DOCUMENT RESUME

ED 093 901 TM 003 716

| | |
|---|---|
| AUTHOR | Pohlmann, John T.; McShane, Michael G. |
| TITLE | Applying the General Linear Model to Repeated Measures Problems. |
| PUB DATE | [74] |
| NOTE | 51p.; Paper presented at the Annual Meeting of the American Educational Research Association (59th, Chicago, Illinois, April 1974) |
| EDRS PRICE | MF-$0.75 HC-$3.15 PLUS POSTAGE |
| DESCRIPTORS | *Hypothesis Testing; Matrices; *Models; Post Testing; Predictor Variables; Pretesting; *Research Design; Statistical Analysis; *Tests of Significance |
| IDENTIFIERS | General Linear Model; Repeated Measures |

ABSTRACT

The purpose of this paper is to demonstrate the use of the general linear model (GLM) in problems with repeated measures on a dependent variable. Such problems include pretest-posttest designs, multitrial designs, and groups by trials designs. For each of these designs, a GLM analysis is demonstrated wherein full models are formed and restrictions are placed on the full models that reflect various research questions. The restricted models and full models are then compared with an F test to ascertain whether a significant reduction in the squared multiple correlation was realized as a result of the restriction. (Author)

# APPLYING THE GENERAL LINEAR MODEL TO REPEATED MEASURES PROBLEMS

John T. Pohlmann and Michael G. McShane

Southern Illinois University, Carbondale

A PAPER PRESENTED AT THE ANNUAL MEETING OF THE AMERICAN EDUCATIONAL RESEARCH ASSOCIATION'S MULTIPLE LINEAR REGRESSION SPECIAL INTEREST GROUP

Chicago, Illinois
April 1974

# APPLYING THE GENERAL LINEAR MODEL TO REPEATED MEASURES PROBLEMS

John T. Pohlmann
and
Michael G. McShane

Southern Illinois University, Carbondale

The purpose of this paper is to demonstrate the use of the general linear model (GLM) in problems with repeated measures on a dependent variable. Such problems include pretest - posttest designs, multi-trial designs, and groups by trials designs. For each of these designs, a GLM analysis is demonstrated wherein full models are formed and restrictions are placed on the full models that reflect various research questions. The restricted models and full models are then compared with an F test to ascertain whether a significant reduction in $R^2$ was realized as a result of the restriction.

# APPLYING THE GENERAL LINEAR MODEL TO REPEATED MEASURES PROBLEMS

John T. Pohlmann
and
Michael McShane

Southern Illinois University, Carbondale

Among the most difficult types of problems to formulate with the general linear model are those in which the dependent variable has been repeatedly measured on a group of subjects. The primary difficulty encountered is that of properly developing a design matrix which will extract variance in the dependent variable attributed to subject differences. This variance must be extracted since failure to do so results in a violation of the assumption of independence of errors. This assumption, unlike the assumptions of normality and homegeneity of group variances, cannot be countered by large sample sizes, nor by making group sizes equal (Glass et.al., 1973). A failure to allow for the dependence of errors in a multi-trial design can seriously affect the actual probabilities of making type I or type II errors. The dependence of errors in multi-trial designs can be controlled by removing from the analysis, variance in the dependent variable attribitable to individual or subject differences. In the application of the GLM to the analysis of such data, this is accomplished through the use of subject or people vectors.

This paper will demonstrate the use of subject vectors in three types of designs: (1) a one group pretest - post-test design, (2) a one group multiple-trial design with more than two trials, and (3) a multiple group-multiple trial design.

For each of these designs a design matrix was developed which reflected the full model being analyzed. Restrictions reflecting a research question were placed upon the full model and the resulting reduced model was then compared to the full model via on F-ratio to answer the research questions.

For each of the designs examined, the following aspects of the analysis will be shown:

1. The full general linear model, which reflects all of the information about the design, will be given.

2. A design matrix, which demonstrates the way the data would be coded for processing by a regression analysis program, will be given for the full model.

3. A research question will then be posed. Research and statistical (null) hypotheses will be stated, and a restriction will be placed on the full model that will force the model to conform to the statistical hypothesis.

4. The restricted model, which reflects a true statistical hypothesis, will be shown

5. A design matrix, which demonstrates the way the data would be coded for processing by a regression analysis program, will be given for the restricted model,

6. A summary of the results will then be given, and an F-ratio will be derived to conduct a test of significance on the statistical hypothesis.

## I. A One Group - Two Trial Design

A one group - two trial design is the type of design usually associated with a correlated or matched groups t-test; data on some dependent variable is obtained on a group of subjects at two time periods and a research hypothesis is usually stated regarding the relative magnitude of the mean on the dependent variable over the two time periods. For the following presentation, assume that a group of subjects has been administered a pretest on a political attitude scale, the group is then subjected to a series of political television commercials, and finally, the group is post-tested on a parallel form of the attitude scale. The researcher may then state a research question such as "Did the T.V. commercial improve political attitudes?" This research question could be answered by testing the statistical hypothesis

$$H_0: \mu_{pre} = \mu_{post}, \text{ or } \mu_{pre} - \mu_{post} = 0$$

where $\mu_{pre}$ = the expected value of the pretest attitude in the population

$_{post}$ = the expected value of the post-test attitude in the population.

The alternate or research hypothesis implied by the research question is

$$H_A: \mu_{post} > \mu_{pre}, \text{ or } \mu_{post} - \mu_{pre} > 0.$$

---

Insert Table 1 here

---

Sample data that will be used to demonstrate the analysis of this design appears in Table 1, N = 4 (the number of subjects). The full model necessary to reflect the information available to the researcher is:

$$\text{Model 1: } Y = a_0U + a_1X_1 + a_2X_2 + a_3P_1 + a_4P_2 + a_5P_3 + a_6P_4 + E_1$$

where: $Y$ = the attitude score (dependent variable) vector
$U$ = the unit vector
$X_1$ = a vector containing a 1 if the attitude score is from the pre-test, 0 otherwise
$X_2$ = a vector containing a 1 if the attitude score is from the post-test, 0 otherwise
$P_i$ (i=1 to 4) = a vector containing a 1 if the score is from person i, 0 otherwise
$a_0, a_1 \ldots a_6$ = a set of least squares weights derived so as to minimize the sum of the squared elements in the error vector, E.
$E$ = the error vector

The design matrix suggested by this model is shown in Figure 1.

---

Insert Figure 1 here

---

Now the weights $a_1$ and $a_2$ will take on values which will reflect the difference between the pre-test and post-test means, i.e., $a_1 - a_2 = \overline{X}_1 - \overline{X}_2$, where $\overline{X}_1$ and $\overline{X}_2$ are the pre-test and post-test means respectively. Recall that the statistical hypothesis was $\mu_{pre} = \mu_{post}$, hence the restriction on Model 1 required to test the statistical hypothesis is

$$a_1 = a_2$$

Imposing this restriction on Model 1, we obtain Model 2a.

$$\text{Model 2a:}\quad Y = a_0U + a_1X_1 + a_1X_2 + a_3P_1 + a_4P_2 + a_5P_3 + P_4 + E_2$$

Collecting terms with like weights we obtain Model 2b.

$$\text{Model 2b:}\quad Y = a_0U + a_1\,(X_1+X_2) + a_3P_1 + a_4P_2 + a_5P_3 + a_6P_4 + E_2$$

Since the vector $(X_1 + X_2) = U$, the simplest form of the reduced model is Model 2c.

$$\text{Model 2c:}\quad Y = a_0U + a_3P_1 + a_4P_2 + a_5P_3 + a_6P_4 + E_2$$

The design matrix for Model 2c appears in Figure 2.

---

Insert Figure 2 here

---

Associated with each of these models (Model 1 and Model 2c) will be a squared multiple correlation ($R^2$) which may be interpreted as the proportion of variance in the dependent variable, Y, accounted for by the weighted combination of the predictor variables, $U_1\ X_1 \ldots P_4$

$R^2$ for Model 1 is obtained as follows:

(1) $R^2 = 1 - \frac{ESS_1}{SS_y}$

where: $ESS_1$ = sum of the squared elements in the error vector, $E_1$, for Model 1.
$SS_y$ = sum of the squared deviations of the criterion scores (Y) about their grand mean.

In the same fashion an $R^2$ value for Model 2C can be obtained. The difference between these two $R^2$'s then serves as a measure of the contribution of trial differences to accounting for variance in the criterion variable. The difference between the two $R^2$'s may then be tested for significance with an F-ratio. The formula for F is

$$(2) \qquad F = \frac{(R_1^2 - R_2^2) / (l_1 - l_2)}{(1 - R_1^2) / (N - l_1)}$$

where $R_1^2$ = the squared multiple correlation between the prediction set and the dependent variables for Model 1

$R_2^2$ = the squared multiple correlation between the prediction set and the dependent variable for Model 2

$l_1$ = the number of linearly independent vectors (predictors) in Model 1

$l_2$ = the number of linearly independent vectors in Model 2

N = the total number of observations. In multiple trial designs this is the number of subjects x the number of trials

If the statistical hypothesis is true, or if the restricted model is the correct model in the population, F will be distributed as a central F distribution with $(l_1 - l_2)$ and $(N - l_1)$ degrees of freedom.

Numerical Solution Using the Data in Table 1

Table 2 contains the values of the regression weights, $R^2$ , and the F ratio observed in the statistical test of the restriction. Note that the number of linearly independent vectors in Model 1 is 5. This indicates that of the 7 vectors U, $X_1 \ldots P_4$, only 5 are linearly independent and 2 are redundant, or linearly dependent. The use of vector algebra on the data matrix for Model 1 (Figure 1) will show that $X_2 = U - X_1$, and $P_4 = U - P_1 + P_2 + P_3$ leaving U, $X_1$, $P_1$, $P_2$ and $P_3$ as a set of linearly independent vectors. It should be mentioned here that every predictor vector in Model 1 can be expressed as a linear combination of the other vectors in the predictor set, and the selection of the specific vectors which are to be considered as independent is arbitrary. The important point is that only 5 of the 7 predictor vectors are independent. Similarly, of the 5 predictor vectors included in Model 2C only 4 are linearly independent, since $P_4 = U - P_1 + P_2 + P_3$. Consequently the restriction imposed on Model 1 to obtain Model 2C resulted in restricting out one linearly independent vector.

The restriction resulted in a reduction in $R^2$ of .21 and the resulting F ratio was 5.40 which is insignificant, given an alpha level of .05 (one tailed). We therefore should conclude that the television commercials did not result in a significant increase in political attitudes.

II

## A ONE-GROUP - MULTIPLE TRIAL DESIGN WITH MORE THAN TWO TRIALS

The second type of design which will be considered in the present paper is the case in which one group is measured on more than two occassions. For the presentation of this design, assume that a group of students has been given an achievement test at the beginning of a course of study. (Pre-test), Then, at the end of the course the same group is given a parallel form of the same test (Post-test), and after a suitable period is given a third form (also parallel) of the same test as a long-term retention test (LTR). Sample data that will be used to demonstrate the analysis of this design appear in Table 3.

--------------------------
Insert Table 3 Here
--------------------------

Two of the possible research questions which may be of interest to a researcher in this kind of a situation may be a question comparing any two of the three tests, and a question comparing one of the tests to the average of the other two.

The first kind of research question might be worded:

> Is the mean for the Post-test greater
> than the mean for the Pre-test?

The research question could be answered by testing the statistical hypothesis that,

The mean for the Post-test is equal to the mean for the Pre-test,

or,

$$H_0 : \mu_{pre} = \mu_{post}$$

or,

$$\mu_{pre} - \mu_{post} = 0$$

When, $\mu_{pre}$ = the expected value of the pre-test score in the population

$\mu_{post}$ = the expected value of the post-test score in the population

The Full Model necessary to reflect the information in Table 3 is:

Model 3

$$Y = a_0U + a_1X_1 + a_2X_2 + a_3X_3 + a_4P_1 + a_5P_2 + a_6P_3 + E_1$$

where $Y$ = the dependent variable (achievement score)
$U$ = the unit vector
$X_1$ = 1 if the dependent variable was observed on the Pre-test, 0 otherwise
$X_2$ = 1 if the dependent variable was observed on the Post-test, 0 otherwise
$X_3$ = 1 if the dependent variable was observed on the LTR-test, 0 otherwise
$P_i$ (i = 1 to 3) = 1 if the observation is for person i, 0 otherwise
$E_1$ = error vector

$a_0, a_1, \ldots, a_6$ are least squares weighting coefficients calculated so as to minimize the sum of squared elements in the error vector.

The design matrix which reflects this model is presented in Figure 3.

---------------------------
Insert Figure 3 Here
---------------------------

In Model 3, weights $a_1$ and $a_2$ will take on values which will reflect the difference between the pre-test and the post-test means. Since the statistical hypothesis ($\mu_{pre} = \mu_{post}$) is that the pre-test mean and post-test mean are equal, the restriction on Model 3 which produces this state of affairs is,

$$a_1 = a_2$$

Imposing this restriction on Model 3, we obtain Model 4a.

Model 4a: $$Y = a_0U + a_1X_1 + a_1X_2 + a_3X_3 + a_4P_1 + a_5P_2 + a_6P_3 + E_2$$

Collecting terms with like weights we obtain Model 4b.

Model 4b: $$Y = a_0U + a_1 (X_1 + X_2) + a_3X_3 + a_4P_1 + a_5P_2 + a_6P_3 + E_2$$

The design matrix for Model 4b appears in Figure 4.

---------------------------
Insert Figure 4 Here
---------------------------

As in the first example, Model 3 and Model 4b, will each have an associated $R^2$. The difference between these two $R^2$'s serves in this case as a measure of the contribution of the difference between the pre-test and post-test to accounting for variance in the criterion variable. The significance of this difference may again be tested using the F-ratio.

NUMERICAL SOLUTION USING THE DATA IN TABLE 3

Table 4 contains the values of the regression weights, $R^2$'s, and the F-ratio observed in the statistical test of the restrictions made on Model 3. Note that there are 5 linearly independent vectors in Model 3, the Full Model, and 4 linearly independent vectors in Model 4b, the restricted model. Once again, the restriction placed on Model 3 to obtain Model 4b resulted in one linearly dependent vector being removed from the model.

This restriction resulted in a reduction in $R^2$ of .818 from Model 3 to Model 4b. The resulting F-ratio was 51.86 (with 1 and 4 degrees of freedom), which was significant given an alpha of .05 (one tailed). Therefore, it can be concluded that the statistical hypothesis can be rejected, and the research hypothesis that the Post-test mean is greater than the Pre-test mean can be accepted as tenable.

## A Second Question

A second question of interest to a researcher using this type of design might be one which compares one of the tests with the average of the other two tests. For example, a researcher might want to know,

> Is the average of the post-test mean and LTR mean greater than the pre-test mean?

This question would be answered by testing the statistical hypothesis that:

> The average of the post-test mean and the LTR mean is equal to the pre-test mean,

or,

$$\frac{\mu_{post} + \mu_{LTR}}{2} = \mu_{pre}$$

The Full Model which is used to test this hypothesis is Model 3, which reflects all of the information in Table 3. In this case however, the restriction on that model must take into account weights $a_1$, $a_2$, and $a_3$, which reflect the pre-test mean, post-test mean, and LTR mean. The restriction which reflects the statistical hypothesis in this case is

$$a_1 = \frac{a_2 + a_3}{2}$$

This restriction can be represented as

$$a_1 = \tfrac{1}{2}a_2 + \tfrac{1}{2}a_3$$

Imposing this restriction on Model 3, we obtain Model 5a.

Model 5a: $Y = a_0U + (\frac{1}{2}a_2 + \frac{1}{2}a_3)\ X_1 + a_2X_2 + a_3X_3 + a_4P_1 + a_5P_2 + a_6P_3 + E_3$

since,
$$(\tfrac{1}{2}a_2 + \tfrac{1}{2}a_3)\ X_1$$
$$= (\tfrac{1}{2}a_2X_1) + (\tfrac{1}{2}a_3X_1)$$
$$= a_2(\tfrac{1}{2}X_1) + a_3(\tfrac{1}{2}X_1)$$

Simplifying Model 5a, we obtain Model 5b.

Model 5b: $Y = a_0U + a_2(\frac{1}{2}X_1) + a_3(\frac{1}{2}X_1) + a_2X_2 + a_3X_3 + a_4P_1 + a_5P_2 + a_6P_3 + E_3$

Collecting terms with like weights we obtain Model 5c.

Model 5c: $Y = a_0U + a_2(\frac{1}{2}X_1 + X_2) + a_3(\frac{1}{2}X_1 + X_3) + a_4P_1 + a_5P_2 + a_6P_3 + E_3$

The design matrix for Model 5c appears in Figure 5.

--------------------------
Insert Figure 5 Here
--------------------------

In order to answer this research question, the significance of the difference between the $R^2$ associated with Model 3 (the Full Model) and the $R^2$ associated with Model 5c (the Restricted Model) is tested. The F-ratio is again used to test the significance of the difference between $R^2$'s.

NUMERICAL SOLUTION OF THE SECOND TYPE OF QUESTION USING THE DATA IN TABLE 3

Table 5 contains the values of the regression weights, $R^2$'s, and the F-Ratio observed in the restrictions made on Model 3 to obtain Model 5c. As before, there were 5 linearly independent vectors in Model 3, and the restriction placed upon the Full Model restricted out 1 linearly independent vector, leaving 4 linearly independent vectors in Model 5c, the restricted model.

Since the research question posed in this instance was a directional question, before interpreting the results it is necessary to ensure that the results are in the hypothesized direction. This can be done directly from the data in Table 3. In this data it can be seen that for each individual the mean of the Post-test and LTR-test scores were higher than the Pre-test scores. In this case, then, the results are in the hypothesized direction and may be interpreted as in Table 5.

The restriction placed on Model 3 to produce Model 5c resulted in a decrease in $R^2$ of .442. The resulting F-Ratio was 27.99 with 1 and 4 degrees of freedom, which was significant given an alpha of .05. With these results the statistical hypothesis may be rejected and the research hypothesis that the mean of the post-test and the LTR test scores is greater than the mean of the pre-test scores may be retained as tenable.

## III. A MULTI-GROUP AND MULTI-TRIAL DESIGN

For this demonstration, assume that the design to be analyzed contains three groups of subjects (a control group, and two experimental groups), and each group is measured on a pretest and a post-test. This then becomes a three groups by two trials analysis. The data that will be used to demonstrate a numerical solution for this design appear in Table 6.

Statistically, this design is a two component design, in that the total variation in the dependent variable is partitioned into a within subjects component and a between subjects component.

Any test of group differences would be treated as a between subjects contrast, where as any test of trial (pre-post) differences would be treated as a within subjects contrast. The purpose of this generalized partition of the total variance of the dependent variable is that the expected mean squares required in the formation of the F-ratios take different forms depending upon the nature of the contrast tested (Winer, 1962, p. 303). An error term based on subject differences within groups is used to test contrasts between groups, and an error term based on the interaction of subjects by trials is used to test contrasts across trials.

For the design analyzed here the following sources of variation can be isolated:

1. Groups- The variation in the dependent variable attributable to group differences.

2. Trials- The variation in the dependent variable attributable to trial differences.

3. Groups X Trials- The variation in the dependent variable attributable to the interaction of group and trial effects.

4. Subjects Within Groups- The variation in the dependent variable attributable to subjects' deviations from their group means.

5. Trials X Subjects Within Groups- The variation in the dependent variable attributable to the interaction of subjects and trials.

Each of these components of variance are independent, and when the sum of squares for each component are added they will equal the total sum of squares for the dependent variable.

The net effect of the requirement to partition the total variance into two global sources (between subjects and within subjects) is that two full models are required to analyze data from this design. One model will be used to analyze repeated measures contrasts and the other model will be used to analyze the between groups contrasts.

Testing Trial and Groups X Trial Effects

A model that could serve as a full model for repeated measures contrasts (trials and groups x trials) is

Model 6 $Y=a_0U+a_1X_1+a_2X_2+a_3X_3+a_4X_4+a_5X_5+a_6X_6+a_7P_1+a_8P_2+a_9P_3+a_{10}P_4+a_{11}P_5+a_{12}P_6+E_6$

where $Y$ = a vector containing the values of the dependent variable
$U$ = a unit vector
$X_1$ = a vector containing a 1 if the dependent variable observation is from the control group on the pre-test, 0 otherwise
$X_2$ = 1 if observation is from the control group on the post-test, 0 otherwise
$X_3$ = 1 if observation is from the 1st experimental group on the pre-test, 0 otherwise
$X_4$ = 1 if observation is from the 1st experimental group on the post-test, 0 otherwise
$X_5$ = 1 if observation is from the 2nd experimental group on the pre-test, 0 otherwise
$X_6$ = 1 if observation is from the 2nd experimental group on the post-test, 0 otherwise
$P_i(i=1,6)$ = vectors containing 1's if the observation is from subject i, 0 otherwise
$E_1$ = the error vector
$a_0 \ldots a_{12}$ = least squares weighting coefficients.

Figure 6 contains the vector representation of Model 6. With a model this complex it becomes rather difficult to determine the number of linearly independent vectors, therefore one has to approach the problem in stages. Consider first the unit vector, U, and vectors $X_1$ through $X_6$. Proceeding from left to right in Figure 6, vectors U through $X_5$ can be seen to be linearly independent, but vector $X_6$ can be obtained as follows:

$$X_6 = U - (X_1 + X_2 + X_3 + X_4 + X_5)$$

Hence $X_6$ is linearly dependent. Thus far 6 linearly independent vectors have been noted. Now proceeding from left to right starting with $P_1$, $P_1$ is seen to be linearly independent of U through $X_6$. However $P_2$ can be obtained as follows:

$$P_2 = X_1 + X_2 - P_1$$

Hence $P_2$ is linearly dependent. Considering $P_3$ and $P_4$ gives a similar result; $P_3$ is linearly independent, but $P_4$ may be obtained as follows:

$$P_4 = X_3 + X_4 - P_3$$

Similarly, of $P_5$ and $P_6$, $P_5$ is linearly independent and $P_6$ is linearly dependent. For the vectors $P_1$ through $P_6$, only 3 linearly independent vectors are present. Consequently, for Model 6 there are a total of 9 linearly independent vectors.

The hypothesis that we test when trial differences are the contrast of concern is

$$Ho: \mu_{pretest} = \mu_{post\text{-}test}$$

In order to generate a restricted model that will allow this hypothesis to be tested, the following restriction must be imposed on Model 5:

$$\frac{a_1 + a_3 + a_5}{3} = \frac{a_2 + a_4 + a_6}{3}$$

In words, this restirction means we are forcing the pre-test mean averaged over groups to equal the post-test mean averaged over groups. This restriction can be restated in the following form, solving for $a_1$

$$a_1 = a_2 + a_4 + a_6 - a_3 - a_5$$

When this restriction is imposed upon Model 5, we obtain Model 7a.

Model 7a: $Y=a_0U+(a_2+a_4+a_6-a_3-a_5)X_1+a_2X_2+a_3X_3+a_4X_4+a_5X_5+a_6X_6+a_7P_1+a_8P_2+a_9P_3+a_{10}P_4+a_{11}P_5+a_{12}P_6+E_7$ .

Distributing $X_1$ over the weights contained in the parentheses and then collecting terms with like weights we obtain Model 7b as the restricted model.

Model 7b: $Y=a_0U+a_2(X_1+X_2)+a_3(X_3-X_1)+a_4(X_1+X_4)+a_5(X_5-X_1)+a_6(X_1+X_6)+a_7P_1+a_8P_2+a_9P_3+a_{10}P_4+a_{11}P_5+a_{12}P_6+E_7$

Figure 7 contains the design matrix for model 7b. Again, it is difficult to determine at a glance the number of linearly independent vectors in Model 7b, but if one proceeds in stages the process is simplified. Consider first the vectors U through $(X_1+X_6)$. The vectors U, $(X_1+X_2)$, $(X_3-X_1)$, $(X_1+X_4)$ and $(X_5-X_1)$ are all linearly independent. However the vector $(X_1+X_6)$ is linearly dependent, since

$$(X_1+X_6)=U-(X_1+X_2)+(X_3-X_1)+(X_1+X_4)+(X_5-X_1)$$

Therefore there are 5 linearly independent vectors in the set of vectors containing U through $(X_1+X_6)$. As in Model 6. there are 3 linearly independent vectors composed of $P_1$ through $P_6$. Consequently Model 7b contains a total of 8 linearly independent vectors.

Model 6 can also be used as a full model to test the groups X trials interaction question. The hypothesis of no interaction can be stated as,

$$Ho: \mu_2 - \mu_1 = \mu_4 - \mu_3 = \mu_6 - \mu_5$$

where $\mu_1$ = population pretest mean for group 1

$\mu_2$ = population post-test mean for group 1

$\mu_3$ = population pre-test mean for group 2

$\mu_4$ = population post-test mean for group 2

$\mu_5$ = population pre-test mean for group 3

$\mu_6$ = population post-test mean for group 3

This hypothesis states that the trial differences (post-test-pre-test) are equal for the three groups. The restriction imposed on model 6 to obtain a restricted model for testing the groups X trials interaction is,

$$a_2 - a_1 = a_4 - a_3 = a_6 - a_5$$

This restriction can be restated as follows by solving for $a_2$ and $a_6$:

$$a_2 = a_1 + a_4 - a_3$$

$$a_6 = a_5 + a_4 - a_3$$

Imposing these restrictions on model 6 we obtain model 8a.

Model 8a: $Y = a_0U + a_1X_1 + (a_1+a_4-a_3)\ X_2 + a_3X_3 + a_4X_4 + a_5X_5 + (a_5+a_4-a_3)\ X_6$

$+ a_7P_1 + a_8P_2 + a_9P_3 + a_{10}P_4 + a_{11}P_5 + a_{12}P_6 + E_8$

Distributing vectors $X_2$ and $X_6$ over the weights contained in the parentheses and then collecting terms with like weights we obtain model 8b as a restricted model.

Model 8b: $Y = a_0U + a_1(X_1+X_2) + a_3(X_3-X_2-X_6) + a_4(X_2+X_4+X_6) + a_5(X_5+X_6) + a_7P_1$

$+ a_8P_2 + a_9P_3 + a_{10}P_4 + a_{11}P_5 + a_{12}P_6 + E_3$

Figure 8 contains the design matrix for model 8b. The determination of the number of linearly independent vectors in model 8b is, again, somewhat difficult, but if the procedures described for models 6 and 7b are followed for model 8b one finds that there are 7 linearly independent vectors in model 7b. Hence, the groups X trials restriction restricted out two independent vectors from model 6.

THE EQUIVALENCE OF THE GROUPS X TRIALS TEST AND A DIFFERENCE SCORE ANALYSIS

The groups X trials interaction analysis can be shown to be indentical to an analysis of difference scores. That is, the same results would be obtained if the following full model was used.

$$\text{Model 9:} \quad Y_0 = a_0U + a_1X_1 + a_2X_2 + a_3X_3 + E_9$$

where $Y_0$ = a vector of difference scores (post-test - pre-test)

$U$ = a unit vector

$X_1$ = a vector containing a 1 if the difference score is for a subject in the control group, 0 otherwise

$X_2$ = a vector containing a 1 if the difference score is for a subject in the first experimental group, 0 otherwise

$X_3$ = a vector containing a 1 if the difference score is for a subject in the second experimental group, 0 otherwise

$E$ = an error vector

$a_0 \ldots a_3$ = least squares weights.

This model can be obtained from model 6 by subtracting rows of the design matrix corresponding to pre-test observations, from the rows of the design matrix corresponding to post-test observations. Row 1 in Figure 9 was obtained by subtracting row 1 (pre-test for subject 1 in group 1) from row 3 (post-test for subject 1 in group 1) in the design matrix for model 6 (Figure 6). By repeating this procedure for each subject we obtain the design matrix in Figure 9. Note that the subject vectors ($P_1 \ldots P_6$) cancel out, and that vectors $X_1 \ldots X_6$ become three sets of redundent group vectors.

The following equalities can be noted in Figure 9:

$$X_1 = -X_2$$
$$X_3 = -X_4$$
$$X_5 = -X_6$$

Consequently $X_1$, $X_3$ and $X_5$ can be removed from the model because they are linearly dependent. The design matrix in Figure 9 is therefore equivalent to the design matrix implied by model 9.

The groups X trials interaction test can be achieved with model 9 as a full model, by testing the hypothesis that all three experimental groups have a common mean difference score.

Symbolically, the hypothesis tested is,

$$Ho: \mu_{D1} = \mu_{D2} = \mu_{D3}$$

where $\mu_{D1}$ = population mean difference score for the control group.

$\mu_{D2}$ = population mean difference score for the first experimental group.

$\mu_{D3}$ = population mean difference score for the second experimental group.

The restriction imposed on model 9 to test this hypothesis is,

$$a_1 = a_2 = a_3$$

When this restriction is placed on the full model the following restricted model is obtained:

Model 10: $Y_D = a_0 U + E_{10}$

The equivalence of the groups X trials interaction test and the test of differences between mean difference scores can be further established by calculating a difference score model by subtracting pre-test and post-test rows in the design matrix for model 8 (Figure 8). This has been done in Figure 10. Note that in Figure 10 the subject vectors $(P_1 \ldots P_6)$ have zeroed out, and so have the vectors $(X_1 + X_2)$ and $(X_5 + X_6)$. The two vectors which have non-zero elements $(X_3 - X_2 - X_6)$ and $(X_2 + X_4 + X_6)$, are linearly dependent on the unit vector. Hence the design matrix in Figure 10 represents model 10. It follows that the F-ratio obtained comparing models 7 and 8b will equal the F-ratio obtained comparing models 9 and 10. This will be demonstrated numerically in a later section of this paper.

TESTING BETWEEN GROUPS CONTRASTS

A complete general linear model analysis would require the removal of the subject vectors $(P_1 \ldots P_6)$ in model 6 and replacing them with a set of vectors coding the subject by trial effect. This procedure can become very complex, and fortunately we can conduct this analysis by summing trial scores for each subject and then conducting a simple between groups analysis of variance on the trial sums.

The full model for between groups comparisons is very similar to model 9. The difference between the two models is that the dependent variable is a trial sum, rather than a trial difference. The full model for testing group differences is,

Model 11: $$Y_s = a_0U + a_1X_1 + a_2X_2 + a_3X_3 + E_{11}$$

where $Y_s$ = a vector containing the sum of scores for pre and post test observations for each subject

$U$ = a unit vector

$X_1$ = a vector containing a 1 if the criterion sum is from a subject in the control group, 0 otherwise

$X_2$ = a vector containing a 1 if the criterion sum is from a subject in the first experimental group, 0 otherwise

$X_3$ = a vector containing a 1 if the criterion sum is from a subject in the second experimental group, 0 otherwise

$E_{11}$ = the error vector

$a_0 \ldots a_3$ = least squares regression weights

The design matrix implied by model 11 is shown in Figure 11. As can be seen in Figure 11, there are three linearly independent vectors in model 11.

In order to test the hypothesis that the groups differ, the following statistical hypothesis is posed:

$$Ho: \mu_{S1} = \mu_{S2} = \mu_{S3}$$

where $\mu_{S1}$ = the population mean of the trial sum for the control group

$\mu_{S2}$ = the population mean of the trial sum for the first experimental group

$\mu_{S3}$ = the population mean of the trial sum for the second experimental group.

The restriction imposed on model 11 required to test this hypothesis is,

$$a_1 = a_2 = a_3$$

The restricted model which follows when this restriction is imposed is,

Model 12 : $Y_s = a_0U + E_{12}$

Model 12 is a model containing only the unit vector.

NUMERICAL SOLUTIONS FOR SECTION III

The numerical results obtained from the analysis of the data in Table 6 are presented in Tables 7, 8 and 9. Table 7 contains the results of the analysis for the within subjects components of the problem. Models 6, 7b and 8b were the models used as the full model, the restricted model for the trial effect, and the restricted model for the groups x trial interaction effect respectively.

The F ratios presented in Table 7 obtained for the trial and the groups x trial effect are both significant ($\alpha = .05$). The results obtained for the groups x trial effect were the same when original observations were used (Table 7) and when difference scores were used (Table 8). This finding serves to further demonstrate the equivalence of the difference score analysis and the groups x trial interaction analysis.

The results of the analysis of the groups effect is presented in Table 9. As can be seen in Table 9 the groups effect was also significant ($\alpha = .05$).

SUMMARY

The purpose of this paper was to demonstrate the use of the general linear model to answer research questions with designs containing repeated measures on subjects. Three types of design were presented, 1. a one group-two trial design, 2. a one group-multi-trial design and 3. a multi-group-two trial design. For each design a series of research questions were posed, a full linear model was stated, restrictions consistent with the hypotheses to be tested were stated, and these restrictions were then imposed on the full model to obtain statistical tests of the hypotheses. For each design numerical solutions were provided which demonstrated applications of the generalized procedures discussed in the text of the paper.

# REFERENCES

Glass, G., Peckham, P. D. and Sanders, J. R., Consequenses of failure to meet assumptions underlying the fixed effects analysis of variance and covariance. Review of Educational Research, 1972, 42 (3), p. 237-288.

Winer, B. J., Statistical principles in experimental design. New York: McGraw-Hill, 1962.

Table 1

Sample Data for a One Group-Two Trial Analysis

| Person | Pre-test Political Attitudes | Post-Test Political Attitudes |
|---|---|---|
| 1 | 15 | 17 |
| 2 | 18 | 18 |
| 3 | 12 | 18 |
| 4 | 20 | 24 |

Table 2

Least Squares Weights and RSQ's for Models 1 and 2C Derived from Data in Table 1

| Vector | Model 1 Weights | Model 2C Weights |
|---|---|---|
| U | 17.4 | 15.9 |
| $X_1$ | -3.0 | ---* |
| $X_2$ | 0.0 | ---* |
| $P_1$ | .1 | .1 |
| $P_2$ | 2.1 | 2.1 |
| $P_3$ | -.9 | -.9 |
| $P_4$ | 6.2 | 6.1 |
| RSQ | .883 | .673 |
| Number of linearly Independent Vectors | $\ell_1 = 5$ | $\ell_2 = 4$ |

$$F = \frac{(.883 - .673)/(5-4)}{(1 - .883)/8-5} = 5.40, \text{ df} = 1,3$$

* Vectors $X_1$ and $X_2$ are not in Model 2C.

Table 3

Sample Data for a One Group - Multiple Trial Analysis with More than Two Trials

| Person | Pre-Test | Post-Test | Long-Term Retention Test |
|---|---|---|---|
| 1 | 5 | 15 | 10 |
| 2 | 8 | 20 | 8 |
| 3 | 4 | 15 | 8 |

Table 4

Least Squares Weights and $R^2$'s for Models II.1 and II.2b derived from Data in Table 3

| Vector | Model 1 Weights | Model 2b Weights |
|---|---|---|
| U | 8.32 | 7.33 |
| $X_1$ | -2.99 | 2.50 - vector $(X_1 + X_2)$ |
| $X_2$ | 7.99 | |
| $X_3$ | 0.0 | -.0 |
| $P_1$ | 0.0 | .99 |
| $P_2$ | 2.02 | 2.99 |
| $P_3$ | -.98 | 0.0 |
| $R^2$ | .937 | .119 |
| Number of linearly independent vectors | 5 | 4 |

$$F = \frac{(.937 - .119)/(5-4)}{(1 - .937)/(9-5)} = 51.86, \text{ df } 1, 4$$

Directional probability = .001

Table 5

Least Squares Weights and $R^2$'s for Model II.1 and II.3c from Data in Table 3

| Vector | Model II.1 Weights | Model II.3c Weights |
|---|---|---|
| U | 8.32 | 5.01 |
| $X_1$ | -2.99 | --- |
| $X_2$ | 7.99 | --- |
| $X_3$ | 0.0 | --- |
| $P_1$ | 0.0 | .99 |
| $P_2$ | 2.02 | 2.98 |
| $P_3$ | -.98 | 0.0 |
| $(\frac{1}{2}X_1 + X_2)$ | --- | .66 |
| $(\frac{1}{2}X_1 + X_3)$ | --- | 0.0 |
| $R^2$ | .937 | .495 |
| Number of linearly independent vectors | 5 | 4 |

$$F = \frac{(.937 - .498)/(5 - 4)}{(1 - .937)/9 - 5} = 27.99, \text{ df } 1, 4$$

Directional probability = .003

Table6

Sample Data for a Three Group-Two Trial Analysis

| Group | Subject | Trial: Pre-Test Score | Trial: Post-Test Score |
|---|---|---|---|
| Control Group | 1 | 3 | 4 |
| | 2 | 4 | 3 |
| Experimental Group 1 | 3 | 5 | 9 |
| | 4 | 4 | 7 |
| Experimental Group 2 | 5 | 3 | 12 |
| | 6 | 5 | 15 |

Table 7

Numerical Solutions for Models 6, 7b and 8b to Analyze the Trial Effect and the Groups x Trial Interaction (Number of Observations = 12)

| Model 6 Vectors | Model 6 Weights | Model 7b Vectors | Model 7b Weights | Model 8b Vectors | Model 8b Weights |
|---|---|---|---|---|---|
| $X_1$ | 0 | $X_1+X_2$ | -4.06 | $X_1+X_2$ | -3.75 |
| $X_2$ | 0 | $X_3-X_1$ | .52 | $X_3-X_2-X_6$ | - .23 |
| $X_3$ | .15 | $X_1+X_4$ | .29 | $X_2+X_4+X_6$ | 4.10 |
| $X_4$ | 3.66 | $X_5-X_1$ | -.10 | $X_5+X_6$ | .16 |
| $X_5$ | -.07 | $X_1+X_6$ | 5.05 | | |
| $X_6$ | 9.43 | | | | |
| $P_1$ | 0 | $P_1$ | 0 | $P_1$ | 0 |
| $P_2$ | 0 | $P_2$ | 0 | $P_2$ | 0 |
| $P_3$ | 1.60 | $P_3$ | 1.51 | $P_3$ | 0 |
| $P_4$ | .11 | $P_4$ | 0 | $P_4$ | -1.53 |
| $P_5$ | -.65 | $P_5$ | -.35 | $P_5$ | .08 |
| $P_6$ | 1.81 | $P_6$ | 2.14 | $P_6$ | 2.58 |
| U | 3.49 | U | 5.39 | U | 5.09 |

| | Model 6 | Model 7b | Model 8b |
|---|---|---|---|
| $R^2$ = | .991 | .655 | .716 |
| Number of linearly independent vectors | 9 | 8 | 7 |

| | Trial Effect | Groups x Trial Effect |
|---|---|---|
| F ratio | 112.6 | 46.2 |
| degrees of Freedom | 1, 3 | 2, 3 |

Table 8

Numerical Solutions for Models 9 and 10 to Analyze the Groups x Trials Interaction Using Difference Scores (Number of Observations = 6)

| | Model 9 | | Model 10 | |
|---|---|---|---|---|
| | Vectors | Weights | Vectors | Weights |
| | $X_1$ | -3.50 | | |
| | $X_2$ | 0 | | |
| | $X_3$ | 6.00 | | |
| | U | 3.50 | U | 4.33 |
| $R^2$ | .969 | | .00 | |
| Number of Linearly Independent Vectors | 3 | | 1 | |

| | Groups x Trial Effect |
|---|---|
| F ratio | 46.2 |
| degrees of Freedom | 2, 3 |

Table 9

Numerical Solutions for Models 11 and 12
to Analyze the Groups Effect
(Number of Observations = 6)

| | Model 11 Vectors | Model 11 Weight | Model 12 Vector | Model 12 Weight |
|---|---|---|---|---|
| | $X_1$ | -5.50 | | |
| | $X_2$ | 0 | | |
| | $X_3$ | 5.00 | | |
| | U | 12.50 | U | 12.33 |
| $R^2$ | | .866 | | .00 |
| Number of Linearly Independent Vectors | | 3 | | 1 |

| | Groups Effect |
|---|---|
| F ratio | 9.73 |
| degrees of Freedom | 2,3 |

| $Y$ | $U$ | $X_1$ | $X_2$ | $P_1$ | $P_2$ | $P_3$ | $P_4$ |
|---|---|---|---|---|---|---|---|
| 15 | 1 | 1 | 0 | 1 | 0 | 0 | 0 |
| 18 | 1 | 1 | 0 | 0 | 1 | 0 | 0 |
| 12 | 1 | 1 | 0 | 0 | 0 | 1 | 0 |
| 20 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| 17 | 1 | 0 | 1 | 1 | 0 | 0 | 0 |
| 18 | 1 | 0 | 1 | 0 | 1 | 0 | 0 |
| 18 | 1 | 0 | 1 | 0 | 0 | 1 | 0 |
| 24 | 1 | 0 | 1 | 0 | 0 | 0 | 1 |

Figure 1. The design matrix implied by Model 1.

| $Y$ | $U$ | $P_1$ | $P_2$ | $P_3$ | $P_4$ |
|---|---|---|---|---|---|
| 15 | 1 | 1 | 0 | 0 | 0 |
| 18 | 1 | 0 | 1 | 0 | 0 |
| 12 | 1 | 0 | 0 | 1 | 0 |
| 20 | 1 | 0 | 0 | 0 | 1 |
| 17 | 1 | 1 | 0 | 0 | 0 |
| 18 | 1 | 0 | 1 | 0 | 0 |
| 18 | 1 | 0 | 0 | 1 | 0 |
| 24 | 1 | 0 | 0 | 0 | 1 |

Figure 2. The design matrix for Model 2C.

| $Y$ | $U$ | $X_1$ | $X_2$ | $X_3$ | $P_1$ | $P_2$ | $P_3$ |
|---|---|---|---|---|---|---|---|
| 5 | 1 | 1 | 0 | 0 | 1 | 0 | 0 |
| 8 | 1 | 1 | 0 | 0 | 0 | 1 | 0 |
| 4 | 1 | 1 | 0 | 0 | 0 | 0 | 1 |
| 15 | 1 | 0 | 1 | 0 | 1 | 0 | 0 |
| 20 | 1 | 0 | 1 | 0 | 0 | 1 | 0 |
| 15 | 1 | 0 | 1 | 0 | 0 | 0 | 1 |
| 10 | 1 | 0 | 0 | 1 | 1 | 0 | 0 |
| 8 | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| 8 | 1 | 0 | 0 | 1 | 0 | 0 | 1 |

Figure 3. The design matrix implied by Model 3.

| $Y$ | $U$ | $(X_1+X_2)$ | $X_3$ | $P_1$ | $P_2$ | $P_3$ |
|---|---|---|---|---|---|---|
| 5 | 1 | 1 | 0 | 1 | 0 | 0 |
| 8 | 1 | 1 | 0 | 0 | 1 | 0 |
| 4 | 1 | 1 | 0 | 0 | 0 | 1 |
| 15 | 1 | 1 | 0 | 1 | 0 | 0 |
| 20 | 1 | 1 | 0 | 0 | 1 | 0 |
| 15 | 1 | 1 | 0 | 0 | 0 | 1 |
| 10 | 1 | 0 | 1 | 1 | 0 | 0 |
| 8 | 1 | 0 | 1 | 0 | 1 | 0 |
| 8 | 1 | 0 | 1 | 0 | 0 | 1 |

Figure 4. The design matrix implied by Model 4B.

| $Y$ | $U$ | $(\frac{1}{2}X_1 + X_2)$ | $(\frac{1}{2}X_1 + X_3)$ | $P_1$ | $P_2$ | $P_3$ |
|---|---|---|---|---|---|---|
| 5 | 1 | .5 | .5 | 1 | 0 | 0 |
| 8 | 1 | .5 | .5 | 0 | 1 | 0 |
| 4 | 1 | .5 | .5 | 0 | 0 | 1 |
| 15 | 1 | 1 | 0 | 1 | 0 | 0 |
| 20 | 1 | 1 | 0 | 0 | 1 | 0 |
| 15 | 1 | 1 | 0 | 0 | 0 | 1 |
| 10 | 1 | 0 | 1 | 1 | 0 | 0 |
| 8 | 1 | 0 | 1 | 0 | 1 | 0 |
| 8 | 1 | 0 | 1 | 0 | 0 | 0 |

Figure 5. The design matrix implied by Model 5C.

| observation | Y | U | $X_1$ | $X_2$ | $X_3$ | $X_4$ | $X_5$ | $X_6$ | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_6$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 3 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | 4 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 3 | 4 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 4 | 3 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 5 | 5 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 6 | 4 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 7 | 9 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 8 | 7 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 9 | 3 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 10 | 5 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11 | 12 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| 12 | 15 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |

Figure 6. The design matrix implied by Model 6.

| Y | U | $(X_1+X_2)$ | $(X_3-X_1)$ | $(X_1+X_4)$ | $(X_5-X_1)$ | $(X_1+X_6)$ | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_6$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 1 | 1 | -1 | 1 | -1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| 4 | 1 | 1 | -1 | 1 | -1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| 4 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 3 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 5 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 4 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 9 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 7 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 3 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 5 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 12 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| 15 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |

Figure 7. The design matrix for Model 7b.

| observation | Y | U | $(X_1+X_2)$ | $(X_3-X_2-X_6)$ | $(X_2+X_4+X_6)$ | $(X_5+X_6)$ | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_6$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 3 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | 4 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 3 | 4 | 1 | 1 | -1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 4 | 3 | 1 | 1 | -1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 5 | 5 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 6 | 4 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 7 | 9 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 8 | 7 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 9 | 3 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| 10 | 5 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| 11 | 12 | 1 | 0 | -1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| 12 | 15 | 1 | 0 | -1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |

Figure 8. The design matrix for Model 8b.

| Observation Differences Required to Obtain Gain Scores | $Y_D$ Difference Scores | U* | $X_1$ | $X_2$ | $X_3$ | $X_4$ | $X_5$ | $X_6$ | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_6$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3-1 | 4 - 3 = 1 | 1 | -1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4-2 | 3 - 4 = -1 | 1 | -1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7-5 | 9 - 5 = 4 | 1 | 0 | 0 | -1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8-6 | 7 - 4 = 3 | 1 | 0 | 0 | -1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11-9 | 12 - 3 = 9 | 1 | 0 | 0 | 0 | 0 | -1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12-10 | 15 - 5 = 10 | 1 | 0 | 0 | 0 | 0 | -1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |

Figure 9. The design matrix derived from Model 6 to analyze difference scores.

* The unit vector is included here even though it cancels ouf of the model since the unit vector is included by most regression programs for every model.

| Observation Differences Required to Obtain Difference Scores | $Y_D$ Difference Scores | U* | $(X_1+X_2)$ | $(X_3-X_2-X_6)$ | $(X_2+X_4+X_6)$ | $(X_5+X_6)$ | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_6$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3-1 | 4 - 3 = 1 | 1 | 0 | -1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4-2 | 3 - 3 = 0 | 1 | 0 | -1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7-5 | 9 - 5 = 4 | 1 | 0 | -1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8-6 | 7 - 4 = 3 | 1 | 0 | -1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11-9 | 12 - 3 = 9 | 1 | 0 | -1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12-10 | 15 - 5 = 10 | 1 | 0 | -1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Figure 10. The design matrix derived from Model 8b to analyze difference scores.

* The unit vector is included here even though it cancels out of the Model since the unit vector is included by most regression programs for every Model.

| $Y_s$ (trial sum) | U | $X_1$ | $X_2$ | $X_3$ |
|---|---|---|---|---|
| 7 | 1 | 1 | 0 | 0 |
| 7 | 1 | 1 | 0 | 0 |
| 14 | 1 | 0 | 1 | 0 |
| 11 | 1 | 0 | 1 | 0 |
| 15 | 1 | 0 | 0 | 1 |
| 20 | 1 | 0 | 0 | 1 |

Figure 11. The design matrix implied by Model 11.