monte Carlo methods were employed to compare $\hat{\omega}^2$, R^2_c, and R^2 in terms of bias and precision. R^2_c and $\hat{\omega}^2$ produced estimates which were negligibly biased. The bias in R^2, while consistently positive, decreased as sample size increased and was too small to be of practical importance when $n \geq 50$. $\hat{\omega}^2$, R^2_c and R^2 were all most precise with large samples and least precise when treatment effects were moderate in magnitude.

Presented at A.E.R.A., New York, NY, April, 1977.

Shrinkage in R^2 and Unbiased Estimates of Treatment Effects Using ω^2

William Hays (1963) encouraged investigators to provide estimates of the magnitude of treatment effects in addition to the statistics and probability levels which are customarily reported in analysis of variance. He offered omega squared (ω^2) as an unbiased estimator of the parameter eta squared (η^2).

$$\omega^2 = \frac{SS_B - (J-1)MS_\omega}{SS_T + MS_\omega}$$

SS_T = sum of squares total

SS_B = sum of squares between groups

MS_w = mean of squares within groups

J = number of groups

and

$$\eta^2 = \frac{\sigma_y^2 - \sigma_\varepsilon^2}{\sigma_y^2}$$

σ_y^2 = variance of the dependent variable

σ_ε^2 = the common homogeneous variance of the Y_{ij} about y_j

η^2 = a population measure of the magnitude of treatment effects relative to the total variance in the experiment

Carroll and Nordholm (1975) have demonstrated that another estimate of η^2 (epsilon squared) developed earlier by Kelley (1935) is as accurate as ω^2. Nevertheless ω^2 appears to be most popular among researchers who use analysis of variance.

Researchers who use regression analysis for testing the same hypotheses tested in analysis of variance (e.g., McNeil, Kelly, and McNeil, 1975; Namboodiri, Carter and Blalock, 1975) use the squared multiple correlation coefficient (R^2) to estimate the proportion of variance accounted for by treatment (i.e., η^2). Because R^2 is positively biased a shrinkage formula is commonly used to provide a value of R^2 corrected for bias (R_c^2).

$$R_c^2 = 1 - (1-R^2)\left(\frac{N-1}{N-k-1}\right)$$

R_c^2 = R^2 after correction for bias

N = number of subjects

k = number of independent variables

In the past, the selection of R_c^2 or $\hat{\omega}^2$ to estimate η^2 has been determined by the researcher's preference for traditional analysis of variance or regression analysis. Because no one has compared these statistics in terms of bias and precision, no evidence is available for selection on the basis of accuracy. The primary purpose of this study, therefore, is to compare the sampling distributions of R^2, R_c^2, and $\hat{\omega}^2$ in terms of bias and precision.

Procedure

Using Monte Carlo methods, the study was conducted in the context of a three level, one way analysis variance. It was assumed the results would generalize directly to higher order analysis of variance. The values of η^2 selected were .00, .05, .15, .40. .75, and .90. Sample sizes ranged from N=30 (10,10,10) to N=600 (200, 200,200). Error terms were randomly selected from normal populations with homogeneous variances. Carroll and Nordholm (1975) studied the effect of violating the assumptions of normality and homogeneity when estimating η^2 using $\hat{\omega}^2$. The reader interested in this problem is referred to their work. For each value of η^2 and for each sample size 1000 simulated experiments were conducted. Data were generated by the pseudorandom number generator in the IBM scientific package.

Results

Bias. For each value of η^2 and for each sample size the mean and standard error (SE) were calculated for R^2, R^2_c, and $\hat{\omega}^2$ (Table 1). R^2_c and $\hat{\omega}^2$ produced estimates of η^2 which were negligibly

Insert Table 1 about here

biased. The bias in R^2, while consistently positive, decreased as sample size increased and was too small to be of practical importance when $n \geq 50$.

Precision. As reflected in the standard errors, the imprecision of all three statistics decreased as sample size increased. Imprecision was greatest in the middle of the range of η^2 (i.e., for $\eta^2=.15$ and $\eta^2=.40$). For small sample sizes R^2 was very slightly, yet consistently more precise than R_c^2 and $\hat{\omega}^2$.

Discussion

Any attempt to pinpoint a tendency in R_c^2 or $\hat{\omega}^2$ to be positively or negatively biased, based upon the present data, would probably be misleading. The mean values of R_c^2 and $\hat{\omega}^2$ always fell within two SEs of the population value (η^2) and 92% of the time both estimates were within one standard error of η^2. The mean value of $\hat{\omega}^2$ never differed from the population value by more than one SE. By these standards $\hat{\omega}^2$ is slightly less biased than R_c^2, however, this conclusion is overshadowed by the tendency of both measures to yield accurate estimates in terms of bias.

When Hays developed $\hat{\omega}^2$ he stressed the importance of estimating the magnitude of treatment effects when sample sizes were large. "Virtually any study can be made to show significant results if one uses enough subjects, regardless of how nonsensical the content may be." It is indeed ironic that the use of $\hat{\omega}^2$ (or

R_c^2) instead of the familiar R^2 is least necessary when sample size is large. In this case the estimates of experimental effects using R^2 show only negligible bias. In the present study the mean value of R^2 ($n \geq 50$) never differed from η^2 by more than one standard error. This degree of bias should not concern most applied researchers. In addition, R^2 is routinely produced by many packaged computer programs and consequently is more readily available than R_c^2 and $\hat{\omega}^2$.

In conclusion, it appears that R_c^2 and $\hat{\omega}^2$ may serve equally well in estimating the magnitude of experimental effects. Both are superior to R^2 in this regard when $n \leq 30$. For large samples R^2 shows little bias and in many circumstances may be sufficient as an estimate of $\hat{\omega}^2$.

References

Carroll, R. M. & Nordholm, L. A. Sampling characteristics of Kelley's and Hays' . Educational and Psychological Measurement, 1975, 35, 541-554.

Hays, W. L. Statistics for Psychologists. New York: Holt, Rinehart and Winston, 1963.

Kelly, T. L. An unbiased correlation ratio measure. Proceedings of the National Academy of Sciences, 1935, 21, 554-559.

McNeil, K., Kelly, F., & McNeil, J. Testing Research Hypotheses Using Multiple Linear Regressions. Carbondale: Southern Illinois University Press, 1975.

Namboodiri, N. K., Carter, L. F., & Blalock, H. M., Jr. Applied Multivariate Analysis and Experimental Design. New York: McGraw Hill, 1975.

Table 1

Means and Standard Errors (SE) of the Multiple Correlation Coefficient (R^2), Omega Squared ($\hat{\omega}^2$), and the Corrected Multiple Correlation Coefficient (R_c^2) for Each Condition

		Eta Squared (η^2)											
	Estimates	.00		.05		.15		.40		.75		.90	
Sample size	of η^2	X	SE	X	SE	X	SE	X	SE	X	SE	X	SE
	R^2	.069	.065	.116	.093	.209	.117	.443	.118	.771	.059	.909	.024
10	$\hat{\omega}^2$	.000	.068	.049	.097	.147	.123	.394	.125	.748	.064	.899	.017
	R_c^2	.000	.070	.051	.099	.150	.126	.402	.126	.754	.063	.902	.026
	R^2	.023	.022	.070	.048	.167	.068	.411	.073	.755	.037	.903	.015
30	$\hat{\omega}^2$	.000	.023	.049	.049	.146	.069	.395	.075	.748	.038	.899	.015
	R_c^2	.000	.023	.049	.049	.148	.070	.397	.075	.750	.037	.900	.015
	R^2	.013	.014	.063	.035	.161	.051	.408	.054	.753	.027	.901	.011
50	$\hat{\omega}^2$	.000	.014	.050	.036	.149	.051	.398	.055	.749	.027	.899	.011
	R_c^2	.029	.014	.023	.037	.126	.053	.382	.056	.743	.028	.897	.011
	R^2	.007	.007	.057	.024	.157	.036	.405	.039	.753	.020	.901	[illegible]
100	$\hat{\omega}^2$	.000	.007	.051	.024	.151	.036	.401	.039	.750	.020	.900	[illegible]
	R_c^2	.014	.007	.038	.025	.139	.037	.393	.040	.748	.020	.899	.008
	R^2	.004	.004	.055	.020	.154	.030	.404	.032	.752	.016	.901	.007
150	$\hat{\omega}^2$	.000	.004	.050	.020	.150	.030	.401	.032	.751	.016	.900	.007
	R_c^2	.000	.004	.050	.020	.151	.030	.401	.032	.751	.016	.901	.007
	R^2	.003	.003	.054	.017	.154	.026	.403	.028	.752	.014	.901	.006
200	$\hat{\omega}^2$	.000	.003	.050	.017	.151	.026	.401	.028	.751	.014	.900	.006
	R_c^2	.000	.003	.050	.017	.151	.026	.401	.028	.751	.014	.901	.006

Multiple Linear Regression Viewpoints
Vol. 7 No. 3 1977

THE USE OF MULTIPLE REGRESSION ANALYSIS IN PREDICTING SUCCESS IN THE COUNSELING PRACTICUM *

JOSEPH M. WALTON
ASSISTANT DEAN OF GRADUATE STUDIES
THE UNIVERSITY OF AKRON

Abstract... The present exploratory study investigated the relationship between several predictor variables and the criterion of success in the counseling practicum among 93 recent graduates of a counselor education program. The investigation revealed that the best predictor of success in the counseling practicum was the square of the graduate grade point average ($ggpa^2$). This suggests the possibility of a curvilinear relationship between this predictor and the criterion. The interaction of female by Miller Analogies Test score (MAT) and the single variable of undergraduate grade point average (Ugpa) also appeared early in the equation. Type of undergraduate institution, type of graduate degree earned, and sex as a single independent variable demonstrated little relationship to the criterion.

Introduction

The counseling practicum in most counselor education programs is conceived as a comprehensive experience in which the counselor education student is asked to demonstrate his or her ability to counsel in a practical sense, while using the methods and techniques that he or she has acquired in previous courses. It is a pre-service experience which normally occurs after substantive course work in counselor education has been completed, but before the student is certificated as a practicing counselor.

While one's rating in the counseling practicum may bear questionable relationship to one's ultimate success in the counseling field, the counseling practicum may be a fairly good barometer of the student's current level of achievement in the practical aspects of counseling, and an indicator of potential success in the field, if one assumes minimum bias in the subjective judgements of the practicum supervisor.

The present study is concerned with determining possible predictor variables that will be useful in predicting success in the counseling practicum. A selected review of recent literature reveals some studies which are relevant to

* non-reviewed

the present study.

James and Dumas (1976) examined college grade point average as a predicator of teaching competency and found a positive relationship between grade point average and competency ratings of student teachers. The authors emphasized that low risk students tended to have lower overall grades, as well as lower student teaching competency ratings.

Richards' (1974) study sought to predict performance in a combined undergraduate and medical education program. He found that the best predictor for a given criterion variable was previous standing on that same variable, and that academic performance in the program was largely independent of other criteria.

A later study by Wallace and Schwab (1976) compared five models which were used to predict graduate admission committee decisions. These investigators found that grade point average and Graduate Record Examination score tended to be weighted about the same from one model to another.

The studies cited above investigated linear relationships. However, the present study attempted to extend beyond the investigation of linear relationships and to explore curvilinear relationships through squared variables.

The present exploratory study investigated possible predictor variables for the criterion of success in the counseling practicum. The independent variables included type of undergraduate institution, sex, type of graduate degree earned, undergraduate grade point average (ugpa) graduate grade point average (ggpa) and Miller Analogies test score, (MAT). In addition, the interaction of several of the above variables was included. Specific interaction variables were: type of graduate degree by MAT; sex by ugpa, sex by ggpa, sex by MAT. Also, ugpa, ggpa, and MAT were squared in order to determine if curvilinear relationships existed.

Thus, the purpose of the present investigation was to examine possible predictors of counseling practicum success, while using a multiple regression approach. The intent was to derive a regression equation which would be useful in a replicative study.

Methods and Procedures

The present study was completed with 93 subjects who recently received the Master of Science in Education or Master of Arts in Education degree in the area of counseling from a large urban, midwestern university. The subjects, 54 females and 39 males, were randomly selected from the complete lists of recent graduates, and any subject with missing information, such as MAT score was excluded from the study. Forty-one of the subjects received the M.A. degree in Education, while the remaining 52 received the M.S. degree in Education. Type of undergraduate institution included: (1) the home institution, (2) other public institutions within the state, (3) private institutions within the state, and (4) public and private institutions outside of the state.

The subjects were divided into groups of "most successful" and "least successful" on the basis of final ratings which were assigned by the practicum supervisors. These two classifications were used as the criterion.

Data Analysis

The data were analyzed by the use of the Statistical Package for the Social Sciences (SPSS) multiple regression subprogram (Klecka, Nie and Hull, 1975). Input data included the independent and dependent variables enumerated above. The SPSS program computed a sequence of multiple linear regression equations in a stepwise manner. At each step the independent variable which made the greatest reduction in the error sum of squares was added to the regression equation. The results are presented in the following section. The present study was completed as ex post facto research, and bears the limitations normally associated with such research (Kerlinger, 1973).

Results

The findings reveal that the predictor variables in the full model controlled 32.3 percent of the total variance. As indicated in Table 1 below, ggpa^2 accounted for most (23.7 percent) of the controlled variance. The interaction variable of female by MAT was nexted added to the regression equation to increase the R^2 by 2.3 percent. Ugpa was entered on step three, followed by ugpa^2. MAT^2 was added at step five to raise the R^2 to .295.

Beyond step five, additional variables were added to the regression equation in the following order: MAT, female, male by ugpa, ggpa, type of undergraduate institution, and male by ggpa. At this point the F- level became insufficient for further computation.

Newman (1973) pointed out that "multiple correlations (R) tend to be biased upward..." and that ..."R tends to be higher in the sample than in the population from which the sample is drawn." In order to correct this condition, he suggests the use of shrinkage estimates. This was not done in the present study, because of its exploratory nature. However, in a replicative study it will be important to make use of shrinkage estimates.

Discussion and Conclusions

As indicated above, graduate grade point average (squared) accounted for most of the variance in the regression models. This finding is somewhat similar to those of James and Dumas, who found a positive linear relationship between gpa and competency ratings of student teachers. However, while James and Dumas found linear relationships, the present study appears to reveal a curvilinear relationship between ggpa^2 and counseling practicum rating.

The interaction variable of female by MAT and the single variable of ugpa were also presented early in the regression equation. This in part seems consistent with the study by Wallace and Schwab, who found consistency in the weighting of undergraduate gpa and Graduate Record Examination score, when they tried to predict graduate admissions committee decisions. It is pointed out here that the primary bases for graduate school admission among students in the present study was undergraduate grade point average and MAT score.

The Richards study concluded in essence that the best predictor for a criterion variable was previous standing on that variable prior to admission to the desired program. This suggests that characteristics which practicum supervisors seek to develop should be sought in preliminary screening devices when students initially apply for counselor enucation programs.

Conclusively, while the present study did not generate a prediciton equation which is generalizable to other populations, it provided additional insight on ways of predicting success in the counseling practicum, by introducing the possibility that a second degree curvilinear relationship exists when ggpa^2 is introduced in a regression equation utilzing counseling practicum rating as the criterion. This investigator plans a replicative study to explore this concept further.

TABLE 1

Stepwise Multiple Regression Analysis for Selected Predictors of Counseling Practicum Success

Step	Predictor Variables	Multiple R	R^2	Increase in R^2
Restricted models				
1	$ggpa^2$	.487 *	.237	
2.	$ggpa^2$, female by MAT	.510 *	.260	.023
3.	$ggpa^2$, female by MAT ugpa	.525 *	.275	.016
4.	$ggpa^2$, female by MAT ugpa, $ugpa^2$, MAT^2	.533 * .538 *	.284 .290	.009 .006
5.	$ggpa^2$, female by MAT ugpa, $ugpa^2$, MAT^2	.538 *	.290	.006
Full Model				
11	$ggpa^2$, female by MAT, ugpa, $ugpa^2$, MAT^2, MAT, Female, Male by ugpa ggpa, type institution, male by ggpa	.568 *	.323	.030

* $p < .001$

References

James, Tery L. and Dumas, Wayne. College gpa as a predictor of teacher competency: a new look at an old question, The Journal of Experimental Education, 1976, 44(4), 40-43.

Kerlinger, Fred N. Foundations of Behavioral Research. New York: Holt, Rinehart and Winston, Inc., 1973.

Klecka, William R., Nie, Norman H. and Hull, C. Hadlai, Statistical Package for the social sciences primer. New York: McGraw-Hill Book Company, 1975.

Richards, James M. Jr. et. al. Predicting preformance in a combined undergraduate and medical education program. Educational and Psychological Measurement, 1974, 923-931.

Newman, Isadore, Variations between shrinkage estimation formulas and the appropriatness of their interpretation, Multiple Linear Regression Viewpoints, 1973, 4(2), 45-48.

Wallace, Marc J. Jr., and Schwab, Donald P. A cross-validated comparison of five models used to predict graduate admissions committee decisions, Journal of Applied Psychology, 1976, 61 (5), 559-563.

Multiple Linear Regression Viewpoints
Vol. 7 No. 3 1977

RELATIONSHIPS BETWEEN RESULTS OBTAINED ON THE ERTL MACHINE AND THE WECHSLER INTELLIGENCE SCALE FOR CHILDREN (WISC)

ROSE K. GANTNER
DEPT. OF PSYCHOLOGY
TROY STATE UNIVERSITY

JOSEPH D. GEORGE
DEPT. OF SPECIAL EDUC.
COLUMBUS COLLEGE

MARK E. MEADOWS
DEPT. OF COUNSELOR EDUC.
AUBURN UNIVERSITY

ABSTRACT

The purpose of this study was to examine relationships between the neural efficiency (NE), symmetry, and time difference (TD) scores on the Ertl machine and WISC scale scores for a group of 22 normal children and a group of 22 children with suspected learning disabilities, all ranging from 8 to 10 years of age. Multiple linear regression techniques were used to analyze the data. Some statistically significant relationships did occur between Ertl machine scores and WISC-V, WISC-P, and WISC-F scale scores for groups 1 and 2. Results supported Ertl's findings that normals and children with learning disabilities (LDs) would have similar NE scores (learning potential). Several symmetry scores (Hemispheric synchronization) and WISC scores correlated significantly in positive directions for both groups. Significant differences occurred between the TD scores (indicator of LDs) but results were in direct contrast to Ertl's claim since group 1 (normals) obtained higher mean scores than group 2.

REVIEW OF LITERATURE

The WISC instrument usually correlates consistently well with other measures of intelligence and has been widely used for diagnosing Children with Learning Disabilities (CLDs) (Sattler, 1974; Coleman & Rasof, 1963). Results of WISC scores have indicated that CLDs often obtain higher WISC-P scores than WISC-V scores. Studies by Rourke, Young, and Flewelling (1971), Wells (1970), and Ackerman, Peters, and Dykman (1971) have concluded that this may occur because CLDs usually have serious reading problems and faulty attention.

Other researchers (Meier, 1970; Kershner & Kershner, 1973; Sperry, 1975) supported the conclusion that CLDs on WISC scores (WISC-P greater than WISC-V) may be due to asymmetry of the brain's two hemispheres. Apparently these children have hemispheric cross integration difficulties which result in problems of general behavior and in academic tasks. Numerous experiments have demonstrated that the two hemispheres of the brain function differently with respect to the kinds of information each hemisphere stores and processes (Milner, 1967; Rosenfield & Klivington, 1975; Delgado, 1975; Sperry, 1975; Gardner, 1975). Studies generally support the conclusion that the left hemisphere is highly verbal and mathematical, performing with computer-like sequential logic, and the right hemisphere is instrumental in high-order integrative visual-spatial activities and cannot yet be simulated by computers. Pines (1973) has reported on new experimental methods being used whereby individuals can learn to integrate the hemispheres of the brain more effectively or to rely more heavily on one hemisphere than the other. This type of training may prove particularly useful for CLDs.

Much remains to be learned about the mechanism of brain function, but several neurophysiological studies have indicated that all behavior may be generated and controlled by a simple neural mechanism. Neurons basically generate electricity and secrete chemicals. The two functions together enable neurons to deal with information received from the environment and to issue appropriate commands to the glands and muscles (Klemm, 1972). With the aid of computer technology, the electroencephalogram (EEG-brain waves) has become one of the most convenient ways to monitor the overall processing reactions in the brain.

Ertl (1973) indicated that the analysis of brain waves can provide information about the learning capacity of an individual. The speed of information transfer occurring within the brain is one of the indicators of brain efficiency and hence the ability to learn. He developed a machine to measure such brain activity.

The Ertl machine (brain wave analyzer) provides neurophysiological measurements (NE, SYM, and TD scores). These measurements are obtained by having electrodes connected to the head and ear lobes of each examinee by an electrode helmet. The helmet detects and amplifies the brain waves by being placed over the left and right sensory areas of the brain. The machine does not measure intelligence as conventionally defined; rather it measures the rate of information transfer occurring within the brain in the nonalpha spectrum. Ertl (1974) reported that the machine also measures the degree of cross correlations between the EEG derived from the right and left hemispheres of the brain. Thus, Ertl (1974) proposes a more objective, culturally fair measurement and questions the use of standardized intelligence tests like the WISC as a measure of assessment. The educational application of this machine may be in its ability to measure learning potential (NE score) as well as detect early learning disabilities (TD score).

Further research on the Ertl machine is minimal at this time since the instrument is relatively new (Ertl, 1975b) and only six existed throughout the U.S.A. during this study. Data, however, by Everhart, China, and Auger (1974) indicated that an inverse relationship existed between the Wechsler Adult Intelligence Scale (WAIS) verbal scores and NE as measured by Ertl's first machine, model 0-1.

The forthcoming decade will undoubtedly observe siginificant progress to resolve the issue of empirical relationships between human intelligence and electrophysiological correlates. Perhaps this type of research may manifest a clearer association between human brain functions and behavior.

STATEMENT OF THE PROBLEM

The primary purpose of the study was to determine relationships between the NE, Symmetry, and TD scores of the Ertl machine and WISC-Verbal (V), Performance (P), and Full scale (F) scores for a group of normal children ages 8 and 9 years (group 1) and a group of children with suspected learning disabilities of comparable ages (group 2). A secondary purpose was to determine whether significant differences existed between two groups of subjects on NE, Symmetry, and TD scores of the Ertl machine.

METHODS AND PROCEDURES

Selection of Subjects

Subjects for this study were 22 normal children (group 1) and 22 children with suspected learning disabilities (LDs) (group 2). Subjects were limited to the 8- to 10-year-old range. The two groups were similar in age, sex and race (Caucasion). The mean for normals was 8 years 9 months; the group included 13 males and nine females. The mean for group 2 was 9 years 3 months. The group included 14 males and eight females.

Procedure

Group 1 (normals) was selected by a random sampling procedure of 50 third grade children attending Brookstone, a private school in Columbus, Georgia. A letter was written and sent to the parents of the 22 children. All parents returned a slip with their approval within a week. Testing was completed during the month of May, 1975. The WISC was initially administered, followed by the Ertl machine. Subjects were not on medication (drugs) or recieving remediation for academic learning problems.

Group 2 (children with suspected learning disabilities) were selected from two sources: (1) the Diagnostic Learning Center at Columbus College, Columbus, Georgia, and (2) the Columbus chapter of the Georgia Association for Children with Learning Disabilities. Twenty-two suspected LDs were obtained from those sources.

Permission to use Columbus College as part of group 2 was approved by the college Vice-President and two assistant directors at the center. All parents who were contacted by phone and who had children in the required age limites cooperated and signed a written statement giving permission to test. The children were all tested at the Diagnostic Learning Center, Columbus College, during the months of May-August, 1975. None of the CLDs were enrolled in special education classes at that time.

Sources of Data

Data consisted of test results from both the WISC and the Ertl machine.

The WISC was standardized on 2200 Caucasian American boys and girls selected to be representative of the 1940 U.S. census. The WISC has five verbal subtests (Information, Comprehension, Arithmetic, Similarities, and Vocabulary) and five performance subtests (Picture Completion, Picture Arrangement, Block Design, Object Assembly, and Coding). Subtests were administered according to the manual's instruction (Wechsler, 1949). Administering the WISC to each child took approximately one hour.

The Ertl machine has three separate subtests which measure various functions. These raw scores are: (a) test A, Phase score--machine emits one pulse for each milisecond that a brain wave is out of phase and displays a count of these pulses; (b) test B--average frequency of all the brain waves (alpha, beta, theta, and delta), and (c) test C--measuring alpha frequency brain rhythms exclusively. All electrophysiological measurements are taken in one complete EEG sampling over a 10-second period.

On the Ertl subtests for each child, five readings were taken. The median score was obtained for test B, with corresponding test scores for subtests A and C. Ertl (1975a) indicated that this method seemed to be most appropriate for research purposes. The digital computer on the Ertl machine then indicated the frequency of brain waves by recording the number of beats per second. The results were then adjusted by Ertl's formulas to investigate NE, Symmetry, and TD scores respectively. The NE score is composed of subtests B and C (with D as an age correction factor). Symmetry score is composed of Test A and the TD score of subtests A and B.

The total administration time per child combined with the proper placement of the electrode helmet was approximately two minutes. Subjects were not required to answer questions or perform any tasks while on the machine, but rather were encouraged to sit comfortably (both physically and psychologically) during the administration.

The Experimental Design

Two research designs were incorporated. The first was the pre-experimental, designs 1, one-shot case study (Campbell & Stanley, 1973). This design was used to investigate the correlations obtained.

The second design was quasi-experimental, non-equivalent-control group design (design 10 in Campbell & Stanley, 1973). Although a pre and post test were not formulated, groups 1 and 2 existed. This design was used to investigate regression models and differences occurring between the groups.

Statistical Treatment of Data

Multiple linear regression techniques were used to analyze the data. These techniques resulted in a matrix of Pearson Product-Moment Correlation Coefficients and in a One-Way Analysis of Variance. The correlations were employed to assess the degree of relationship between the variables studied.

The "F" ratios were used to determine level of differences between groups 1 and 2. The .05 alpha level was used.

RESULTS, DISCUSSION, AND CONCLUSIONS

Correlation coefficients of the Wechsler Intelligence Scale for Children (WISC) and Ertl machine scores for the two groups are presented in Tables 1 and 2. A description of variables used for the regression models is presented in Table 3. Significant differences between Ertl machine scores for groups 1 and 2 are reported in Table 4.

To answer the question asked by Null Hypothesis One, "There will be no significant relationship between the neural efficiency (NE) scores and WISC verbal (WISC-V) scores of groups 1 and 2," the NE and WISC-V scores of group 1 were correlated. Table 1 shows that a nonsignificant *r* of .262 was found. As shown in Table 2, and *r* of -.434 was found between NE and WISC-V scores for group 2 and was significant at the .02 alpha level.

When Null Hypothesis Two, "There will be no significant relationship between the NE scores and WISC performance (WISC-P) of groups 1 and 2, was examined by correlating the NE and WISC-P scores of groups 1 and 2, *r*'s of -.095 and .018 were obtained. Both correlations were nonsignificant.

The question asked by Null Hypothesis Three, "There will be no significant relationship between the Symmetry scores and WISC-V scores of groups 1 and 2", was answered by correlating the Symmetry and WISC-V scores of both groups.

An *r* of -.412 was found when the Symmetry and WISC-V scores of group 1 were correlated. For group 2, an *r* of .436 was found. Although *r*'s for groups 1 and 2 were opposite in directions, both scores were

TABLE 1

CORRELATION MATRIX COMPUTED BETWEEN ERTL MACHINE SCORES AND WISC SCORES FOR GROUP 1 (NORMALS)

	NE		Symmetry		TD	
	r	s	r	s	r	s
VIQ	.262	.12	-.412	.02*	.305	.08
PIQ	-.095	.34	.449	.01**	-.504	.01**
FIQ	.107	.32	-.118	.31	.064	.39

r = Pearson correlation coefficient
s = level of significance

*p .05
**p .01

TABLE 2

CORRELATION MATRIX COMPUTED BETWEEN ERTL MACHINE SCORES AND WISC SCORES FOR GROUP 2 (SUSPECTED LDs)

	NE		Symmetry		TD	
	r	s	r	s	r	s
VIQ	-.434	.02*	.436	.02*	-.305	.08
PIQ	.018	.47	.070	.38	-.050	.41
FIQ	-.309	.08	.369	.05*	-.279	.11

r = Pearson correlation coefficient

s = level of significance

*p .05

significant at .02 alpha.

Symmetry and WISC-P scores of group 1 were correlated to investigate Null Hypothesis Four, "There will be no significant relationship between the symmetry scores and WISC-P scores of groups 1 and 2". A significant *r* of .449 was obtained for group 1 and a nonsignificant *r* of .070 was found for group 2 (see Tables 1 and 2).

To examine Null Hypothesis Five, "There will be no significant relationship between the time difference (TD) scores and WISC-V scores of groups 1 and 2", the TD and WISC-V scores were correlated. An *r* of .305 for group 1, and an *r* of -.305 was obtained for group 2. Both *r*'s approached significance (*p* .08) as shown in Tables 1 and 2.

The testing of Null Hypothesis Six, "There will be no significant relationship between the TD scores and WISC-P scores of groups 1 and 2", resulted in an *r* of -.504 for group 1, and an *r* of -.050 for group 2. As shown in Tables 1 and 2 the relationship between the TD scores and the WISC-P scores for group 1 were significant at the .01 level. Those for group two were not significant.

Null Hypothesis Seven, reported in Table 4, was a comparison of the NE scores between the groups. The F ratio obtained was .1905 with 1/43 *df*; the probability of that being a chance occurance was .6647.

To test the differences between the symmetry scores for groups 1 and 2, Null Hypothesis Eight, as reported in Table 4, was investigated. Here, the F ratio calculated was .3607 with 1/43 *df*; the probability of that being a chance occurance was .5514.

When comparing the TD scores between the groups to test Null Hypothesis Nine (reported in Table 4), the F ratio computed was 4.093 with 1/43 *df* and *p* .05.

TABLE 3

A DESCRIPTION OF THE VARIABLES USED

Where the Full Model is:

$$Y_1 = a_oU + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + E$$

The Variables are:

Y_1 = A criterion, NE score on the Ertl machine - predicted from WISC-F

a_o = a_1 through a_5 = Partial regression weights

U = the Unit Vector (1 for each sample)

x_2 = Symmetry score (SYM)

x_3 = Time Difference score (TD)

x_4 = Group (1, Normals)

x_5 = Group (2, Suspected LDs)

E = Error vector, difference between predicted and actual score.

TABLE 4

MODELS, F-RATIOS, AND R^2 FOR PREDICTING NE SCORES FROM WISC-F SCORES

MODELS AND EXPLANATION	R^2	df	alpha	F	P
Null Hypothesis Seven	.0045	1.43	.05	.19054	.6647
There will be no significant difference between the NE scores of groups 1 and 2.	.0000				
Full Model:					
$Y_1 = a_0U + a_4x_4 + a_5x_5$					
Restriction:					
$a_4 = a_5 = a_o$					
Restricted:					
$Y_1 = a_oU = E$					
Null Hypothesis Eight	.0085	1.43	.05	.36065	.5514
There will be no significant difference between the symmetry scores of groups 1 and 2.	.0000				
Full Model:					
$Y_2 = A_oU + a_4x_4 + a_5x_5$					
Restriction:					
$a_4 = a_5 = a_o$					
Restricted:					
$Y_2 = a_oU + e$					

*p .05

TABLE 4 (cont)

MODELS, F-RATIOS, AND R^2 FOR PREDICTING NE SCORES

FROM WISC-F SCORES

MODELS AND EXPLANATION	R^2	df	alpha	F	P
<u>Null Hypothesis Nine</u>	.0887	1.43	.05	4.0928	.0495*
There will be no significant difference between the RD scores of groups 1 and 2	.0000				
Full Model:					
$Y_3 = a_oU + a_4x_4 + a_5x_5$					
Restriction:					
$a_4 = a_5 = a_o$					
Restricted:					
$Y_3 = a_oU + E$					

*<u>p</u> .05

Note: See Table 3 for a description of variables used.

Although some statistically significant relationships did exist between Ertl machine scores and WISC scores, the intercorrelations were relatively low. WISC correlations with other intellectual measures generally correlated higher than results obtained with the Ertl machine scores. Results on the machine did support significant relationships opposite in sign between scores obtained on NE and WISC-V and Symmetry and WISC-V, but the variances accounted for by the correlations are inconclusive. Interestingly the two groups did differ significantly on the WISC scores for intelligence but did not differ significantly on the NE scores (learning potential) for the Ertl machine (Table 4). It appears that the low intercorrelations may be caused by greater variations on the WISC than the Ertl Machine. With regard to comparing differences between groups 1 and 2, the TD score was significant. While empirical evidence for the existence of meaningful electrophysiological correlates of human intelligence did occur in some instances, further research seems warranted utilizing children known to possess specific learning disabilities.

The results of the present study support the following conclusions:

1. A significant negative relationship existed between NE and WISC-V scores for group 2 and correlations between the variables for the groups were opposite in sign.

2. The NE score was nonsignificant when comparing differences between the groups and supported Ertl's original hypothesis (1974) that NE scores would be similar for both normals and CLDs.

3. Several Symmetry scores and WISC scores were significant and correlated in a positive direction for groups 1 and 2.

4. The relationship between Symmetry and WISC-V scores were significant and opposite in sign for both groups.

5. Comparison of TD scores and WISC scores revealed a significant negative relationship for group 1 between TD and WISC-P.

6. The difference between the TD scores for groups 1 and 2 were significant but not in the direction Ertl predicted.

It seems apparent that additional research needs to be conducted regarding the relationships between scores obtained on the Ertl machine and human intelligence as measured by the WISC. Perhaps the present design could be replicated with these changes: 1) An increased sample size; 2) The experimental group could be diagnosed learning disabled children, rather than suspected LDS; and 3) The sample studied could include more children from lower socioeconomic levels, including minority group students.

It must be remembered that generalizations from the present research are limited because of the lack of reliability and validity data on the Ertl machine. Moreover, though this work identified some correlations between the Ertl machine and WISC, the WISC is a highly language oriented instrument. Future studies should be conducted with the Ertl machine and other measures of intelligence. Finally, WISC subtests, in addition to WISC scale scores, should be carefully examined in relationship to Ertl machine socres.

REFERENCES

Ackerman, P., Peters, J., & Dykman, R. Children with specific learning disabilities: WISC profiles. Journal of Learning Disabilities, 1971, 4, 150-166.

Campbell, D. T., & Stanley, J. Experimental and quasiexperimental designs for research. Chicago: Rand McNally College Publishing Company, 1973.

Coleman, J.C., & Rasof, B. Intellectual factors in learning disorders. Perceptual and Motor Skills, 1963, 16, 139-152.

Delgado, J. Jose Delgado: Exploring inner space. In R. Restak (ed.), Saturday Review, 1975, 216, 21-25.

Ertl, J. P. Fourier analysis of evoked potentials and human intelligence. Nature, 1971, 230, 525-528.

Ertl, J. P. IQ evoked responses and Fourier analysis. Nature, 1973, 241, 209-210.

Ertl, J. P. Operating manual for brain wave analyzer. Toronto, Canada: Neural Models Limited, 1974.

Ertl, J. P. Operating manual for brain wave analyzer (0-3). Toronto, Canada: Neural Models Limited, 1975. (a)

Ertl, J. P. Research newsletter, Toronto, Canada: Neural Models Limited, 1975. (b)

Everhart, J. D., China, C. L. & Auger, R. A.. Measures of EEG and verbal intelligence: An inverse relationship. Physiological Psychology, 1974, 2, 374-378.

Gardner, H. The shattered mind. New York: Knopf Co., 1975.

Kershner, J., & Kershner, B. Dual brain asymmetry: Another cause of learning disorder? Academic Therapy, 1973, 3 (4), 391-393.

Klemm, W. R. Science, the brain, and our future. New York: The Bobbs-Merrill Company, Inc., 1972.

Meier, N. Effects of focal cerebral lesions on contralateral visuomotor adaption to reversals and inversion of visual feedback. Neurophsychologia, 1970, 8, 269-279.

Milner, P.M. The cell assembly: Mark II. Psychological Review, 1957, 64, 242-252.

Newman, I. An introduction to the basic concepts of measurement and evaluation. Akron: Akron University Press, 1973.

Pines, M. The riddles of recall and forgetfullness. Saturday Review, 1975, 216, 16.

Rosenfield, A., & Klivingston, K.W. Inside the brain: The last frontier. Saturday Review, 1975, 13-15.

Rourke, B.P., Young, G.C., Flewelling, R.W. The relationships between WISC verbal-performance discrepancies and selected verbal, auditory-perceptual, visual-perceptual, and problem solving abilities in children with learning disabilities. Journal of Clinical Psychology, 1971, 27, 475-479.

Sattler, J. M. Assessment of children's intellegence. Philadelphia: London: Toronto: W.B. Saunders Company, 1974.

Sperry, R. W. Left-brain, right brain. Saturday Review, 1975, 30-33.

Wechsler, D. Manual for the Wechsler intelligence scale for children. New York: Psychological Corporation, 1949.

Wells, C.G. A comparative study of children grouped by three basic score patterns on the Wechsler Intelligence Scale for Children (Doctoral dissertaion, University of Northern Colorado, 1970). Dissertation Abstracts International, 1970. (University Microfilms NO. 7-14, 543)

If you are submitting a research article other than notes or comments, I would like to suggest that you use the following format, as much as possible:

Title

Author and affiliation

Single-space indented abstract (the rest of the manuscript should be double-spaced)

Introduction (purpose—short review of literature, etc.)

Method

Results

Discussion (conclusion)

References

All manuscripts should be sent to the editor at the above address. (All manuscripts should be camera-ready copy.)

It is the policy of the sig=multiple linear regression and of *Viewpoints* to consider for publication articles dealing with the theory and the application of multiple linear regression. Manuscripts should be submitted to the editor as an original, double-spaced typed copy. A cost of $1 per page should be sent with the submitted paper. Reprints are available to the authors from the editor. Reprints should be ordered at the time the paper is submitted and 20 reprints will cost $.50 per page of manuscript. Prices may be adjusted as necessary in the future.

"A publication of the *Multiple Linear Regression Special Interest Group* of the American Educational Research Association, published primarily to facilitate communication, authorship, creativity, and exchange of ideas among the members of the group and others in the field. As such it is not sponsored by the American Educational Research Association nor necessarily bound by the Association's regulations.

"Membership in the *Multiple Linear Regression Special Interest Group* is renewed yearly at the time of the American Educational Research Association Convention. Membership dues pay for a subscription to the *Viewpoints* and are divided into two categories: individual = $2.00; and institutional (libraries and other agencies) = $12.00. Membership dues and subscription requests should be sent to the Secretary of the MLRSIG."

THE UNIVERSITY OF AKRON
AKRON, OHIO 44325

157PD 77
SPANER STEVEN D.
BEHAVIORAL STUDIES
UNIV OF MO-ST LOUIS
ST LOUIS, MISSOURI 63121

AKRON OHIO
MAR 1 '77
U.S. POSTAGE
25
PB

TABLE OF