METHODS OF ANALYZING THREE-PERIOD TWO-TREATMENT CROSSOVER DESIGNS WITH INCOMPLETE DATA

by

C. David Shen

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1451

August 1983
ACKNOWLEDGEMENTS

I wish to thank my advisor, Dr. Ronald W. Helms, for his deep patience and constant encouragement in the preparation of this study. I would also like to thank the other members of my committee, Drs. Dana Quade, Gary G. Koch, James Hosking and Elliot Cramer for their suggestions and ideas.

Generous financial support for this study has been provided by the National Institute of Mental Health.

Finally, I would like to express my thanks to my wife, Dannie, and my parents, for their continuing patience and sacrifice during my study period in Chapel Hill.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS .............................................. ii
LIST OF TABLES ................................................. v

CHAPTER I: INTRODUCTION AND LITERATURE REVIEWS ............. 1

1.1. Introduction .................................................. 1
1.2. Literature Reviews ........................................... 4
   1.2.1 Crossover designs - designs .......................... 5
   1.2.2 Crossover designs - analysis ........................ 8
   1.2.3 Analysis of incomplete crossover data ............... 12

1.3. Linear Model Notations and Abbreviations ................. 21
   1.3.1 GLUM - General linear univariate model ............. 21
   1.3.2 GLMM - General linear multivariate model ........... 22
   1.3.3 MGLMM - More general linear multivariate model .... 24

CHAPTER II: LINEAR MODEL ANALYSIS OF COMPLETE DATA ......... 27

2.1. Traditional Combined-Period Model ......................... 27
   2.1.1 The model .............................................. 27
   2.1.2 Estimation and hypothesis testing .................... 30
   2.1.3 Numerical example .................................... 31
   2.1.4 Inter-subject analysis ................................ 37

2.2. Separate-Period Univariate Model .......................... 40
   2.2.1 The model .............................................. 40
   2.2.2 Estimation and hypothesis testing .................... 42
   2.2.3 Numerical example .................................... 43
   2.2.4 Discussion ............................................. 46

2.3. Separate-Period Multivariate Model ....................... 46
   2.3.1 The model .............................................. 46
   2.3.2 Estimation and hypothesis testing .................... 48
   2.3.3 Numerical example .................................... 49
CHAPTER III: MULTIVARIATE ANALYSIS OF INCOMPLETE DATA .... 52

3.1. Introduction 52
3.2. Incomplete Data Models 53
3.3. Estimation and Hypothesis Testing 55
   3.3.1 Σ known and unrestricted 55
   3.3.2 Σ unknown and unrestricted 57
3.4. Numerical Example 60
3.5. Algorithmic Issues 63

CHAPTER IV: MULTIVARIATE ANALYSIS OF CROSSOVER DATA WITH
             SELECTED COVARIANCE MATRICES ....... 71

4.1. Introduction 71
4.2. Covariance Models Selection 71
   4.2.1 The homogeneous variance models 72
   4.2.2 The heterogeneous variance models 75
4.3. Numerical Example 79
4.4. Discussion 81

CHAPTER V: SUMMARY AND SUGGESTIONS FOR FURTHER RESEARCH .... 89

5.1. Summary 89
5.2. Suggestions for Further Research 91

APPENDIX A ........................................... 93
APPENDIX B ........................................... 94
BIBLIOGRAPHY ......................................... 101
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Schematic Representation of Some Crossover Designs.</td>
<td>19</td>
</tr>
<tr>
<td>1.2</td>
<td>Treatment Specifications for Eight-Sequence Three-Period Two-Treatment Crossover Design.</td>
<td>32</td>
</tr>
<tr>
<td>2.2</td>
<td>List of Complete Mungbean Data.</td>
<td>33</td>
</tr>
<tr>
<td>2.3</td>
<td>List of Transformed Data $A_1^Y$ and LTFR Design Matrix $A_1^X$.</td>
<td>34</td>
</tr>
<tr>
<td>2.4</td>
<td>Full Rank Design Matrix $Z_2$ Reparametrized from $A_1^X$.</td>
<td>35</td>
</tr>
<tr>
<td>2.5</td>
<td>Transformation Matrix $H_{12}$.</td>
<td>36</td>
</tr>
<tr>
<td>2.6</td>
<td>Coefficients of Incidence Matrix $Q$ for Estimating Parameters from the Intra-subject Model.</td>
<td>38</td>
</tr>
<tr>
<td>2.7</td>
<td>OLS Estimators from the Intra-subject Model.</td>
<td>38</td>
</tr>
<tr>
<td>2.8</td>
<td>OLS Estimators from the Inter-subject Model.</td>
<td>39</td>
</tr>
<tr>
<td>2.9</td>
<td>Intra-subject ANOVA Table.</td>
<td>39</td>
</tr>
<tr>
<td>2.10</td>
<td>Inter-subject ANOVA Table.</td>
<td>39</td>
</tr>
<tr>
<td>2.11</td>
<td>Full Rank Design Matrix $X_3$ from Inter-subject Analysis.</td>
<td>40</td>
</tr>
<tr>
<td>2.12</td>
<td>Subscript of Cell Mean $u_{ptuv}$ within Period.</td>
<td>42</td>
</tr>
<tr>
<td>2.13</td>
<td>Separate-Periods ANOVA Table.</td>
<td>45</td>
</tr>
<tr>
<td>2.14</td>
<td>OLS Estimators and Bonferroni Confidence Intervals from Separate-Periods Analysis.</td>
<td>45</td>
</tr>
<tr>
<td>2.15</td>
<td>OLS Estimators from Separate-Period Multivariate Model.</td>
<td>51</td>
</tr>
<tr>
<td>3.1</td>
<td>List of Mungbean Data with 20% Missing.</td>
<td>66</td>
</tr>
<tr>
<td>3.2</td>
<td>Estimates of Parameters in MSPC Model from Three Least Squares Procedures.</td>
<td>67</td>
</tr>
<tr>
<td>3.3</td>
<td>Sample Variance-Covariance Matrices in MSPC Model.</td>
<td>68</td>
</tr>
<tr>
<td>3.4</td>
<td>The Comparisons of Standard Errors of Estimates from MSPC Model</td>
<td>68</td>
</tr>
</tbody>
</table>
3.5 Estimates of Secondary Parameters in MSCP Model .................. 69
3.6 The ANOVA Table in MSCP Model from Incomplete Mungbean Data ................. 69
3.7 Estimates of Parameters in MCPC Model .................................. 70
3.8 Sample Variance-Covariance/Correlation Matrix from MCPC Model .................. 70
3.9 The ANOVA Table in MCPC Model from Incomplete Mungbean Data .................. 70
4.1 Estimates of Covariance/Correlation Matrices from Selected Covariance Models .................. 83
4.2 Estimates of Primary and Secondary Parameters from Selected Covariance Models .................. 84
4.3 Test Statistics for Important Hypotheses in the Selected Covariance Models .................. 86
4.4 The Variance-Covariance Matrices Generated from Six Random Missing Samples in Mungbean Experiment .................. 87
4.5 Comparison of Wald Statistics from Three Different Missing Patterns .................. 88
4.6 Test Statistics for Important Hypotheses in the Heteroscedastic Models using Approximation Methods .................. 88
C. DAVID SHEN. Methods of Analyzing Three-Period Two-Treatment Crossover Designs with Incomplete Data. (Under the direction of RONALD W. HELMS.)

Some of the problems which affect traditional crossover designs and analyses are confounding (arising from both the incompleteness of traditional designs and from less-than-full-rank models used for their analysis), heteroscedasticity, and complications arising from combinations of missing data and intraclass correlations.

This research addresses some of these problems for two-treatment three-period crossover designs. After a review of the traditional designs and models we describe the use of a separate period univariate model which handles the confounding and missing data problems but does not permit within-subject between-periods comparisons. A standard multivariate model is described which handles confounding and permits within-subject comparisons but does not permit missing data. Both of these models produce finite (non-asymptotic) tests.

Application of Kleinbaum's more general linear multivariate model (MGLMM) and weighted least squares leads to an analysis which handles confounding, missing data and heteroscedasticity, and permits within-subject between-period comparisons. Estimators from this procedure are BAN and the Wald test statistics have asymptotic properties.

Finally, we consider the incomplete data crossover model in which the within-subject covariance matrix has known pattern. Weighted least squares estimators and Wald tests are derived for this situation.

Numerical examples are presented for each of the proposed methods.
sensitivity to treatment time may be largely reduced. When experimental subjects receive the same treatments but in different order, the design is often called a crossover design. Crossover designs are also known as change-over designs or switch-back trials in the literature. In many circumstances, the crossover design is related to a Latin square or sets of Latin squares, where Latin square columns represent sequences of treatment administration and rows represent periods of observation. The crossover design is also similar to the split-plot design, in that whole plots may be represented by subjects and subplots may be represented by treatment application period. When two treatments, A and B, are given such that each experimental subject receives either treatment A or treatment B in the first period, followed by treatment B or treatment A in the second period, respectively, the design is called a two-treatment, two-period, two-sequence crossover design. This design is particularly popular in pharmacological studies when studying the effect of a test drug versus a standard drug on a selected group of subjects.

One important issue in crossover design is the residual effect (or carryover effect). The residual effect is defined to be the effect of treatment lasting from the period of application into a subsequent treatment period. The residual effect can be observed as a learning effect or conditioning effect in many psychological and educational studies. One controversy about residual effect involves design efficiency. Grizzle (1965) presented an ANOVA linear model for this design and found that period effects are not estimable in the presence of residual effects. He also suggested that the estimation of treatment effects suffers a significant loss of efficiency if residual effects are simultaneously present in the model. Brown (1980) compares the two-treatment two-period two-sequence crossover design
with a completely randomized design and concludes that in the presence of residual effects to attain the same relative bias and the same power, the crossover design requires 10 times as many subjects as the completely randomized design.

The validity of statistical assumptions in the crossover design seems to be another important issue in the literature. Almost all crossover models assume no subject by treatment interaction effect. Furthermore, Hills and Armitage (1979) indicate that most models in crossover designs assume no treatment by period interactions, i.e., the response to a treatment during the second period should not be influenced by the treatment which was given in the first period. Brown (1980) suggests that crossover designs not be used if there is any a priori evidence of treatment by period interaction. Moreover, many authors, especially those prior to Grizzle (1965) assumed the data to be independent, identically and normally distributed, or at least equally correlated. These assumptions are worrisome since the correlations among within-subject observations are sometimes significant and since the data may not always show a bell shaped distribution.

A third issue in crossover designs concerns missing data. Longitudinal studies including crossover designs of free-living biological subjects often suffer the disadvantage of losing subjects during the study. Subjects may drop out of the study due to illness, drug toxicity, vacations and many other unpredictable events, leading to "missing data" and significant difficulties in statistical analysis. Data may be missing due to a random process independent of the experiment and its effects, a process dependent on the experiment, or because the design is purposely incomplete. Methods of handling messy data in crossover designs have not been extensively presented in the literature.
The major thrust of this study is to develop linear models for the crossover design without ignoring residual effects and using these models together with current missing data methodology to analyze incomplete crossover data.

A literature review including the application of developed missing data methods to crossover designs is presented in this chapter. Methods of analyzing complete data using both the traditional crossover model and a new full rank model are given in Chapter II. The full rank model is particularly useful when the first- and second-order residual effects should not be ignored. The traditional model also allows up-to-second-order residual effects by placing certain constraints on the parameters. Chapter III introduces two incomplete data models for three-period two-treatment crossover designs. The analysis procedures using these models are illustrated by a numerical example. A computer algorithm is developed for the analysis. Studies of variance-covariance matrix structure in crossover designs are given in Chapter IV. Situations of having homogeneous variances and heterogeneous variances across time periods are discussed. Theoretically-modelled covariance structures are assumed and analysis results under different covariance models are presented and compared with the unrestricted covariance model through a numerical example. Finally, Chapter V summarizes the results of this study and indicates possible directions for further research.

1.2 Literature review

The literature review of this study will be divided into two components: crossover designs, and the analysis of incomplete crossover data. In the literature of crossover designs with residual effects, the early development concentrated on the changes of design patterns. These classic papers did
not explicitly assume a general linear model until Berenblut (1964), but
the data analysis they performed implied the general linear univariate
ANOVA model with error terms assumed to be equally correlated, indenti-
cally and normally distributed. These classic papers were not concerned
with the validity of these assumptions. The literature on the analysis of
crossover designs with residual effects was developed later in order to ad-
dress the importance of these critical assumptions. In addition, Grizzle
(1965) included the assumption of a mixed model in the analysis of two-
period two-treatment two-sequence crossover designs, under which the subject
effects and error terms are random effects. This approach seems to concen-
trate on the simple designs. In the literature of incomplete data analysis,
there are two predominant and useful approaches: univariate and multivariate
linear models, including models for both continuous and categorical data.

1.2.1 Crossover designs - design

Cochran and Cox (1950) described the crossover design, developed by
earlier authors, as sets of Latin squares, where rows represent periods,
columns represent sequences, and squares represent a blocking factor. The
example they gave consists of two replicates of a three-period three-treat-
ment design with residual effects. They showed how estimates of effects
and sums of squares can be obtained by analysis of variance techniques.
They noted that the sum of squares used for hypothesis testing of treat-
ment effects and that for residual effects are not orthogonal and need to
be calculated separately. Furthermore, they cited the idea of balanced de-
signs from Williams (1949) and demonstrated that balanced designs for first-
order residual effects can consist of a single Latin square if the number
of treatment is even, two Latin squares if the number of treatment is odd.
Finally, they used switchback designs from Brandt (1938) to reduce the ex-
perimental error due to the change of periods.

Lucas (1957) introduced the extra-period design by repeating the treatment in the last period of the design. This idea was originally from Yates (1947). The extra-period design (Table 1.1.2) permits the direct treatment effects and residual effects to be orthogonal. It also allows the residual effects to be orthogonal to sequences. In contrast, the direct treatment effects are not orthogonal to subject effects in these designs, but the degree of non-orthogonality is not great. In using these designs, the net result is that of increased efficiency in measuring residual and cumulative treatment effects at the expense of some loss in efficiency on direct treatment effects. Lucas's work was extended by Patterson and Lucas (1959) to a wide class of extra-period crossover designs.

The concept of tied-double-change-over designs (Table 1.1.3) described by Federer (1955) was generalized by Federer and Atkinson (1964). The construction of these designs is complicated. The designs are useful when the estimates of direct and residual effects are equally important, because the designs tend to give equal estimated variances for both effects when q, the number of periods, increases.

Berenblut (1964) redeveloped a class of crossover designs from Quenouille (1953). The designs are almost completely balanced for residual effects except that residual effects are partially confounded with sequences. Consequently, the residual effects and direct effects are orthogonal and the variance of contrasts of direct effects actually attains minimum value since they are not confounded with any other effects at all. A disadvantage is that many periods are required; for t treatments, 2t periods and $t^2$ subjects are required for constructing such a design (Table 1.1.4). Sharma (1980) extended the designs to the case when the second-order residual effects are
also balanced.

In order to reduce the number of periods required for the balanced designs, Berenblut (1967) presented a four-treatment four-period, eight-sequence crossover design (Table 1.1.5). He assumed the treatment factor is at four equally spaced levels in this design. By examining the linear, quadratic and cubic components of direct and residual effects, he found that all three degrees of freedom for direct effects are orthogonal to linear and cubic components of the residual effects. He suggested using the designs if the quadratic term of residual effects is not statistically significant since the estimates of contrasts of direct effects can still share the same property of minimum variance as his previous designs. Berenblut (1968) extended his work to the more general situation by including direct by residual interaction effects in the analysis.

Atkinson (1966) presented designs to be used when the effect of a sequence of treatments is the quantity of interest. His designs (Table 1.1.6) consist of applying one treatment K times followed by a distinct treatment K times. For K=2 and 2 treatments, the design is simply a subset of designs developed from Berenblut (1964). These designs are particularly applicable when the residual effects persist for a long time.

Balaam (1968) developed two-period crossover designs with t treatments and $t^2$ subjects. The designs are suitable for experiments when the periods by treatment interactions are important and there are no residual effects, and when subjects receive only two-treatment periods. The two-treatment design (Table 1.1.7) discussed in his paper is obtained by simply discontinuing that of Berenblut (1964) at the end of the second period.

Davis and Hall (1969) introduced a class of cyclic crossover designs. These designs are extended from the cyclic incomplete block designs and exist
for any number of treatments and periods. The efficiencies of these designs compare favorably with previously existing designs and in general the cyclic designs require fewer subjects. But orthogonality of direct and residual effects is not achieved in their designs (Table 1.1.8).

Some algebraic results in the theory of serial factorial designs were discussed by Kok and Patterson (1976). They defined the general conditions when direct and residual effects are orthogonal in these designs. Furthermore, special "R-orthogonal" designs are defined in order that the residual effects can also be orthogonal to direct by residual interactions. Both the complete balanced designs discussed by Berenblut (1964) and the designs of Patterson (1970) are R-orthogonal.

Kershner (1979) presented an extensive bibliography of the crossover design literature. Kershner and Federer (1980) presented a class of two-treatment crossover designs and set up a general model for use with virtually any kind of crossover designs. Furthermore, they compared the efficiency and effect of complete random designs with extra-period designs and demonstrated that complete random designs will be no better than the three or four-period crossover alternatives in the presence of residual effects or period by treatment interactions.

1.2.2 Crossover design - analysis

Box (1954) considered the problem of the effects of inequality of variance and of correlation between errors in the two-way analysis of variance, e.g. with g groups by k treatments when the observations in each group are multivariate normal with arbitrary variance-covariance matrix. He found the the sum of squares for treatment effects is not stochastically independent of that for error. His results reveal the difficulties of
hypothesis testing with dependent data. He developed a conservative F test for testing treatment effects; under \( H_0 \), the F is distributed as \( F(\varepsilon(k-1), \varepsilon(k-1)(N-g)) \) where \( \varepsilon \) is a function of \( \Sigma \). His work is the theoretical basis for testing multivariate data with a univariate model.

Geiser and Greenhouse (1958) extended his result to a two-way mixed model with \( g>2 \). They noted that \( \varepsilon > 1/(k-1) \) and suggested approaching the analysis under the setting of a general \( \Sigma \) by computing the usual univariate sums of squares and modifying the degrees of freedom by using the lower bound on \( \varepsilon \). Although there is no direct connection with crossover design analysis, their work actually gives the theoretical foundation for hypothesis testing in crossover design analysis.

Chassan (1964) indicated a correct procedure in testing for treatment effect in crossovers when the numbers of subjects in the two sequence groups are not equal, as is often the case in clinical studies in which drop-outs for reasons other than treatment effects may occur. He suggested taking the difference between two observations from the same subject. The new, transformed data can be used to test treatment effects with students' t test. This test is not confounded with order of treatment.

Grizzle (1965) developed a linear model and addressed the problem of estimating direct and residual effects in the two-period two-treatment crossover designs. He assumed a mixed model where the subject and error terms are distributed as \( N(0,\sigma_s^2) \) and \( N(0,\sigma_e^2) \) respectively. Consequently the variance of an observation is \( \sigma_s^2 + \sigma_e^2 \) and two observations on a subject have covariance \( \sigma_s^2 \). Observations made on different subjects are independent. In his model, period effects are not estimable in the presence of residual effects. The residual effects are also completely confounded with sequence effects. In addition, the estimation of the treatment effect
suffers a significant loss of efficiency if residual effects are simultaneously present in the model, since the estimate of treatment effect is then based only on first period data. For hypothesis testing, he found that the subjects-within-sequence mean square is the proper term for testing the equality of residual effects. Furthermore, he found that there is no appropriate error term to test the significance of treatment effects in the presence of residual effects. His work reveals the analytical difficulties of two-treatment, two-period, two-sequence crossover designs. More important, he showed that a better strategy for handling crossover data is to examine residual effects first, then analyze the crossover data in the absence of residual effects. The analysis proposed by Grizzle (1965) uses only first period data to estimate treatment effects, but in a correction (Grizzle, 1974), treatment effects are estimated using data from both periods.

Balaam (1968), discussed above under "designs", also presented an interesting analysis procedure. He performed both intra- and inter-experimental subject analysis by using raw data and taking the sum and difference of observations from the same subject respectively. He then showed that the intra-subject analysis is more efficient than the inter-subject analysis since the between-subject variability is eliminated by taking sums and differences of observations. However, his model is different from Grizzle's by including period by treatment interaction effects and eliminating residual effects.

Gart (1969) considered the problem when the response in the crossover design is dichotomous. He used Cox's (1958) logistic model in the two-treatment two-period two-sequence crossover design and showed the exact conditional distribution of the second period data to be hypergeometric. Hence the exact tests for treatment and residual effects are equivalent to
Fisher's exact tests in this case.

Koch (1972) presented a nonparametric method for analysis of the two-treatment two-period two-sequence crossover design. He assumed Grizzle's mixed model, except for distribution assumptions, and analyzed the rank scores with a Wilcoxon statistic. In particular, the test of equality of residual effects is performed by analyzing the sum of the two observations on the same subjects in order to make the subjects in two different sequences satisfy the same model. Moreover, the dependence of within-subject observations is eliminated since the new (sum) observations are from different subjects. In the same manner, the test of equality of direct effects when residual effects are absent can be performed by taking the difference of the two observations on the same subject because the between-subject variability is removed at this time. Koch's work, together with Grizzle's (1965) paper, represents the majority of the presently published theory relevant to estimating and testing in the two treatment case when residual effects are present.

Wallenstein and Fisher (1977) extended Grizzle's (1965) and Koch's (1972) ideas to the two periods repeated measurements crossover designs. In addition to the current parameters in the crossover designs model, they included time effects and the interaction of time with sequence, period and treatment (direct and carryover) effects. Consequently, the carryover effects and time by carryover interaction effects can be tested by taking the sum of the two observations from the same subject at the same time interval whereas in the absence of carryover effects, the treatment, period, time effects and time by treatment, time by period interaction effects can be tested by taking the difference as the new observations. This paper differs from Koch (1972) in that the error terms used for testing time
effects and interaction effects are not the same as those for testing treatment and carryover effects.

Wallenstein (1979) introduced the use of baseline values in the analysis of crossover designs. The procedure involves taking pre- and post-treatment measurements from each period of the experiment. The treatment effects can be estimated by analyzing the pre-post differences and the carryover effects can be estimated by analyzing the pre-treatment values only. He also found that the estimate of period effects is confounded with the sum of carryover effects. Furthermore, one needs to assume the improvement over baseline is the same for both periods and for each subject. His design is actually similar to the designs which allow rest periods (washout periods) before the treatments are given.

In all the literature on the analysis of crossover designs with residual effects there seems to be no definite strategy to treat the inherent residual effect problem. These authors did not consider the three-period or full-sequence alternatives. As will be shown in the next chapter, these alternatives or their combinations can reduce or eliminate some of the confounding inherent in the two-period designs and also improve the interpretability of some of the results.

1.2.3 Analysis of incomplete crossover data

In the literature of analysis of crossover designs, little has been done with the missing data problem. When missing data occur, the inter-subject analysis procedure suggested by Balaam (1968) and Koch (1972), which uses the sum and difference of observations from the same subject, is more difficult to use since either (1) the subject must be dropped out from the analysis even if only one observation from the subject is missing,
or (2) the missing values might be replaced by estimates, which are also very difficult to obtain. Lucas (1957) and Balaam (1968) presented a heuristic procedure to estimates missing values by replacing them with a value which minimizes the error sum of squares in the analysis. But they also warn the users that the procedure is formidable. In order to consider the missing data problems together with intra-subject correlation in the crossover designs, several papers which may be applicable in this case are selected from the literature of incomplete multivariate data analysis and are reviewed here.

Srivastava and Khatri (1979) described conditions on the covariance matrix of a vector of correlated observations under which the usual univariate general linear model, least squares estimate of $\hat{\beta}$ is identical to the BLUE of $\hat{\beta}$. The procedure assumes the linear model to be

$$E(Y) = X\beta, \quad V(Y) = \Sigma$$

(1.2.1)

The weighted least squares estimates of $\beta$ from (1.2.1) will be identical to the unweighted estimates if the covariance matrix $\Sigma$ satisfies

$$\Sigma = \sigma^2 P + (I - P_X) G_6 X + X' G_6 (I - P_X) + (I - P_X) G_4 (I - P_X)$$

(1.2.2)

where

$$P_X = X'(X'X)^{-1} X'$$

$$G_6 = (G_5 + \sigma^2 (X'X)^{-1}) G_1$$

$$G_4 (I - P_X) = 0$$

$G_4$ is an arbitrary positive-definite matrix

$G_1$ and $G_5$ are arbitrary

The proof was developed by Rao (1967) and Mitra and Rao (1968). It is not
known if there exist analogous conditions for hypothesis testing, and little work has been done to determine if there is any situation in practice satisfying these conditions.

A second approach to the analysis of incomplete multivariate data uses the multivariate model. Casewise deletion, where all data from the same subject are deleted when there is at least one missing value, seems to be the traditional method to deal with the problem since the very early days of multivariate analysis. The application of casewise deletion to crossover designs would imply the deletion of all data from any subject for whom any one observation is missing.

Glasser (1964) presented the pairwise deletion method. For estimating the means and variances of the $j^{th}$ variable in the model, he suggested using all available data in the $j^{th}$ variable. As for the covariance of the $i^{th}$ and $j^{th}$ variables, one can estimate by using those observations where both variables are complete and deleting the others. Although the resulting estimate of $\Sigma$ is unbiased, the method can give an estimate of $\Sigma$ with negative eigenvalues (indefinite) which could lead to negative variance estimates for certain linear combinations of variables. The application of the pairwise deletion method to crossover designs can be described under two conditions: (1) For two-period designs, the pairwise deletion uses all data to estimate means and variances, but for estimating covariances it deletes all data from subjects which have missing data in one period. (2) For more-than-two-period designs, pairwise deletion does not delete all data from a subject unless there is only one observation available from the subject.

"Missing data" can usually be divided into three classes: structurally missing data, randomly missing data and censored data. A structurally missing value occurs when a subject is dropped from the study, in accordance
with the protocol: all "missing data" after the subject is dropped are structurally missing data. A "randomly missing value" occurs when the protocol specifies that data ought to be collected but, due to uncontrolled factors, data were not collected or were lost. The literature relevant to several important methods of handling either case is described below. In general, these methods may be superior to both casewise deletion and pairwise deletion methods in the sense that, instead of deleting observations from the data, they either use all available data or replace missing values by some estimates and use all data to perform the analysis.

Trawinski and Bargmann (1964) considered the structurally missing data problem in which the extant data have hierarchical structure. The data are first grouped according to their missing-value patterns. They derived a non-iterative maximum likelihood (ML) method to estimate means, variances and covariances and the likelihood ratio linear hypothesis test procedure for the hierarchically (structurally) missing data problem. The method is restricted to special (hierarchical) missing data patterns. The application of this method to crossover data uses all available data to estimate means, variances and covariances.

Hocking and Smith (1968) developed a heuristic method for estimating parameters in the multivariate normal distribution with structurally missing data. The data are first grouped according to their missing-value patterns. The estimates of variances and covariances are obtained as a weighted average of the estimates of parameters in each group. The resulting estimates are proved to be consistent and have the same large sample properties as the MLE. In addition, the estimates converge to the C-R lower bound when the number of observations in the complete data group is large. The results do not generalize to models with concomitant variables.
Hartley and Hocking (1971) presented a simple taxonomy for incomplete data analysis and at the same time developed the MLE for the structurally-missing data problem. They used an iterative method to estimate the mean and the variance-covariance matrix from the marginal likelihood function of the existing data; the estimates are weighted averages of group means, variances and covariances respectively. They did not develop hypothesis testing procedures. The results do not extend to models with concomitant variables.

Orchard and Woodbury (1970) presented maximum likelihood estimates for the incomplete multivariate model data based on the "missing information principle" (MIP). The basic idea of MIP is to consider the conditional distribution of observed data given that missing data are random variables, replace the missing data with MLE's from the conditional likelihood function, and then use the real data plus estimated missing values with standard MLE methods (i.e., as if the data were complete) to estimate parameters, and repeat until the iterative process converges. They did not develop general linear multivariate hypothesis procedure. But a big difference between the two methods is that the Orchard-Woodbury method can be applied to both structurally missing data and randomly missing data problems while the other cannot.

Woodbury and Hasselblad (1970) gave an empirical example to illustrate the MLE based on MIP. They presented an algorithm for the method described by Orchard and Woodbury and compared their method with the analytical method developed by Anderson (1957) in the "nested" missing pattern condition. The results from both methods agree to 5 decimal places after 20 iterations. The application of the Woodbury-Orchard-Hasselblad method to crossover data uses all available data plus estimated missing data, and may be a favored method since the inter-subject variability might be eliminated in this way.
Kleinbaum (1970, 1973) proposed a generalization of the standard GLMM model which he called the more general linear model (MGLMM), and developed large-sample techniques for both estimation and hypothesis testing based on the MGLMM. His technique is applicable to both structurally-missing and randomly-missing situations. The basic idea is to "roll-out" the $\chi$ and $\beta$ matrices and re-shape $\chi$ in the GLMM into a form such that the univariate approach can be applied. Kleinbaum derived a class of BAN, unbiased estimators for $\beta$ based on generalized least squares using any consistent estimator of $\Sigma$. ($\Sigma$ can be estimated by pairwise deletion, for example.) Kleinbaum described both iterative and non-iterative BAN estimation procedures for $\beta$ and $\Sigma$. The iterative procedures always converge to the MLE but simulation showed that in some cases the non-iterative procedure is superior in small samples to the MLE. Furthermore, Kleinbaum developed hypothesis testing theory in MGLMM by using weighted least squares to generate Wald ($\chi^2$) statistics. He verified that Wald statistics in the MGLMM are asymptotically equivalent to likelihood ratio tests (LRT). He also demonstrated that in the standard multivariate case without missing data, the Wald statistic is equivalent to Hotelling's $T^2$ statistic and has desirable small sample properties in this special case. However, Kleinbaum's method is asymptotic, with unknown small-sample properties. It also requires substantial amounts of computation and may not be practical. The application of Kleinbaum's technique to crossover data is straightforward. It applies all available data to estimate parameters. The only disadvantage is that to attain a certain power for hypothesis testing, quite a large sample may be necessary.

Hosking (1980) generalized the Hartley-Hocking-Smith (HHS) procedure and Woodbury-Orchard-Hasselblad (WOH) technique to the general linear multivariate model. He found that in certain circumstances the two approaches
are equivalent in the sense that the systems of equations being solved are
equivalent. He also found that the algorithm fails for the HHS procedure
unless each group of data with the same pattern of missing values contains
sufficient data for all relevant rows of $\hat{\beta}$ to be estimable. He compared
the HHS, WOH and Kleinbaum techniques in a Monte Carlo study and found that
for estimating $\Sigma$, Kleinbaum's technique is better, in his situation, than the
the WOH or pairwise techniques, which are roughly comparable. For esti-
mating $\hat{\beta}$, the WOH was found to be better than Kleinbaum and pairwise tech-
niques. Overall the Kleinbaum and WOH are better than casewise deletion.
The designs analyzed by Hosking were not crossover-type designs, nor were
the patterns in $\Sigma$ typical of covariance matrices found in crossover designs.
### TABLE 1.1 Schematic Representation of Some Crossover Designs

1) **Standard two-period two-treatment two-sequence design**

<table>
<thead>
<tr>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Period 1</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

2) **Extra-period two-treatment design** (Lucas 1957)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>3</td>
<td>B</td>
<td>A</td>
</tr>
</tbody>
</table>

3) **Tied-double-change-over design** (Federer and Atkinson 1964)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>3</td>
<td>C</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>4</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>5</td>
<td>C</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>6</td>
<td>B</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>7</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
</tbody>
</table>

4) **Design completely balanced for first-order residual effects** (Berenblut 1964)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>3</td>
<td>B</td>
<td>A</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>4</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>A</td>
</tr>
</tbody>
</table>

5) **Completely balanced design at four equally spaced levels of treatment** (Berenblut 1967)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>A</td>
<td>D</td>
<td>C</td>
<td>C</td>
<td>D</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>3</td>
<td>D</td>
<td>C</td>
<td>B</td>
<td>A</td>
<td>D</td>
<td>C</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>4</td>
<td>C</td>
<td>D</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>A</td>
<td>D</td>
<td>C</td>
</tr>
</tbody>
</table>
6) Crossover design balanced for first-order residual effects (Atkinson 1962)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>B</td>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>2</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>B</td>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>3</td>
<td>C</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>4</td>
<td>C</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>5</td>
<td>B</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>6</td>
<td>B</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
</tbody>
</table>

7) Two-period two-treatment four-sequence design (Balaam 1968)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>B</td>
</tr>
</tbody>
</table>

8) Cyclic change-over design (Davis and Hall 1969)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>F</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
</tr>
<tr>
<td>2</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>3</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>1</td>
<td>F</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
</tr>
<tr>
<td>2</td>
<td>E</td>
<td>F</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>3</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
</tr>
</tbody>
</table>
1.3 Linear model notations and abbreviations

We shall use the notations and abbreviations described in this section for discussions in the succeeding chapters.

1.3.1 GLUM - General linear univariate model

We shall use the abbreviations GLUM \((Y, X\beta, \sigma^2 I_N)\) to represent the general linear univariate model,

\[
E(Y) = X\beta
\]

\[
V(Y) = \sigma^2 I_N
\]  

(1.3.1)

The parameter vector \(\beta(q \times 1)\) is called the primary vector, \(Y(N \times 1)\) is called the response vector, \(X(N \times q)\) is called the design matrix, and \(\sigma^2\) is the variance parameter. \(\beta\) and \(\sigma^2\) are unknown. We shall be interested in estimable secondary parameters of the form

\[
\delta = C\beta - \theta_0
\]  

(1.3.2)

where \(C(a \times q)\) and \(\theta_0(a \times 1)\) are known, fixed constant matrices, and in testable hypotheses of the form

\[
H_0: \delta = 0.
\]  

(1.3.3)

A slightly more general version of the model is denoted by GLUM \((Y, X\beta, \sigma^2 Y)\), where

\[
E(Y) = X\beta
\]

\[
V(Y) = \sigma^2 Y
\]  

(1.3.4)

and \(Y(N \times N)\) is a known positive definite matrix.
1.3.2 GLMM - General linear multivariate model

We shall use GLMM(\(\mathbf{Y}, \mathbf{X}\beta, \mathbf{\Sigma}\)) to represent the general linear multivariate model

\[
E(\mathbf{Y}) = \mathbf{X}\beta
\]

\[
\text{V}(\mathbf{Y}) = \mathbf{I}_N \otimes \mathbf{\Sigma}.
\]  

(1.3.5)

The elements of the model have names similar to those of GLUM except that the dimensions vary. \(\mathbf{Y}(N \times p)\) is the observed response matrix; the elements of each row of \(\mathbf{Y}\) are assumed to have \(p \times p\) covariance matrix \(\mathbf{\Sigma} = (\sigma_{ij})\) which is usually unknown. Any two different rows of \(\mathbf{Y}\) are independent. \(\mathbf{\beta}(q \times p)\) is the matrix of unknown primary parameters, and \(\mathbf{X}(N \times q)\) is the design matrix as in GLUM.

Letting \(\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_p)\) and \(\mathbf{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)\), the GLMM assumptions discussed above ensure that any \(\mathbf{Y}_j\) satisfies the GLM \((\mathbf{Y}_j, \mathbf{X}_j\beta_j, \sigma_{jj}\mathbf{I}_N)\) assumptions. Note that \(\sigma_{jj}\) is the \(j\)th diagonal element of \(\mathbf{\Sigma}\). We shall be interested in estimable GLMM secondary parameters specified in the form

\[
\mathbf{Q} = \mathbf{G} \mathbf{Y} - \mathbf{Q}_0
\]

(1.3.6)

where \(\mathbf{G}(a \times p), \mathbf{Y}(q \times b)\) and \(\mathbf{Q}_0(a \times b)\) are known, fixed constant matrices, and in testable GLMM hypotheses of the form

\[
H_0: \mathbf{Q} = \mathbf{0}
\]

(1.3.7)

GLMM manipulations are often more convenient when the response matrix \(\mathbf{Y}\) is rolled out by rows or columns. Let

\[
\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_p) = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N)^T
\]

(1.3.8)

so that \(\mathbf{Y}_j(N \times 1)\) is the \(j\)th column of \(\mathbf{Y}\) and \(\mathbf{y}_i(p \times 1)\) is the transpose of the \(i\)th row of \(\mathbf{Y}\).
\( Y^*(Np \times 1) = (Y_1^T, Y_2^T, \ldots, Y_p^T)^T \) \hspace{1cm} (1.3.9)

then the **column-wise version** of GLMM \((Y, X, \beta, \Sigma)\) can be shown as

\[
E(Y^*) = X^* \beta^*
\]
\[
V(Y^*) = \Sigma \otimes I_N
\]

with

\[
\begin{align*}
X^*(Np \times pq) &= I_p \otimes X, \\
\beta^* (pq \times 1) &= (\beta_1^T, \beta_2^T, \ldots, \beta_p^T)^T
\end{align*}
\]

(1.3.10) \hspace{1cm} (1.3.11) \hspace{1cm} (1.3.12)

where \( \beta_j(q \times 1) \) is the \( j \)th column of \( \beta \) and \( \otimes \) is the right Kronecker product.

Secondary parameters in the form (2.1.6) may be converted to the rolled out representation of the GLMM, denoted by

\[
\beta^* = \mathcal{L}^* \beta^* - \beta_0^*
\]

(1.3.13)

where \( \beta^*(ab \times 1) \) is the appropriately rolled out version of \( \beta(a \times b) \), and \( \mathcal{L}^*(ab \times pq) \) is

\[
\mathcal{L}^* \equiv \Sigma \otimes \mathcal{U}^T
\]

(1.3.14)

Similarly, the **row-wise version** of GLMM \((Y, X, \beta, \Sigma)\) can be represented as

\[
E(Y^\#) = X^\# \beta^\#
\]
\[
V(Y^\#) = I_N \otimes \Sigma
\]

by letting

\[
\begin{align*}
Y^\# (Np \times 1) &= (u_1^T, u_2^T, \ldots, u_N^T)^T, \\
X^\# (Np \times pq) &= X \otimes I_p^*, \\
\beta^\# (pq \times 1) &= (Y_1^T, Y_2^T, \ldots, Y_q^T)
\end{align*}
\]

(1.3.15) \hspace{1cm} (1.3.16) \hspace{1cm} (1.3.17) \hspace{1cm} (1.3.18)

where \( Y_i(p \times 1) \) is the \( i \)th row of \( \beta \).

Secondary parameters may be converted to this rolled out representation of the GLMM, denoted by
\[ \Theta^\# = L^\# \hat{\Theta}^\# - \Theta_0^\# \]  
(1.3.19)

where \( \Theta^\# \) is (1.5.6) rolled out by rows and \( L^\# \) (ab×pq) is

\[ L^\# = U^T \otimes \zeta. \]  
(1.5.20)

1.3.3 MGLMM - More general linear multivariate model

Both versions of GLMM, (2.1.10) and (2.1.15), were further generalized by Kleinbaum (1973) to the "More general linear multivariate model", which allows different design matrices for different columns of \( \mathbf{y} \) and can also be used to handle missing values in the multivariate model.

The column-wise version is denoted as MGLM(\( \mathbf{y}^*, \mathbf{D}^*, \mathbf{E}^*, \Theta^* \)), which represents

\[ \mathbf{E}(\mathbf{y}^*) = \mathbf{D}^* \mathbf{E}^* \]  
(1.3.21)

\[ \mathbf{V}(\mathbf{y}^*) = \Theta^* \]  
(1.3.22)

\[ \mathbf{D}^*(\sum_{j=1}^{p} \mathbf{D}^*_j \otimes \mathbf{Q}_j) = \begin{bmatrix} \mathbf{D}_1 & \mathbf{Q}_1 \otimes \mathbf{Q}_2 & \mathbf{Q}_2 \otimes \mathbf{Q}_3 & \cdots & \mathbf{Q}_p \otimes \mathbf{Q}_p \\ \mathbf{Q}_2 \otimes \mathbf{Q}_1 & \mathbf{D}_2 & \mathbf{Q}_3 \otimes \mathbf{Q}_3 & \cdots & \mathbf{Q}_2 \otimes \mathbf{Q}_p \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \mathbf{Q}_p \otimes \mathbf{Q}_1 & \mathbf{Q}_p \otimes \mathbf{Q}_2 & \cdots & \mathbf{D}_p \end{bmatrix} \]  
(1.3.23)

and

\[ \mathbf{E}^*(\mathbf{E} \mathbf{q}_j \times 1) = (\mathbf{E}^T \mathbf{q}_1, \mathbf{E}^T \mathbf{q}_2, \ldots, \mathbf{E}^T \mathbf{q}_p) \]  
(1.3.24)

where \( \mathbf{y}^*_j \) (n×1) is the vector of non-missing responses in the \( j \)th column of \( \mathbf{y} \), \( \mathbf{q}_j \) (q×1) is the primary parameter corresponding to \( \mathbf{y}^*_j \), and \( \mathbf{D}_j \) (n×q) is the design matrix corresponding to \( \mathbf{y}^*_j \). In addition,
\( Q^*(\Sigma_{n_j} \times \Sigma_{n_j}) = \begin{bmatrix} \sigma_{11} \ldots \sigma_{1p} \\
1 & 1 & \ldots & 1 \\
\sigma_{21} \ldots \sigma_{2p} \\
\vdots \\
\sigma_{p1} \ldots \sigma_{pp} \end{bmatrix} \) (1.3.25)

where \( Q_{ij} = (n_i \times n_j) \) is an \((n_i \times n_j)\) matrix such that

\[
 q_{rs}(ij) = \begin{cases} 
 1 & \text{if the } r\text{th element of } \chi_i^* \text{ and the } s\text{th element of } \chi_j^* \text{ are correlated;} \\
 0 & \text{otherwise}
\end{cases}
\]

Note that (1.3.21) is a more general case of (1.3.5) in the sense that (1.3.21) may allow different missing data patterns in the different columns of \( \chi \) and the corresponding design matrix of \( \chi_j^* \) may be different from that of \( \chi_i^* \) for \( i \neq j \). Secondary parameter in the form (1.3.6) may be converted to

\[
\theta^* = H \xi^* - \theta_0^*
\] (1.3.26)

where \( H(h \times Q_j) \) is the incidence matrix, constant for any linear combinations of \( \xi^* \). Again (1.3.26) is more general than (1.3.6) since (1.3.26) allows all possible combinations of parameters from \( \xi^* \), whereas (1.3.6) does not allow linear combinations of two parameters with distinct columns and rows.

Similarly, we can define the row wise version as \( MGLM(\chi^#, D^#, \xi^#, \omega^#) \).

This can be transformed from model (1.3.21) by letting \( P(\Sigma_{n_j} \times \Sigma_{n_j}) \) be a permutation matrix such that

\[
\chi^# = P \chi^* = (\chi_1^T, \chi_2^T, \ldots, \chi_N^T)^T
\] (1.3.27)

Note that \( \chi_i^#(p_i \times 1) \) contains the non-missing responses of \( \mu_i \) (p\times1) in (1.3.16).

Hence we can define \( D^# \) as

\[
D^# = P \Sigma^*
\] (1.3.28)
and the model is represented as
\[
\mathbb{E}(\mathbf{\chi}^\#) = \mathbf{D}^\# \mathbf{\varepsilon}^\# = \mathbf{P} \mathbf{\Sigma} \mathbf{P}'^\#
\]
(1.3.29)
\[
\mathbb{V}(\mathbf{\chi}^\#) = \mathbf{\Omega}^\# = \mathbf{P} \mathbf{\Omega} \mathbf{P}'
\]
(1.3.30)
\(\mathbf{\Omega}^\#\) can actually be seen in a simple form as
\[
\mathbf{\Omega}^\# = \begin{bmatrix}
\Sigma_1, & 0, & \ldots, & 0 \\
0, & \Sigma_2, & \ldots, & 0 \\
\vdots, & \vdots, & \ddots, & \vdots \\
0, & 0, & \ldots, & \Sigma_N
\end{bmatrix}
\]
(1.3.31)
where \(\Sigma_1(p_1 \times p_1)\) is the variance-covariance matrix of \(u_1^\#\).
CHAPTER II
LINEAR MODEL ANALYSIS OF COMPLETE DATA
FROM CROSSOVER DESIGNS

This chapter contains a summary of the models and methods used for the analysis of complete data in the three-period two-treatment crossover designs. This provides a common framework for comparing the various types of analysis in current use and a foundation for the analysis of incomplete data as discussed in the next chapter.

2.1 Traditional combined-period univariate model

The traditional linear model used for the two-treatment crossover design is called a "combined-period" model in the sense that primary parameters are defined as average effects across all periods, e.g. parameters corresponding to treatment effects are the averages of the two treatment effects across three periods. There is usually no period by treatment interaction effect assumed for this type of model.

2.1.1 The model

Following Kershner and Federer (1981), a class of traditional linear models for crossover designs can be represented as

\[ y_{ptuvk} = \mu + \pi_p + \lambda_{tuv} + \xi_k + \delta_{1p}\{\tau_t + \pi_{1t}\} \]

\[ + \delta_{2p}\{\tau_u + \alpha_t + \tau_{ut} + \pi_{2u}\} \]

\[ + \delta_{3p}\{\tau_v + \alpha_u + \beta_t + \tau_{vu} + \pi_{3v}\} + \epsilon_{ptuvk} \]  

(2.1.1)
where $Y_{ptuvk}$ = observed response in period $p$ for subject $k$, which receives treatment $t,u,v$ in period 1, 2, 3 respectively. The index $k(=1,2,\ldots,n)$ uniquely identifies subjects. [Note that the period index $p$ has a different meaning from the letter $p$ in the multivariate model in Section 1.3].

$\mu$ = overall mean.

$\pi_p$ = effect due to period $p$.

$\lambda_{tuv}$ = effect due to sequence $(t,u,v)$.

$\xi_k$ = random effect due to subject $k$, distributed with mean 0 and variance $\sigma^2_s$.

$\tau_v$ = effect due to treatment $v$.

$\alpha_u$ = first-order residual effect due to treatment $u$.

$\beta_t$ = second-order residual effect due to treatment $t$.

$\pi_{pt}$ = interaction effect due to period $p$ and treatment $t$.

$\tau_{ut}$ = interaction effect due to treatment $u$ and first-order residual effect from treatment $t$.

$\delta_{ij} = 1$ if $i=j$; 0 otherwise.

$\epsilon_{ptuvk}$ = random error, distributed with mean 0 and variance $\sigma^2_e$.

The assumption of random subject effects leads to a variance-component (mixed) model. Since the interaction terms are usually assumed to be negligible and hence omitted from the model, we will concentrate on the following reduced model without interactions for our analysis:

$$E(Y_{ptuvk}) = \mu + \pi_p + \lambda_{tuv} + \delta_{1p} \tau_t + \delta_{2p} \{\tau_u + \alpha_t\} + \delta_{3p} \{\tau_v + \alpha_u + \beta_t\}. \quad (2.1.2)$$
The covariances of the $Y_{ptuvk}$'s are denoted by

$$\text{Cov}(Y_{ptuvk}, Y_{p't'u'v'k'}) = \sigma^2 \text{ for } k=k' \text{ and } p=p',$$
$$= \rho \sigma^2 \text{ for } k=k' \text{ and } p \neq p',$$
$$= 0 \text{ otherwise}$$

(2.1.3)

Note that $\sigma^2 = \sigma_s^2 + \sigma_e^2$, $\rho = \sigma_s^2 / (\sigma_s^2 + \sigma_e^2)$, and the "subject effects" are represented in the covariance matrix.

The GLUM setting of model (2.1.2) can be expressed as

$$E(Y) = X\beta$$
$$V(Y) = \Sigma \otimes I_N$$

(2.1.4)

where $X(3n \times 1) = (Y_{1t1u1v1}, Y_{2t1u1v1}, Y_{3t1u1v1}, \ldots, Y_{3t_su_sv_s})^T$, 

$\beta(p \times 1) = (\mu, \pi_1, \pi_2, \pi_3, \ldots, \lambda_{tsu_sv_s})^T$

and $\Sigma(3 \times 3) = \sigma^2 \begin{bmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{bmatrix}$

(Symm)

$$= \begin{bmatrix} \sigma_s^2 + \sigma_e^2 & \sigma_s^2 & \sigma_s^2 \\ \sigma_s^2 & \sigma_s^2 + \sigma_e^2 & \sigma_s^2 \\ \sigma_s^2 & \sigma_s^2 & \sigma_s^2 + \sigma_e^2 \end{bmatrix}$$

(2.1.5)

Since period effects are not estimable unless residual effects are restricted, Kershner (1980) recommends putting constraints on the residual parameters such as

$$\alpha_A + \alpha_B = 0, \quad \beta_A + \beta_B = 0$$

(2.1.6)

to ensure the estimability of period effects in the traditional ANOVA without
further adjustment.

In general, analysis can be performed by first making a linear transformation of the observations within each subject. This leads to a new set of data which are i.i.d. Let $A(3n \times 3n)$ be a nonsingular matrix partitioned into two submatrices $A_1$ and $A_2$ where $A = [A_1^T, A_2]^T$,

$$A_1(2n \times 3n) = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{6} \end{bmatrix}$$

and

$$A_2(n \times 3n) = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix}.$$


The transformed GLUM model can be expressed as

$$E(Z_i) = \bar{Z}_i X \beta$$

$$V(Z_i) = \sigma_i^2 I_{n_i} \quad \text{(i=1 or 2)} \quad (2.1.7)$$

where $n_1 = 2n$, $n_2 = n$, $\sigma_1^2 = \sigma^2(1-\rho)$ and $\sigma_2^2 = \sigma^2(1+2\rho)$. $Z_1$ satisfies less-than-full-rank (LTFR) GLUM assumptions and may be analyzed to examine treatment, carryover and period effects. $Z_1$ contains no information about the overall mean or sequence effects:

$$\sqrt{2} \; z_{1tuvk} = (\pi_1 - \pi_2) + (\tau - \tau_v) - \alpha_u - \beta_t + (\epsilon_{1tuvk} - \epsilon_{3tuvk}),$$

$$\sqrt{6} \; z_{2tuvk} = (\pi_1 - 2\pi_2 + \pi_3) + (\tau - 2\tau_u + \tau_v) - 2\alpha_t + \alpha_u + \beta_t$$

$$+ (\epsilon_{1tuvk} - 2 \epsilon_{2tuvk} + \epsilon_{3tuvk}). \quad (2.1.8)$$

Analysis of $Z_1$ is called a "within-subject" analysis.

2.1.2 Estimation and hypothesis testing

Since the model for the within-subject data vector $Z_1 = A_1Y$ is LTFR, it is convenient to reparametrize to a full rank (FR) model before estimating
parameters. Let $X_1 = A_1X(n_1 \times q)$, the within-subject design matrix and $\beta_1 = \beta(q \times 1)$, so that $E(Z_1) = X_1\beta_1$. By standard reparametrization methods we can obtain an equivalent full rank model, $E(Z_1) = X_2\beta_2$, $V(Z_1) = \sigma^2_1I$, where $X_2$ has full column rank, $X_2 = X_1H_{12}$, and if $\theta = \zeta_1\beta_1$ is estimable then $\theta = \zeta_1H_{12}\beta_2$ in terms of the full rank model parameters. The BLUE of $\beta_2$ is $\hat{\beta}_2 = (X_2'X_2)^{-1}X_2'Z_1$ and if $\theta$ is estimable its BLUE is $\hat{\theta} = \zeta_1H_{12}\beta_2$. The variance parameter $\sigma^2_1 = \sigma^2(1-\rho)$ can be estimated using standard techniques.

A testable hypothesis $H_0: \zeta_1\beta_1 = \theta$ is tested in the reparametrized model by the identical hypothesis $H_0: \zeta_1H_{12}\beta_2 = \theta$.

2.1.3 An example

A two-treatment three-period crossover study was conducted to determine the effect of salt on the growth of mungbean sprouts. The sprouts were exposed to a moderately harmful dose of salt and tap water during three separate 24-hour periods, and the root length of each bean at the end of each period was measured. Since change in root length of mungbean was believed to have a non-normal distribution, the average change in root length of a pair of beans was the observed response. Six bean pairs were randomly assigned to each of the eight sequence groups. The eight sequence groups differed according to the type of treatment they received ($0$ = tap water placebo, $1$ = salt water) in each of the three periods. That is, one group received the placebo in all three periods, one group received salt water in all three periods and the remaining six groups received both placebo and salt depending upon the periods. Table 2.1 shows all eight sequences applied in the experiment. The three-period design is diagrammed in Figure 2.1. Measurements of sprout length were taken at $t_0 = 24$ hours, $t_1 = 48$ hours, etc., to $t_5 = 144$ hours. The response variable was sprout length growth during the
treatment period, i.e., length at \( t_1 \) minus length at \( t_0 \), length at \( t_3 \) minus length at \( t_2 \), and length at \( t_5 \) minus length at \( t_4 \).

**TABLE 2.1** Treatment Specifications for Eight-Sequence Three-Period Two-Treatment Crossover Design

<table>
<thead>
<tr>
<th>Period</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5 6 7 8</td>
</tr>
<tr>
<td>1</td>
<td>A A A B B B B</td>
</tr>
<tr>
<td>2</td>
<td>A A B B A A B</td>
</tr>
<tr>
<td>3</td>
<td>A B A B A B A</td>
</tr>
</tbody>
</table>

*A denotes tap water (placebo)  
B denotes salt water ("treatment")

**Figure 2.1** Diagram of the Three-Period Two-Treatment Experiment

<table>
<thead>
<tr>
<th>Baseline</th>
<th>Period 1 Treatment</th>
<th>Washout</th>
<th>Period 2 Treatment</th>
<th>Washout</th>
<th>Period 3 Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>24</td>
<td>48</td>
<td>72</td>
<td>96</td>
<td>120 144 Hours</td>
</tr>
<tr>
<td></td>
<td>( t_0 )</td>
<td>( t_1 )</td>
<td>( t_2 )</td>
<td>( t_3 )</td>
<td>( t_4 ) ( t_5 )</td>
</tr>
</tbody>
</table>

The data with 6 measurements per subject are presented in Table 2.2. Data in columns \( Y_1 \), \( Y_2 \) and \( Y_3 \) of Table 2.2 are the means (over a bean pair) of root length growths during the three time periods. If the \((48 \times 3)\) matrix of \( Y_1 \), \( Y_2 \) and \( Y_3 \) were rolled out by rows, one would have the \( Y \) matrix in model (2.1.4). The transformed data \( A_1 Y \) of model (2.1.7), and the corresponding design matrix \( A_1 \tilde{X} \) are shown in Table 2.3. Obviously, linear dependence can be observed among the columns of \( A_1 \tilde{X} \). The matrix \( \tilde{X}_2 \) in Table 2.4 is of full rank, converted from \( A_1 \tilde{X} \) through reparametrization. The number of columns in \( \tilde{X}_2 \) is the rank of matrix \( A_1 \tilde{X} \), equal to 5 in our example. The transformation matrix \( H_{12} \) is shown in Table 2.5.

The interesting estimable effects are the period effects, treatment
<table>
<thead>
<tr>
<th>AT1</th>
<th>AX1</th>
<th>AX2</th>
<th>AX3</th>
<th>AX4</th>
<th>AX5</th>
<th>AX6</th>
<th>AX7</th>
<th>AX8</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0.707107</td>
<td>-0.86602549</td>
<td>0.707107</td>
<td>0</td>
<td>0.707107</td>
<td>-0.707107</td>
<td>0.707107</td>
</tr>
<tr>
<td>-4.08248</td>
<td>0.408248</td>
<td>0.86602549</td>
<td>-0.707107</td>
<td>-0.707107</td>
<td>0.707107</td>
<td>0.707107</td>
<td>0.707107</td>
<td>0.707107</td>
</tr>
<tr>
<td>-2.46949</td>
<td>0.408248</td>
<td>0.86602549</td>
<td>-0.707107</td>
<td>-0.707107</td>
<td>0.707107</td>
<td>0.707107</td>
<td>0.707107</td>
<td>0.707107</td>
</tr>
<tr>
<td>1.53166</td>
<td>-0.408248</td>
<td>-0.86602549</td>
<td>0.707107</td>
<td>0.707107</td>
<td>-0.707107</td>
<td>-0.707107</td>
<td>0.707107</td>
<td>0.707107</td>
</tr>
<tr>
<td>-1.43259</td>
<td>0.408248</td>
<td>0.86602549</td>
<td>-0.707107</td>
<td>-0.707107</td>
<td>0.707107</td>
<td>0.707107</td>
<td>0.707107</td>
<td>0.707107</td>
</tr>
<tr>
<td>2.36396</td>
<td>-0.408248</td>
<td>-0.86602549</td>
<td>0.707107</td>
<td>0.707107</td>
<td>-0.707107</td>
<td>-0.707107</td>
<td>0.707107</td>
<td>0.707107</td>
</tr>
<tr>
<td>5.30720</td>
<td>-0.408248</td>
<td>-0.86602549</td>
<td>0.707107</td>
<td>0.707107</td>
<td>-0.707107</td>
<td>-0.707107</td>
<td>0.707107</td>
<td>0.707107</td>
</tr>
<tr>
<td>4.08248</td>
<td>0.408248</td>
<td>0.86602549</td>
<td>-0.707107</td>
<td>-0.707107</td>
<td>0.707107</td>
<td>0.707107</td>
<td>0.707107</td>
<td>0.707107</td>
</tr>
<tr>
<td>0.408248</td>
<td>0.408248</td>
<td>0.86602549</td>
<td>-0.707107</td>
<td>-0.707107</td>
<td>0.707107</td>
<td>0.707107</td>
<td>0.707107</td>
<td>0.707107</td>
</tr>
<tr>
<td>0.86602549</td>
<td>0.86602549</td>
<td>0.707107</td>
<td>0.707107</td>
<td>-0.707107</td>
<td>-0.707107</td>
<td>0.707107</td>
<td>0.707107</td>
<td>0.707107</td>
</tr>
<tr>
<td>0.408248</td>
<td>0.408248</td>
<td>0.86602549</td>
<td>-0.707107</td>
<td>-0.707107</td>
<td>0.707107</td>
<td>0.707107</td>
<td>0.707107</td>
<td>0.707107</td>
</tr>
<tr>
<td>0.86602549</td>
<td>0.86602549</td>
<td>0.707107</td>
<td>0.707107</td>
<td>-0.707107</td>
<td>-0.707107</td>
<td>0.707107</td>
<td>0.707107</td>
<td>0.707107</td>
</tr>
<tr>
<td>X1</td>
<td>X2</td>
<td>X3</td>
<td>X4</td>
<td>X5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.57735</td>
<td>0</td>
<td>-0.57735</td>
<td>1.1547</td>
<td>-0.57735</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.57735</td>
<td>1</td>
<td>0.57735</td>
<td>1.1547</td>
<td>-0.57735</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.57735</td>
<td>0</td>
<td>-0.57735</td>
<td>1.1547</td>
<td>-0.57735</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.57735</td>
<td>1</td>
<td>0.57735</td>
<td>1.1547</td>
<td>-0.57735</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.57735</td>
<td>0</td>
<td>-0.57735</td>
<td>1.1547</td>
<td>-0.57735</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.57735</td>
<td>1</td>
<td>0.57735</td>
<td>1.1547</td>
<td>-0.57735</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.57735</td>
<td>0</td>
<td>-0.57735</td>
<td>1.1547</td>
<td>-0.57735</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.57735</td>
<td>1</td>
<td>0.57735</td>
<td>1.1547</td>
<td>-0.57735</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.57735</td>
<td>0</td>
<td>-0.57735</td>
<td>1.1547</td>
<td>-0.57735</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.57735</td>
<td>1</td>
<td>0.57735</td>
<td>1.1547</td>
<td>-0.57735</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.57735</td>
<td>0</td>
<td>-0.57735</td>
<td>1.1547</td>
<td>-0.57735</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.57735</td>
<td>1</td>
<td>0.57735</td>
<td>1.1547</td>
<td>-0.57735</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.57735</td>
<td>0</td>
<td>-0.57735</td>
<td>1.1547</td>
<td>-0.57735</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.57735</td>
<td>1</td>
<td>0.57735</td>
<td>1.1547</td>
<td>-0.57735</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.57735</td>
<td>0</td>
<td>-0.57735</td>
<td>1.1547</td>
<td>-0.57735</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.57735</td>
<td>1</td>
<td>0.57735</td>
<td>1.1547</td>
<td>-0.57735</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.57735</td>
<td>0</td>
<td>-0.57735</td>
<td>1.1547</td>
<td>-0.57735</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.57735</td>
<td>1</td>
<td>0.57735</td>
<td>1.1547</td>
<td>-0.57735</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
effects and first- and second-order residual effects. The $G$ matrices for these effects (with respect to $\beta_1$) are shown in Table 2.6. For example, to test the equality of treatment effects, $G_1$ matrix has the form

$$ G_1 = (0 \ 0 \ 0 \ -1 \ 1 \ 0 \ 0 \ 0). $$

The least squares estimator and standard errors of the effects specified in Table 2.6 are shown in Table 2.7. The same results are obtained from traditional ANOVA methods (Grizzle, 1965).

Hypothesis testing results are summarized in the ANOVA table shown in Table 2.9. From the within-subject analysis we find that there is no significant difference between residual effects of two treatments beyond the period of application. The period effects are not significant either. On the other hand, we reject the hypothesis of no difference in the root growth of bean sprouts between tap water and salt water at a 5% level of significance. Salt water has a significant detrimental effect on the root growth of bean sprouts in our experiment.

<table>
<thead>
<tr>
<th>Table 2.5 Transformation matrix $H_{12}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FR MODEL PARAMETERS ($\beta_2$)</strong></td>
</tr>
<tr>
<td>$\pi_{1-\pi_3}$</td>
</tr>
<tr>
<td>----</td>
</tr>
<tr>
<td>$\mu$</td>
</tr>
<tr>
<td>$\pi_1$</td>
</tr>
<tr>
<td>$\pi_2$</td>
</tr>
<tr>
<td>$\pi_3$</td>
</tr>
<tr>
<td>$\tau_A$</td>
</tr>
<tr>
<td>$\tau_B$</td>
</tr>
<tr>
<td>$\alpha_{B-\alpha_A}$</td>
</tr>
<tr>
<td>$\beta_{B-\beta_A}$</td>
</tr>
</tbody>
</table>
2.1.4 Inter-subject analysis

There is additional information $A_2\gamma$ from the data collected, that has not been used yet. The matrix $A_2\gamma$ takes the sum of all observations from the same subject and is known as the inter-subject information. Koch (1972) tests the residual effects based on the sums of two scores in the two-period two-treatment crossover design; but he assumed random subject effects and no sequence effects in his model. In the eight-sequence full designs we can express the model equations of $A_2\gamma$ as

$$\sqrt{3} Z_{3tuvk} = \{3\mu + \sum_{p=1}^{3} \pi_p + 3 \lambda_{tuv} + 3 \xi_k + (\tau_t + \tau_u + \tau_v)
+ (\alpha_t + \alpha_u) + \beta_t + (\sum_{p=1}^{3} \epsilon_{ptuv})\}$$

(2.1.9)

The variance of $Z_{3tuvk}$ is $\sigma^2(1+2\rho)$. Obviously, period effects are not estimable in the inter-subject model. If we reparametrize the linear model $E(Z_2) = A_2\chi, V(Z_2) = \sigma^2(1+2\rho) I$, the full rank design matrix $X_3$ is shown to be exactly the same as the design matrix of eight sequence parameters (see Table 2.11). Any additional parameters in the model will be totally confounded with the sequence effects. For the effects of interest, the between-subject data can not provide the additional information unless the sequence effects are zero. Assuming zero sequence effects, the analysis of variance based on $A_2\gamma$ is presented in Tables 2.8 and 2.10. The estimates of $\tau_B - \tau_A$, $\alpha_B - \alpha_A$ and $\beta_B - \beta_A$ are similar in magnitude to the estimates in the intra-subject analysis but the standard errors are relatively large, the effect of a positive intra-class correlation coefficient. The P-values for corresponding hypothesis tests are similar to these in the within-subject analysis except for the second-order residual effects, which are significant in the inter-subject analysis but not in the intra-subject analysis. Note that the inter-subject analysis assumes zero sequence effect.
TABLE 2.6 Coefficients of Incidence Matrix $\zeta$ for Estimating Parameters from the Intra-Subject Model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\pi_1$</th>
<th>$\pi_2$</th>
<th>$\pi_3$</th>
<th>$\tau_A$</th>
<th>$\tau_B$</th>
<th>$\alpha_A$</th>
<th>$\alpha_B$</th>
<th>$\beta_A$</th>
<th>$\beta_B$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Period Effects</td>
<td>1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\bar{\theta}_1$</td>
<td>1</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Treatment Effects</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\bar{\theta}_2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>First-order Residuals</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\bar{\theta}_3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Second-order Residuals</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>1</td>
</tr>
</tbody>
</table>

TABLE 2.7 OLS Estimators from the Intra-Subject Model

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>$\hat{\theta}$</th>
<th>s.e.$(\hat{\theta})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi_1 - \pi_2$</td>
<td>-0.958</td>
<td>0.857</td>
</tr>
<tr>
<td>$\pi_1 - \pi_3$</td>
<td>-0.227</td>
<td>0.857</td>
</tr>
<tr>
<td>$\tau_B - \tau_A$</td>
<td>-2.457</td>
<td>1.089</td>
</tr>
<tr>
<td>$\alpha_B - \alpha_A$</td>
<td>-0.648</td>
<td>0.683</td>
</tr>
<tr>
<td>$\beta_B - \beta_A$</td>
<td>1.097</td>
<td>0.921</td>
</tr>
</tbody>
</table>
TABLE 2.8  OLS Estimates from the Inter-Subject Model

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\theta}$</th>
<th>s.e.$(\hat{\theta})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_B - \tau_A$</td>
<td>-5.485</td>
<td>2.515</td>
</tr>
<tr>
<td>$\alpha_B - \alpha_A$</td>
<td>0.082</td>
<td>1.778</td>
</tr>
<tr>
<td>$\beta_B - \beta_A$</td>
<td>-3.367</td>
<td>1.778</td>
</tr>
</tbody>
</table>

TABLE 2.9  Intra-Subject ANOVA Table

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>VR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Period Effects</td>
<td>2</td>
<td>24.076</td>
<td>12.038</td>
<td>0.675</td>
</tr>
<tr>
<td>Treatment Effects</td>
<td>1</td>
<td>89.681</td>
<td>89.681</td>
<td>5.032*</td>
</tr>
<tr>
<td>First-order Residuals</td>
<td>1</td>
<td>15.881</td>
<td>15.881</td>
<td>0.891</td>
</tr>
<tr>
<td>Second-order Residuals</td>
<td>1</td>
<td>25.029</td>
<td>25.029</td>
<td>1.404</td>
</tr>
<tr>
<td>Error</td>
<td>90</td>
<td>1603.98</td>
<td>17.822</td>
<td></td>
</tr>
</tbody>
</table>

95

TABLE 2.10  Inter-Subject ANOVA Table

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>VR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment Effects</td>
<td>1</td>
<td>120.34</td>
<td>120.34</td>
<td>4.756</td>
</tr>
<tr>
<td>First-order Residuals</td>
<td>1</td>
<td>0.051</td>
<td>0.051</td>
<td>0.002</td>
</tr>
<tr>
<td>Second-order Residuals</td>
<td>1</td>
<td>90.686</td>
<td>90.686</td>
<td>3.584</td>
</tr>
<tr>
<td>Error</td>
<td>44</td>
<td>1087.98</td>
<td>25.303</td>
<td></td>
</tr>
</tbody>
</table>

47
TABLE 2.11  Full-rank Design Matrix $\mathbf{X}_3$ from Inter-Subject Analysis

\[
\mathbf{X}_3 = (48\times8)
\]

\[
\begin{bmatrix}
1_{6\times1} & 0_{6\times1} & 0_{6\times1} & 0_{6\times1} & 0_{6\times1} & 0_{6\times1} & 0_{6\times1} & 0_{6\times1} \\
0_{6\times1} & 1_{6\times1} & 0_{6\times1} & 0_{6\times1} & 0_{6\times1} & 0_{6\times1} & 0_{6\times1} & 0_{6\times1} \\
0_{6\times1} & 0_{6\times1} & 1_{6\times1} & 0_{6\times1} & 0_{6\times1} & 0_{6\times1} & 0_{6\times1} & 0_{6\times1} \\
0_{6\times1} & 0_{6\times1} & 0_{6\times1} & 1_{6\times1} & 0_{6\times1} & 0_{6\times1} & 0_{6\times1} & 0_{6\times1} \\
0_{6\times1} & 0_{6\times1} & 0_{6\times1} & 0_{6\times1} & 1_{6\times1} & 0_{6\times1} & 0_{6\times1} & 0_{6\times1} \\
0_{6\times1} & 0_{6\times1} & 0_{6\times1} & 0_{6\times1} & 0_{6\times1} & 1_{6\times1} & 0_{6\times1} & 0_{6\times1} \\
0_{6\times1} & 0_{6\times1} & 0_{6\times1} & 0_{6\times1} & 0_{6\times1} & 0_{6\times1} & 1_{6\times1} & 0_{6\times1} \\
0_{6\times1} & 0_{6\times1} & 0_{6\times1} & 0_{6\times1} & 0_{6\times1} & 0_{6\times1} & 0_{6\times1} & 1_{6\times1}
\end{bmatrix}
\]

2.2  Separate-period univariate model

In addition to the traditional crossover model, Helms and Shen (1980) discussed the possibility of analyzing three-period crossover data one period at a time based on what is called "separate-period" univariate model. The model is of full rank and easy to interpret without reparametrizing as in the traditional LTFR model. Therefore, computation becomes straightforward and simple. Furthermore, the crossover model they generated can be used to obtain a clear description of the confounding patterns in the traditional model.

2.2.1  The model

The separate-period models for the three-period two-treatment crossover data are quite simple. Again let $Y_{p,tuv,k}$ denote the $p^{th}$ period observation from subject $k$, which was assigned to sequence $(t,u,v)$. Then for the three
period we have:

\[ p = 1: \quad E(\gamma_{1stuvk}) = \pi_1^+ (-1)^{12} \tau_1, \]
\[ p = 2: \quad E(\gamma_{2stuvk}) = \pi_2^+ (-1)^{12} \rho_{12} + (-1)^{12} \tau_2, \]
\[ p = 3: \quad E(\gamma_{3stuvk}) = \pi_3^+ (-1)^{12} \rho_{13} + (-1)^{12} \rho_{23} + (-1)^3 \tau_3 \]  

(2.2.1)

The parameters are described in the following sentences; their definitions as linear combinations of cell means, \( \mu_{ptuv} = E[\gamma_{ptuvk}] \), are given in Table 2.12.

- \( \pi_p \) is the "period p mean".

- \( \tau_p \) is the differential effect of treatment B compared with (minus) treatment A, in period p. Note that this model does not require that the treatment effect be identical in all periods.

- \( \rho_{pp'} \) is the carryover differential treatment effect (treatment B vs treatment A carryover) from period p to period p'.

Within each period one uses a general linear univariate model,

\[ E(\gamma_p) = X_p \beta_p, \quad V(\gamma_p) = \sigma_p^2 I, \]  

(2.2.2)

where \( \gamma_p \) is the vector of complete observations from period p, \( \beta_p \) is the vector of period p parameters (\( \pi_p, \tau_p, \) etc.), \( X_p \) is the appropriate period p design matrix, and \( \sigma_p^2 \) is the variance of period p observations.
TABLE 2.12 Subscript of Cell Mean \( \mu_{ptuv} \) within Period

<table>
<thead>
<tr>
<th>Period</th>
<th>Parameter</th>
<th>AAA</th>
<th>AAB</th>
<th>ABA</th>
<th>ABB</th>
<th>BAA</th>
<th>BAB</th>
<th>BBA</th>
<th>BBB</th>
<th>Row Divisor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \pi_1 )</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>( \tau_1 )</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>( \pi_2 )</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>( \tau_2 )</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>( p_{12} )</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>( \pi_3 )</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>( \tau_3 )</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>( p_{23} )</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

2.2.2 Estimation and hypothesis testing

To analyze the data with model (2.2.2), all parameters in \( \beta_p \)'s are estimable and the OLS estimators are BLUE's within each period. Hypothesis testing within each period is also straightforward, the one-sample exact t-test can be utilized to obtain the level of significance for within-period hypotheses. This one-model-per-period approach does not impose certain constraints which might be found in model (2.1.4). For instance, in the traditional combined-period model one might assume \( \tau_1 = \tau_2 = \tau_3 \) and only test for average treatment effects \( \tau = (\tau_1 + \tau_2 + \tau_3)/3 \). For the purpose of testing overall effects across periods in model (2.2.1), one would be interested in testing \( H_0: \tau_1 = \tau_2 = \tau_3 = 0 \) vs \( H_a: \) at least one \( \tau_p \neq 0 \). This hypothesis is easy to test using a Bonferroni approach. Similarly, overall residual effects are also tested by the Bonferroni procedure.

The separate-periods analysis does not support estimation or testing of period effects. Period effects can be analyzed using one of the other models.
2.2.3 An example

The mungbean data were analyzed using the separate-period model, with
parameters estimated by the OLS procedure.

For $p = 1$: \[ \hat{\beta}_1 = \begin{bmatrix} \pi_1 \\ \tau_1 \end{bmatrix} = \begin{bmatrix} 7.729 \\ -3.104 \end{bmatrix} \text{ with s.e. = 0.354,} \]

For $p = 2$: \[ \hat{\beta}_2 = \begin{bmatrix} \pi_2 \\ P_{12} \\ \tau_2 \end{bmatrix} = \begin{bmatrix} 8.687 \\ -2.271 \\ -0.896 \end{bmatrix} \text{ with s.e. = 0.576,} \]

For $p = 3$: \[ \hat{\beta}_3 = \begin{bmatrix} \pi_3 \\ P_{13} \\ P_{23} \\ \tau_3 \end{bmatrix} = \begin{bmatrix} 7.956 \\ -0.652 \\ -1.036 \\ -2.208 \end{bmatrix} \text{ with s.e. = 0.893} \]

Note that in this particular complete, balanced design, all estimates within
any one period have the same standard error.

To test the treatment effect hypotheses,

\[ H_0: \tau_p = 0 \quad \text{vs} \quad H_a: \tau_p \neq 0, \]

one for each period, one-sample t-tests reveal that treatment effects are
significantly different at period 1 and period 3, but only a moderate differ-
ence between tap water and salt water can be detected at period 2 ($p = 0.13$).
This indicates that there may be treatment by period interaction effects.
These interaction effects are usually assumed to be zero in the traditional
crossover model.
To examine the hypotheses of residual effects

\[ H_0: \rho_{pp'} = 0 \quad \text{vs} \quad H_a: \rho_{pp'} \neq 0, \]

one for each pair of \( p \) and \( p' \), \( t \)-statistics can be calculated similarly.

The results show that strong first-order residual effects are found in the second period but the residual effects are not significant in the third period. The ANOVA summarizing these results is shown in Table 2.13.

To test the combined-periods treatment effects,

\[ H_0: \tau_1 = \tau_2 = \tau_3 = 0 \quad \text{vs} \quad H_a: \text{at least one } \tau_p \neq 0, \]

the 90% and 95% Bonferroni joint confidence intervals are shown in Table 2.14. Even with 95% confidence, the data would still reject \( H_0 \) since treatment effects are significant at both period 1 and period 3. Similarly, in testing the hypothesis

\[ H_0: \rho_{12} = \rho_{23} = \rho_{13} = 0 \quad \text{vs} \quad H_a: \text{at least one } \rho_{pp'} \neq 0, \]

we have the same conclusion: reject \( H_0 \) with 95% confidence. Again this may be because of the strong level of significance on testing \( \rho_{12} = 0 \).

Note that the variance estimates for the three periods (the mean squared errors from the ANOVA table) increase from period to period. If the data from different periods were independent, which is not the case here, the ratio of these MSE's would have \( F \) distributions. The corresponding \( F \)-statistics were computed as a matter of curiosity and if there were independence, there would be strong evidence that the population variances increase from period to period. Traditional models, including the within-subjects model described earlier, assume homoscedasticity, an assumption which may well be violated by these data. This topic is discussed further in a subsequent section.
### TABLE 2.13  Separate-Periods ANOVA Tables

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>VR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Period 1 Overall mean</td>
<td>1</td>
<td>462.52</td>
<td>462.52</td>
<td>76.82*</td>
</tr>
<tr>
<td>Period 1 Treatment effects</td>
<td>1</td>
<td>462.52</td>
<td>462.52</td>
<td></td>
</tr>
<tr>
<td>Period 1 Error</td>
<td>46</td>
<td>276.96</td>
<td>6.02</td>
<td></td>
</tr>
<tr>
<td>Period 2 Overall mean</td>
<td>1</td>
<td>38.52</td>
<td>38.52</td>
<td>2.41</td>
</tr>
<tr>
<td>Period 2 Treatment effects</td>
<td>1</td>
<td>247.50</td>
<td>247.50</td>
<td>15.51*</td>
</tr>
<tr>
<td>Period 2 First-order residuals</td>
<td>1</td>
<td>247.50</td>
<td>247.50</td>
<td></td>
</tr>
<tr>
<td>Period 2 Error</td>
<td>45</td>
<td>718.27</td>
<td>15.96</td>
<td></td>
</tr>
<tr>
<td>Period 3 Overall mean</td>
<td>1</td>
<td>232.50</td>
<td>232.50</td>
<td>6.08*</td>
</tr>
<tr>
<td>Period 3 Treatment effects</td>
<td>1</td>
<td>51.49</td>
<td>51.49</td>
<td>1.35</td>
</tr>
<tr>
<td>Period 3 First-order residuals</td>
<td>1</td>
<td>20.43</td>
<td>20.43</td>
<td>0.53</td>
</tr>
<tr>
<td>Period 3 Second-order residuals</td>
<td>1</td>
<td>1683.81</td>
<td>38.27</td>
<td></td>
</tr>
</tbody>
</table>

*Significant at level $\alpha=0.05$

### TABLE 2.14  OLS Estimators and Bonferroni Confidence Intervals from Separate-Periods Analysis

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>$\hat{\theta}$</th>
<th>s.e.($\hat{\theta}$)</th>
<th>90% Bonferroni CI</th>
<th>95% Bonferroni CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lower</td>
<td>Upper</td>
</tr>
<tr>
<td>$\tau_1$</td>
<td>-3.104</td>
<td>0.354</td>
<td>-3.377</td>
<td>-2.831</td>
</tr>
<tr>
<td>$\tau_2$</td>
<td>-0.895</td>
<td>0.576</td>
<td>-1.621</td>
<td>-0.170</td>
</tr>
<tr>
<td>$\tau_3$</td>
<td>-2.201</td>
<td>0.893</td>
<td>-3.939</td>
<td>-0.462</td>
</tr>
<tr>
<td>$\rho_{12}$</td>
<td>-2.271</td>
<td>0.576</td>
<td>-2.996</td>
<td>-1.545</td>
</tr>
<tr>
<td>$\rho_{23}$</td>
<td>-1.036</td>
<td>0.893</td>
<td>-2.775</td>
<td>0.703</td>
</tr>
<tr>
<td>$\rho_{13}$</td>
<td>0.652</td>
<td>0.893</td>
<td>-2.391</td>
<td>1.087</td>
</tr>
</tbody>
</table>

|                   |                |                       | Lower             | Upper             |
|                   |                |                       |                   |                   |

-3.409 - 2.798
-1.709 - 0.082
-4.149 - 0.250
-3.085 - 1.457
-2.985 - 0.914
-2.602 - 1.297
2.2.4 Discussion

One disadvantage of the separate-period univariate model is its inadequacy for making intra-subject comparisons. However, the inclusion of pre-treatment measures and washout periods can improve the situation. In contrast, the separate-period univariate model also carries several advantages compared to the traditional combined-period model.

1) Incomplete data present no difficulties. The test statistics are still exact even when missing data occur. But the tests become inappropriate in the combined-period model if there are any missing values in the analysis, and the intra-subjects analysis fails.

2) Intra-subject correlations do not present a problem for the separate-period model as all observations within a given period are uncorrelated.

3) The separate-periods model allows for the possibility of different variances in different periods, which often occurs for biological data.

4) The separate-periods model includes the possibility that the magnitude of treatment effects varied from period to period. This is usually not permitted in traditional models.

5) The extension from three-period designs to any multiple-period designs is straightforward.

2.3 Separate-period multivariate model

Another way to present the separate-periods model is through the GLMM. The multivariate model described below will explore the difficulties associated with the analysis suggested.
2.3.1 The model

Let $u_k = (Y_{1kuv}, Y_{2kuv}, Y_{3kuv})^T$ be the observed responses from subject $k$ ($k=1,2,\ldots,n$); the GLMM setting of the separate-periods analysis can be described as

$$
E(\chi) = \chi \beta
$$
$$
V(\chi) = \Sigma \otimes I_n
$$

(2.3.1)

where $\chi(n \times 3) = [u_1, u_2, \ldots, u_n]^T$ represents the data matrix, and

$$
\beta_0 = [\beta_1^*, \beta_2^*, \beta_3^*]
$$

Note that the $\beta_i^*$'s include the sequence parameters defined within each period that are not confounded with the parameters of interest.

$\Sigma(3 \times 3)$ is the corresponding design matrix,

$$
\Sigma = \begin{bmatrix}
\sigma_{11} & \sigma_{12} & \sigma_{13} \\
\sigma_{21} & \sigma_{22} & \sigma_{23} \\
(Symm) & & \sigma_{33}
\end{bmatrix}
$$
is the covariance matrix of $u_k$. 

Obviously model (2.3.1) satisfies the assumption of GLMM of full rank. However, the interpretation of parameters in the model is rather difficult due to confounding. One may eliminate an entire row of $\beta$ and the corresponding columns of $\chi$ if the corresponding sequence effects are nil. Presumably one would not eliminate any of the first four rows of $\beta$ as these contain the treatment effects. In addition, if some subjects drop out of the study for causes unrelated to the treatments being administered, the complete information from these subjects should be omitted ("casewise deletion") from the analysis under the standard GLMM (2.1.1). With either a combined-period model or a separate-period univariate model, this information is still available in the analysis.

### 2.3.2 Estimation and hypothesis testing

With model (2.3.1) one can estimate $\beta$, $\Sigma$ and any secondary parameters of the form $\theta = \zeta \beta \mu - \theta_0$. For instance, to test the hypothesis of no period effects,

$$H_0: \pi_1 = \pi_2 = \pi_3 \quad \text{vs} \quad H_a: \text{at least one } \pi_i \neq \pi_j \quad (i \neq j),$$

we can let $\zeta = (1, 0, 0, 0, 0, 0, 0)$, 

$$\mu = \begin{bmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}$$

and generate Hotelling's $T^2$ statistic to perform an exact test. But for testing either overall treatment effects or treatment by period interactions effects, it is unfortunately not possible to obtain the elements of $\tau_1$, $\tau_2$ and $\tau_3$ in the form $\theta = \zeta \beta \mu - \theta_0$. This is also true for estimating any combined-period residual effects. In other words, the most straightforward method for performing the combined-period analysis under GLMM is to use the
Bonferroni procedures.

Under model (2.3.1), appropriate multivariate tests may be used to determine whether the data support the assumption that the covariance matrix $\Sigma$ has any particular structure, such as

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{bmatrix}$$  \hspace{1cm} (Morrison 1976).

2.3.3 An example

The mungbean data were analyzed under the multivariate model (2.3.1). The OLS estimates of $\beta$ can be shown as

$$\hat{\beta} = [\hat{\beta}_1^*, \hat{\beta}_2^*, \hat{\beta}_3^*] = \begin{bmatrix} 7.729 & 8.687 & 7.956 \\ -3.104 & -2.271 & -0.652 \\ -0.729 & -0.896 & -1.036 \\ -0.187 & -0.354 & -2.208 \\ 0.437 & 1.146 & -1.144 \\ -0.437 & 0.521 & -1.726 \\ 0.437 & 0.646 & 0.080 \\ -0.187 & -0.146 & 2.056 \end{bmatrix}$$

the s.e. for estimates in $\hat{\beta}_1^*$ is 0.338; for estimates in $\hat{\beta}_2^*$ it is 0.566 and in period 3 it is 0.815. The estimated variance-covariance matrix of $\Sigma$ is

$$\hat{\Sigma} = \begin{bmatrix} 5.5125 & 2.692 & 3.840 \\ & 15.138 & 2.690 \\ & & 31.872 \end{bmatrix}$$ (Symm)

Note that those parameters which were also defined in the separate-period univariate model have the same estimated values under model (2.3.1). Table 2.15 shows the levels of significance for testing each parameter defined in the multivariate model (2.3.1). According to our data, most of these
estimates whose true parameters correspond to sequence effects in $\beta$ are nonsignificant. Only 2 of the 15 tests are strongly significant ($p<0.05$) and one test is moderate ($p=0.05$). The result apparently supports the assumption of no sequence effects in the crossover model.

The estimate of $\Sigma$ suggests that the data are heterogeneous. The correlation matrix $\hat{\Sigma}$ is estimated by

$$\hat{\Sigma} = \begin{bmatrix} 1.00 & 0.29 & 0.29 \\ 0.29 & 1.00 & 0.12 \\ 0.29 & 0.12 & 1.00 \end{bmatrix}_{(Symm)}$$

The hypothesis of compound symmetry of $\Sigma$, say $H_0: \Sigma = \sigma^2 \begin{bmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{bmatrix}$ vs $H_a: \text{otherwise}$, is rejected by the data($\chi^2(4) = 30.69$). If we test the sphericity hypothesis $H_0: \Sigma = \sigma^2 I$ vs $H_a: \text{otherwise}$, the $\chi^2$ statistic also rejects $H_0$ strongly ($\chi^2(5) = 40.61$). These results suggest that the data do not support the usual assumption of homoscedasticity and independence made in the traditional analysis of crossover data.
<table>
<thead>
<tr>
<th>$p = 1$</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi_1$</td>
<td>7.729</td>
<td>0.338</td>
<td>22.81</td>
<td>0.0001*</td>
</tr>
<tr>
<td>$\tau_1$</td>
<td>-3.104</td>
<td>0.338</td>
<td>-9.16</td>
<td>0.0001*</td>
</tr>
<tr>
<td>$\lambda_{1BABA}$</td>
<td>-0.729</td>
<td>0.338</td>
<td>-2.15</td>
<td>0.375</td>
</tr>
<tr>
<td>$\lambda_{1BBA}$</td>
<td>-0.187</td>
<td>0.338</td>
<td>-0.55</td>
<td>0.583</td>
</tr>
<tr>
<td>$\lambda_{1AAA}$</td>
<td>0.437</td>
<td>0.338</td>
<td>1.29</td>
<td>0.204</td>
</tr>
<tr>
<td>$\lambda_{1ABA}$</td>
<td>-0.437</td>
<td>0.338</td>
<td>-1.29</td>
<td>0.204</td>
</tr>
<tr>
<td>$\lambda_{1AAB}$</td>
<td>0.437</td>
<td>0.338</td>
<td>1.29</td>
<td>0.204</td>
</tr>
<tr>
<td>$\lambda_{1AAA}$</td>
<td>-0.187</td>
<td>0.338</td>
<td>0.55</td>
<td>0.58</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$p = 2$</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi_2$</td>
<td>8.687</td>
<td>0.566</td>
<td>15.35</td>
<td>0.001*</td>
</tr>
<tr>
<td>$\rho_{12}$</td>
<td>-2.271</td>
<td>0.566</td>
<td>-4.01</td>
<td>0.001*</td>
</tr>
<tr>
<td>$\tau_2$</td>
<td>-0.896</td>
<td>0.566</td>
<td>-1.58</td>
<td>0.12</td>
</tr>
<tr>
<td>$\lambda_{2BBA}$</td>
<td>-0.354</td>
<td>0.566</td>
<td>-0.63</td>
<td>0.54</td>
</tr>
<tr>
<td>$\lambda_{2BAA}$</td>
<td>1.146</td>
<td>0.566</td>
<td>2.02</td>
<td>0.05*</td>
</tr>
<tr>
<td>$\lambda_{2AAA}$</td>
<td>0.521</td>
<td>0.566</td>
<td>0.92</td>
<td>0.36</td>
</tr>
<tr>
<td>$\lambda_{2AAB}$</td>
<td>0.646</td>
<td>0.566</td>
<td>1.14</td>
<td>0.26</td>
</tr>
<tr>
<td>$\lambda_{2AAA}$</td>
<td>-0.146</td>
<td>0.566</td>
<td>-0.26</td>
<td>0.80</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$p = 3$</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi_3$</td>
<td>7.96</td>
<td>0.815</td>
<td>9.76</td>
<td>0.001*</td>
</tr>
<tr>
<td>$\rho_{13}$</td>
<td>-0.65</td>
<td>0.815</td>
<td>-0.80</td>
<td>0.42</td>
</tr>
<tr>
<td>$\rho_{23}$</td>
<td>-1.035</td>
<td>0.815</td>
<td>-1.27</td>
<td>0.21</td>
</tr>
<tr>
<td>$\tau_3$</td>
<td>-2.208</td>
<td>0.815</td>
<td>-2.70</td>
<td>0.01*</td>
</tr>
<tr>
<td>$\lambda_{3BAA}$</td>
<td>-1.144</td>
<td>0.815</td>
<td>-1.40</td>
<td>0.17</td>
</tr>
<tr>
<td>$\lambda_{3ABA}$</td>
<td>-1.726</td>
<td>0.815</td>
<td>-2.12</td>
<td>0.04*</td>
</tr>
<tr>
<td>$\lambda_{3AAB}$</td>
<td>0.081</td>
<td>0.815</td>
<td>0.10</td>
<td>0.92</td>
</tr>
<tr>
<td>$\lambda_{3AAA}$</td>
<td>2.056</td>
<td>0.815</td>
<td>2.52</td>
<td>0.01*</td>
</tr>
</tbody>
</table>
CHAPTER III

MULTIVARIATE ANALYSIS OF INCOMPLETE DATA
FROM CROSSOVER DESIGNS

3.1 Introduction

The current statistical methods for the analysis of crossover designs are mainly concentrated on the application of the traditional combined-period univariate model. There seems to be little treatment of the missing data problem in the literature beyond Grizzle (1965). Subjects with incomplete data are usually deleted in the within-subject analysis. Deleting the non-missing data from these subjects leads to a loss of information. Although the separate-period analysis is an alternative to handle these difficulties, the method can not fully handle within-subject treatment comparisons. These difficulties seem to suggest that a multivariate analysis should be considered. In particular, the model should be able to handle missing data.

In this chapter two types of incomplete data models will be introduced. These models are derived from the more general linear multivariate (MGLM) model (Kleinbaum 1973) to accommodate the missing data problem in crossover designs. Both models take into account the intra-class correlation problem. The characteristics of these models and their relative efficiencies with respect to the separate-period analysis will be shown by means of a numerical example. Finally, alternative model presentations to deal with large data computations will be described.
3.2 Incomplete data models

The first incomplete data model is based on the separate-period univariate model (2.2.1). Let data be missing at random in the crossover design. Variances may be different from period to period and observations from the same subject are correlated but not necessarily equally correlated.

Let $\tilde{Y}_p^* (n \times 1)$ denote the data collected from period $p$ with missing values omitted. $\tilde{\xi}_p$'s($q\times 1$) are defined as $\tilde{\beta}_p$'s in Table 2.12. $D_p (n \times q)$ is the corresponding design matrix of $Y*$. Define

$\tilde{Y}_p^* = (\tilde{Y}_1^*, \tilde{Y}_2^*, \tilde{Y}_3^*)^T$; a vector of dimensions $\sum_{p=1}^{3} n \times 1$,

$\tilde{\xi}_p^* = (\tilde{\xi}_1^*, \tilde{\xi}_2^*, \tilde{\xi}_3^*)^T$; a vector of dimensions $\sum_{p=1}^{3} q \times 1$,

and $D^* = \begin{bmatrix} D_1 & 0 \\ 0 & D_3 \end{bmatrix}$ a matrix of dimensions $\sum_{p=1}^{3} n \times \sum_{p=1}^{3} q$.

Then the incomplete data model can be expressed as

$E(\tilde{Y}_p^*) = D^* \tilde{\xi}_p^*$

$V(\tilde{Y}_p^*) = \Omega^*$

where $\Omega^* = \begin{bmatrix} \sigma_{11}I_{n_1} & \sigma_{12}Q_{n_1n_2} & \sigma_{13}Q_{n_1n_3} \\ \sigma_{22}I_{n_2} & \sigma_{23}Q_{n_2n_3} \\ (\text{Symm}) & \sigma_{33}I_{n_3} \end{bmatrix}$

$\sigma_{pp}$ is the covariance of $Y_{p'tuvk}$ and $Y_{p'tuvk}$, $I_n$ is an identity matrix of order $n$, $Q_{n,n}$ is an $(n \times n)$ "binary matrix" with 1 in the (rth, sth) element if the rth observation of $Y_p^*$ and the sth observation of $Y_p^*$, are from the same subject; 0 otherwise. This is an MGLM model with p=3.
Further examining the structure of the incomplete data model (3.2.1), we find that the model allows not only different design matrices in different periods but also randomly missing data. Kleinbaum's (1970) "general incomplete model" (GIM) cannot be used here because the design matrix changes from period to period. The "multiple design matrices model" (MDM), or so-called "seemingly unrelated regressions" model cannot be used either, due to missing data. We shall call model (3.2.1) the "multivariate separate-period crossover" (MSPC) model since it is directly converted from the separate-period univariate model.

In general, it may be necessary to include concomitant variables in the model. These variables are frequently quantitative responses and may be assumed to have a common slope to subject responses. For instance, we may want to compare the post-treatment observations by including pre-treatment information as a covariable, or we will be interested in fitting a linear model with only one treatment parameter for all periods when the period by treatment interactions are not significant. In these situations model (3.2.1) may not be appropriate and the second incomplete data model can be introduced here.

Define $Y^*$, $Q^*$ to be the same as in (3.2.1) and (3.2.2) respectively. Let $Z^* = (Z_1^T, Z_2^T, Z_3^T, D_1^T, D_2^T, D_3^T)$; a vector of dimension $\sum_{p=1}^{4} q_p \times 1$,

and $Z_p = \begin{bmatrix} D_1 & 0 & Z_1 \\ D_2 & Z_2 \\ 0 & D_3 & Z_3 \end{bmatrix}$; a matrix of dimension $\sum_{p=1}^{3} n_p \times \sum_{p=1}^{4} q_p$,

where $\xi_p$ $(q_p \times 1)$ is the vector of parameters defined within period $p$, $D_1(q_4 \times 1)$ is so defined that parameters in $D$ are linear combinations of $Y_p$'s from at least two periods. $Z_p(n_p \times q_4)$ is the corresponding design matrix of $D$ in
period \( p \) and \( \tilde{Z} \) is not necessarily block diagonal. Then the incomplete
data model is expressed as

\[
\begin{align*}
E(Y^*) &= \tilde{Z}^* \eta^* \\
V(Y^*) &= \Omega^*
\end{align*}
\] (3.2.3)

We shall call model (3.2.3) the "multivariate combined-period crossover" (MCPC) model in the sense that it allows parameters defined from more than
a single period. Note that MCPC models do not fit in the MGLM model because
of the structure of \( \tilde{Z}^* \). Actually, model (3.2.3) is similar to the general
version of the GLUM model (2.1.4) except that the covariance matrix has the
form

\[
\Omega^* = \begin{bmatrix}
A_1 \tilde{Z} A_1^T & 0 \\
0 & A_2 \tilde{Z} A_2^T \\
& & \ddots \\
& & & A_n \tilde{Z} A_n^T
\end{bmatrix}
\] (3.2.4)

after some data permutation, where \( \tilde{Z} \) represents the unknown covariance
matrix of a complete set of data from a single subject.

3.3 Estimation and hypothesis testing

3.3.1 \( \tilde{Z} \) known and unrestricted

Suppose the covariance matrix of subject \( k \), \( \Sigma_k \), has the
structure of \( \Sigma_k = \sigma^2 V_k \) where \( \sigma^2 \) is an unknown scalar and \( V_k \) is a known
positive-definite symmetric matrix. Then \( V_k^{-1} \) can be partitioned through
Cholesky decomposition as \( V_k^{-1} = T_k^T \) where \( T_k \) is an upper triangular matrix
of the same dimension as \( V_k \). Pre-multiplying \( Y^* \) by \( T \), the MSPC models con-
vert to
\[ E(TY^*) = \tilde{T} D^* \tilde{\xi}^* \]
\[ V(TY^*) = \tilde{T} \Omega T^* = \sigma^2 \tilde{T} \]
\[ (3.3.1) \]
where
\[ \tilde{T} = \begin{bmatrix} \tilde{T}_1 & 0 \\ \vdots & \vdots \\ \tilde{T}_2 & 0 \\ \vdots & \vdots \\ \tilde{T}_n & 0 \end{bmatrix} \cdot \mathcal{P} \]
and \( \mathcal{P} \) is a permutation matrix.

defined as in (2.1.27). The structure of the covariance matrix shows that
\( TY^* \) are identical and independently distributed and model (3.3.1) satisfies
the GLUM assumptions. Hence we can apply the general linear model theory
to estimate parameters and test hypotheses. For instance, one of the contro-
versial issues about crossover designs is the assumption of no period by
treatment interactions. This can be investigated under the MSPC model by
testing the null hypothesis
\[ H_{01} : H_1 \tilde{\xi}^* = 0, \]
where
\[ H_1 = \begin{bmatrix} 0 & -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \]
and
\[ \tilde{\xi}^* = (\pi_1 \quad \tau_1 \quad \pi_2 \quad \rho_{12} \quad \tau_2 \quad \pi_3 \quad \rho_{13} \quad \rho_{23} \quad \tau_3)^T. \]

In addition, by combining treatment parameters from different periods we
can test the hypothesis of no average treatment effects under the same
model by generating
\[ H_{02} : H_2 \tilde{\xi}^* = 0, \]
where
\[ H_2 = \begin{bmatrix} 0 & c_1 & 0 & 0 & c_2 & 0 & 0 & c_3 \end{bmatrix}, c_p's \]are constants
representing weights assigned to each treatment. Letting \( c_p = 1/3 \) (p=1,2,3)
gives the same weight to all treatment effects in the model. We can also
construct the general linear hypothesis
\[ H_{03} : H_{33}^{c*} = 0, \]

where \[ H_3 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \]

and calculate the test statistics for the overall treatment effects. The same strategy can be utilized to test carryover effects. Note that none of these hypotheses has the \( C \bar{Q} U \) form in the standard multivariate model.

The choice of testing for either overall or average treatment effects in the MSPC model depends on the individual's interest. In general, the test for the overall effects is more likely to be rejected than that for average effects. For, instance, if one of the \( \tau_p \)'s were strongly significant, the overall tests would tend to reject \( H_{03} \) even if the other two \( \tau_p \)'s were nonsignificant. But the test statistics for average treatment effects \( H_{02} \) may not be significant since it averages the estimates over all three periods. However, if the error variance is large, the overall tests may not be significant due to the large standard errors. But the average effects can be significant since the inter-subject variability was removed. Note that the testing of average treatment effects does allow separate treatment parameters in the model.

### 3.3.2 \( \Sigma \) unknown and unrestricted

In general, \( \Sigma \) is unknown except for the restriction that it is positive-definite. An exact test for the general linear hypotheses under the MSPC model does not exist but an unbiased estimate of \( H \Sigma^{c*} \) and an asymptotic test can be obtained. Kleinbaum (1973) suggested using pairwise deletion to obtain an unbiased estimate of \( \Sigma \), which in turn provides the unbiasedness of \( \hat{H} \Sigma^{c*} \) by replacing \( \Sigma^* \) with \( \hat{\Sigma}^* \), which is a function of \( \hat{\Sigma} \) (see (3.2.4)), in the
weighted least squares (WLS) procedure. That is, using the special pattern
of design matrices in the MSCP model: $D_{pp'} = [D_p' \circ D_{p'}];$ where $p \geq p'$
and $D_{pp'},$ is a design matrix of period $p$ corresponding to data observed in
both periods $p$ and $p',$ and $R_{p'}$ is the additional matrix in $D_{pp'}$. Let
$\Sigma = (\sigma_{pp}),$ then unbiased estimates of $\sigma_{pp}$, using the pairwise deletion
method are

$$\hat{\sigma}_{pp} = \frac{1}{n_{pp} - \text{Rank}(D_{pp'})} \varphi_{xx}^T \left[ I_{n_{pp}} - D_p (D_p^T D_p)^{-1} D_p^T \right] \varphi_{xx}$$  \hspace{1cm} (3.3.2)

and

$$\hat{\sigma}_{pp'} = \frac{1}{n_{pp'} - \text{Rank}(D_{pp'})} \varphi_{xx}^T \left[ I_{n_{pp'}} - D_{p'} (D_{p'}^T D_{p'})^{-1} D_{p'}^T \right] \varphi_{xx}$$ \hspace{1cm} (3.3.3)

for $p' < p$, where $\varphi_{xx}$ is the data vector corresponding to $D_{pp'}$ and $n_{pp'}$ is
the number of observations in $\varphi_{xx}$. The mathematical proof of (3.3.2) and
(3.3.3) is given in the Appendix. The unbiased estimate of $\mathbb{H}_x^*$ for any con-
stant matrix $\mathbb{H}$ can be expressed as

$$\hat{\mathbb{H}}_x^* = \mathbb{H}(\varphi_{xx}^T \varphi_{xx}^{-1} \varphi_{xx}^T)^{-1} (\varphi_{xx}^T \varphi_{xx}^{-1} \varphi_{xx}^T)$$ \hspace{1cm} (3.3.4)

with $\hat{\mathbb{H}}_x^*$ estimated from (3.3.2) and (3.3.3). The asymptotic minimum variance
of $\hat{\mathbb{H}}_x^*$ can be estimated by

$$\hat{\nu}(\mathbb{H}_x^*) = \mathbb{H}(\varphi_{xx}^T \varphi_{xx}^{-1} \varphi_{xx}^T)^{-1} \mathbb{H}$$ \hspace{1cm} (3.3.5)

Note that, under normality $\hat{\mathbb{H}} = \hat{\mathbb{H}}_x^*$ is a "BAN" estimate of $\mathbb{H}_x^*$. This is
true for any unbiased estimate of $\mathbb{H}_x^*$ in (3.3.4). The procedure described
above is usually known as the "two-stage WLS" procedure since it requires
that $\Sigma$ be estimated by OLS and that $\mathbb{H}_x^*$ be estimated by WLS. Kakwani's (1967)
results, when applied to the MPC model setting, prove that if $\hat{\mathbb{H}}_x^*$ is unbiased
then $\hat{\mathbb{H}}_x^*$ from two-stage WLS is also unbiased.
A method frequently suggested to improve the numerical solution of $H \hat{\xi}^*$ is the iterative WLS (IWLS) technique. It computes an estimate of $\hat{\Omega}^*(\hat{\xi}^*)$ by OLS (pairwise deletion), an estimate of $H \hat{\xi}^*$ by WLS (two-stage WLS) and uses these as $\hat{\Omega}^*(0)$ and $H \hat{\xi}^*(0)$ in the following iterative procedure:

$$\hat{\xi}^{*}(i) = H \hat{\xi}^{*}(i-1) = H(D^*T_{\xi^*}^{-1}(D^*)^{-1}(D^*T_{\xi^*}^{-1} Y^*)$$

(3.3.6)

$$\hat{\xi}^{*}(i) = (Y^* - D^*\hat{\xi}^{*}(i)) (Y^* - D^*\hat{\xi}^{*}(i))^T$$

(3.3.7)

until $\hat{\xi}^{*}(i)$ converges. One difficulty is that $\hat{\xi}^{*}(i)$ may be close to singular, especially when residuals are highly correlated (Telser, 1968). An alternative is to use $\hat{\xi}^{*}(i)$ rather than $\hat{\xi}^{*}(i)$ in (3.3.7), i.e., estimate $\xi$ by taking the residual mean squares from the $i$th iteration. Then $\hat{\xi}^{*}(i)$ will not be singular unless the true $\xi$ is close to singularity. The OLS estimates, WLS estimates and IWLS estimates in the MDM model were shown in a simulation study by Kmenta and Gilbert (1967). Both WLS estimates and IWLS estimates are consistently better than OLS estimates in the small sample case, whereas the comparisons of WLS and IWLS estimates are somewhat mixed. In general, WLS estimates are no more than IWLS estimates except in autoregression models. But the results may not be directly applicable to the incomplete crossover models since missing data were not taken into account in their studies. The comparison of these methods under the general MGLM model is a topic for further research.

To test the hypotheses $H_{01}$, $H_{02}$ and $H_{03}$, the test statistics suggested by Wald (1943) using asymptotic theory are calculated by

$$W = (H \hat{\xi}^{*})^T (H(D^*T_{\xi^*}^{-1}D^*)^{-1} H^T)(H \hat{\xi}^{*})$$

(3.3.8)

Under the null hypothesis, asymptotically $W$ has a central $\chi^2$ distribution.
with rank(\(H\)) degrees of freedom. Note that the pairwise deletion estimate \(\hat{\Sigma}\) is not necessarily positive-definite, which might lead to negative values of \(W\) and raise a difficulty in interpretation. One can use either a "smoothing" technique to replace \(\hat{\Sigma}\) with the "closest" positive-definite covariance matrix (Schwertman and Allen, 1973) or assume a special covariance matrix pattern (LeVange and Helms, 1982), to obtain positive values for the test statistics. In fact any consistent estimate of \(\hat{\Sigma}\) can be used to calculate \(W\) for the purpose of testing hypotheses since the "BAN" theory only requires \(\hat{\Sigma}\) to be consistent.

3.4 Numerical example

To analyze the mungbean data with the MSPC model, 20% of the data were deleted at random. The incomplete data are presented in Table 3.1. One of the subjects (ID=2) in sequence '010' (ABA) was omitted completely. This gives us unbalanced sequences. The estimates of \(\xi^*\)'s and their standard errors from all three methods, are shown in Table 3.2. Table 3.2(A) gives the initial estimates of \(\xi^*\) using OLS procedures; Table 3.2(B), taking into account the information obtained from the correlation with the other two periods, gives the WLS estimates of \(\xi^*\), and Table 3.2(C) uses the iterative method (IWLS). Note that all estimates in the table are unbiased. The values of the WLS and IWLS estimates are similar, and the values of the OLS estimates are the same as the estimates from the separate-period analysis.

The estimated covariance/correlation matrices from OLS and IWLS are shown in Table 3.3. The similarity between \(\hat{\Sigma}_{OLS}\) and \(\hat{\Sigma}_{IWLS}\) reflects the quickness of convergence. In our data the 5th iteration narrows the difference of every element in \(\mid \hat{\xi}_{(5)} - \hat{\xi}_{(4)} \mid\) to within \(10^{-4}\). The correlations in both cases are too large to be ignored. \(\hat{\Sigma}\)'s also show the possibility of heter-
ogeneity and the negative correlation of period 2 and period 3 growth measurements suggests that growth in period 2 produces a reverse growth effect in period 3.

The variance-covariance matrix of an estimate of the coefficient vector, $\hat{\xi}^*$, is unknown and depends upon the variance-covariance matrix of $\hat{Y}^*$. Let $P \hat{\Omega}^* = V(\hat{Y}^*)$ which is block diagonal with $A_k \Sigma A_k$ as the $k^{th}$ block. (See equation 3.2.4). Let

$$\hat{\xi}^*_m = (D^* T \hat{\Omega}^*_m - 1 D^*)^{-1} (D^* T \hat{\Omega}^*_m - 1 Y^*)$$ (3.4.1)

where the subscripts $m$ and $r$ denote the estimation method: OLS, WLS or IWLS. Since $\hat{\Omega}^*_{\text{OLS}}$ is a (random) estimate, $V(\hat{\xi}^*_m)$ is difficult to ascertain. However, we can obtain an estimate of $V(\hat{\xi}^*)$ conditional upon $\hat{\Omega}^*_{\text{OLS}}$. Assuming $\hat{\Omega}^*_{\text{IWLS}}$ is in some sense the "best" of the three estimates (i.e. better than $\hat{\Omega}^*_{\text{OLS}}$ or $\hat{\Omega}^*_{\text{WLS}}$), we have the following estimators:

$$\hat{V}_{\text{IWLS}}(\hat{\xi}^*_m) = (D^* T \hat{\Omega}^*_{\text{IWLS}} - 1 D^* T \hat{\Omega}^*_{\text{IWLS}} D^* (D^* T \hat{\Omega}^*_{\text{IWLS}} - 1 D^* T \hat{\Omega}^*_{\text{IWLS}} D^* - 1)^{-1}$$ (3.4.2)

$$\hat{V}_{\text{IWLS}}(\hat{\xi}^*_{\text{OLS}}) = (D^* T \hat{\Omega}^*_{\text{OLS}} - 1 D^* T \hat{\Omega}^*_{\text{IWLS}} D^* (D^* T \hat{\Omega}^*_{\text{IWLS}} - 1 D^* T \hat{\Omega}^*_{\text{IWLS}} D^* - 1)^{-1}$$ (3.4.3)

$$\hat{V}_{\text{IWLS}}(\hat{\xi}^*_{\text{WLS}}) = (D^* T \hat{\Omega}^*_{\text{WLS}} - 1 D^* T \hat{\Omega}^*_{\text{IWLS}} D^* (D^* T \hat{\Omega}^*_{\text{IWLS}} - 1 D^* T \hat{\Omega}^*_{\text{IWLS}} D^* - 1)^{-1}$$ (3.4.4)

The true, unknown $V(\hat{\xi}^*)$ values depend upon $V(Y^*)$ and the method; there is no guarantee that the above estimators are better than others one might use, or even that these are "reasonably good" estimators. However, comparing their values is useful and interesting.

The standard errors of the elements of $\hat{\xi}^*$, computed from the estimated variance formula shown above, are presented in Table 3.4. Apparently both weighted methods typically produce smaller variations for the elements of $\hat{\xi}^*$ than the unweighted method. The differences between WLS estimates and IWLS estimates are somewhat mixed. The IWLS estimates have consistenly
smaller standard errors than the WLS estimates in the 20% missing data case, but the pattern is not so persistent for smaller missing data patterns. For all listed missing patterns, the standard errors of \( \hat{\tau}_2 \), \( \hat{\tau}_3 \) and \( \hat{\beta}_3 \) from WLS and IWLS are smaller than those of the other elements in the same \( \mathcal{F}_P^* \) (p=2,3). This may be caused by the special design matrix pattern in the MSPC model in which the standard error of \( \hat{\tau}_3 \) uses only information contained in \( \sigma_{33} \) and the standard error of \( \hat{\tau}_2 \) and \( \hat{\rho}_{23} \) involve \( \sigma_{22} \) and \( \sigma_{23} \).

One would expect the estimated secondary parameters shown in Table 3.5 to have similar results. The estimates of the effects are similar for all three methods, but notice considerable differences in the standard errors.

The ANOVA for analyzing the incomplete (20% missing) mungbean data is shown in Table 3.6. Note that all parameters are tested with all the other parameters in the model. There seems to be no period by first-order residual effect (\( \chi^2(1) = 0.929 \)) and no period effects (\( \chi^2(2) = 2.402 \)). Period by treatment interactions are significant (\( \chi^2(2) = 14.49 \)). Overall treatment and average treatment effects are significant with \( \chi^2(3) = 72.0216 \) and \( \chi^2(1) = 25.469 \) respectively. The different levels of significance are due to the significant period by treatment interactions. In particular, there is no treatment effect in period 2. Similarly, the overall residual effect is significant, but this is due to the highly significant first-order residual effect in the second period; the second-order residual effect is moderately significant (\( \chi^2(1) = 2.906 \)) but the third period first-order residual effect is not significant.

Note that the Wald statistics in the ANOVA table have values close to the squares of \( t \) statistics from the separate-period analysis, which suggests that the \( \chi^2 \) approximations to the statistics are fairly good in the cross-over data.
Since the second period treatment effect is not significant ($\chi^2(1) = 0.29$) it is reasonable to omit the parameter $\tau_2$ in the model. Similarly, we can eliminate the residual parameter $\rho_{23}$. In addition, since there is no significant difference between $\tau_1$ and $\tau_3$ ($\chi^2(1) = 0.567$), we shall combine $\tau_1$ and $\tau_3$ and fit a new model with the same procedure. The reduced linear model fits the MCPC setting. The estimates of $\hat{H}_0^*$, $\hat{z}$ and the ANOVA from analyzing the MCPC model are shown in Tables 3.7, 3.8 and 3.9 respectively. The results are similar to the analysis of MSPC models except that the average treatment effect is more significant in the MSPC model ($\chi^2(1) = 25.429$ vs $\chi^2(1) = 52.9987$).

3.5 Algorithmic issues

Since the analysis of incomplete data using the MGLM model is essentially based on large sample theory, theoretically we would like to collect as much information as possible. But the MGLM setting makes computation difficult to perform as the number of observations increase; large memory space and computing time are required to obtain the inverse of $\hat{G}^*$. Although simplifications have been derived for GIM and MDM models, none of these algorithms is applicable to the general cases. For example, in the GIM models, we compute parameter estimates with one design matrix for all periods with the same pattern of missing data. In the MDM model we compute parameter estimates by taking advantage of the fact that there are no missing data. Neither of the incomplete data models discussed, (3.2.1) and (3.2.3), has these characteristics. Therefore alternative algorithms need to be sought to implement the currently used models.

Recall the definition of $\chi^*$ from MSPC model. $\chi^*$ can be reshaped by pre-multiplying by a $(\prod_{p=1}^{3} n_p \times \prod_{p=1}^{3} n_p)$ permutation matrix $P$ such that observations
from the same subject are grouped together in ascending period order, i.e.,

\[ \mathbf{X}^\# = \mathbf{P}^\# \mathbf{X}^* \]

(3.5.1)

with

\[ \mathbf{X}^\# = [\mathbf{u}_1^\# T, \mathbf{u}_2^\# T, \ldots, \mathbf{u}_n^\# T] \]

(3.5.2)

then

\[ \mathbf{D}^\# = \mathbf{P}^\# \mathbf{D}^* \]

where

\[ \mathbf{D}^\# = [\mathbf{D}_1^\# T, \mathbf{D}_2^\# T, \ldots, \mathbf{D}_n^\# T] \]

(3.5.3)

\( \mathbf{D}_k^\# \) is a \((m_k \times q)\) matrix corresponding to \( \mathbf{\xi}_k^\# \). Hence model (3.2.1) can be expressed alternatively as

\[ \mathbf{E}(\mathbf{X}^\#) = \mathbf{D}^\# \mathbf{\xi}_0^\# \]

(3.5.4)

\[ \mathbf{V}(\mathbf{X}^\#) = \mathbf{\Omega}^\# \]

where \( \mathbf{\Omega}^\# \) is defined in (3.2.4), a block diagonal covariance matrix. Model (3.5.4) is known as the row-wise representation of MSPC model. Assuming \( \mathbf{\Sigma} \) is unknown, the WLS estimate of \( \mathbf{H}_0 \mathbf{\xi}^\# \) becomes

\[ \mathbf{H}_0 \mathbf{\xi}^\# = \mathbf{H}(\sum_{k=1}^n \mathbf{D}_k^\# \mathbf{\Sigma}_k^- \mathbf{D}_k^\#)^{-1} (\sum_{k=1}^n \mathbf{D}_k^\# \mathbf{\Sigma}_k^- \mathbf{u}_k^\#) \]

(3.5.5)

which can be expressed as

\[ \mathbf{H}_0 \mathbf{\xi}^\# = \mathbf{H}(\sum_{k=1}^n \mathbf{D}_k^\# \mathbf{\Sigma}_k^- \mathbf{D}_k^\#)^{-1} (\sum_{k=1}^n \mathbf{D}_k^\# \mathbf{\Sigma}_k^- \mathbf{u}_k^\#) \]

(3.5.6)

where \( \mathbf{\hat{\Sigma}}_k \) is the estimated covariance matrix and corresponds to the observed responses of subject \( k \). Furthermore, if we group subjects of the same missing data pattern together before the analysis, then (3.5.6) can be shown to be equivalent to

\[ \mathbf{H}_0 \mathbf{\xi}^\# = \mathbf{H}(\sum_{t=1}^T \sum_{k=1}^n \mathbf{D}_{kt}^\# \mathbf{\Sigma}_{kt}^- \mathbf{D}_{kt}^\#)^{-1} (\sum_{t=1}^T \sum_{k=1}^n \mathbf{D}_{kt}^\# \mathbf{\Sigma}_{kt}^- \mathbf{u}_{kt}^\#) \]

(3.5.7)

where \( t=1,2,\ldots,T \) indexes the missing data patterns, \( \mathbf{u}_{kt}^\# \) is the observed response vector of subject \( k \) in group \( t \) and \( n_t \) is the number of subjects in
group \( t \). The sample covariance matrix of \( \hat{H} \hat{\xi}^\# \) becomes

\[
\hat{\Sigma}(H \hat{\xi}^\#) = H(\sum_{t=1}^{T} \sum_{k=1}^{n_t} D_{kt}^{\#T} \hat{\Sigma}_{t}^{-1} D_{kt}^{\#})^{-1} H^T
\]  

(3.5.8)

The Wald statistic for any general linear hypothesis \( H_0: H \hat{\xi}^\# = \theta_0 \) can be calculated by

\[
W = (H \hat{\xi}^\# - \theta_0)^T [H^T(\sum_{t=1}^{T} \sum_{k=1}^{n_t} D_{kt}^{\#T} \hat{\Sigma}_{t}^{-1} D_{kt}^{\#})H]^{-1}(H \hat{\xi}^\# - \theta_0)
\]

(3.5.9)

Note that all these expressions [(3.5.1) to (3.5.9)] are directly applicable in the MCPC models since the pattern of \( D_k^{\#} \) is not specified.

The computational advantage of calculating \( H \hat{\xi}^\# \) with (3.5.6) is the tremendous saving of memory space in the computer. Let \( N \) = total number of observed responses, \( q \) = number of parameters in the model, and \( m_k \) = number of observed responses in subject \( k \). Calculating \( H \hat{\xi}^\# \) with (3.5.5) requires space to store an \((N \times N)\) matrix \( \hat{\Sigma} \) and an \((N \times q)\) matrix \( D \), which becomes impractical when \( N \) is large. The algorithm in (3.5.6) only needs to store the \((m_k \times m_k)\) matrix \( \hat{\Sigma}_k \) and an \((m_k \times q)\) matrix \( D_k \), where \( m_k \) and \( q \) are usually small relative to \( N \) in crossover studies. To further reduce the computing time, one can use (3.5.7). Using (3.5.6) would require computing \( \hat{\Sigma}_k^{-1} \) \( n \) times to obtain \( H \hat{\xi}^\# \) but using (3.5.7) would only compute \( \hat{\Sigma}_k \) \( T \) times. Note \( T \) is at most \( 2^p - 1 \).
<table>
<thead>
<tr>
<th>ID</th>
<th>SEQ</th>
<th>PERIOD 1</th>
<th>PERIOD 2</th>
<th>PERIOD 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>000</td>
<td>14</td>
<td>19</td>
<td>14</td>
</tr>
<tr>
<td>2</td>
<td>000</td>
<td>14</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>000</td>
<td></td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>000</td>
<td>11</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>5</td>
<td>000</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>000</td>
<td>7</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>001</td>
<td>15</td>
<td>12</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>001</td>
<td>12</td>
<td>15</td>
<td>17</td>
</tr>
<tr>
<td>3</td>
<td>001</td>
<td>11</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>001</td>
<td>12</td>
<td>18</td>
<td>19</td>
</tr>
<tr>
<td>5</td>
<td>001</td>
<td>14</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>001</td>
<td>8</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>010</td>
<td>6</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>010</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>010</td>
<td>12</td>
<td>8</td>
<td>18</td>
</tr>
<tr>
<td>4</td>
<td>010</td>
<td>14</td>
<td>9</td>
<td>18</td>
</tr>
<tr>
<td>5</td>
<td>010</td>
<td>7</td>
<td></td>
<td>14</td>
</tr>
<tr>
<td>6</td>
<td>010</td>
<td>11</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>011</td>
<td>10</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>011</td>
<td>7</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>011</td>
<td>13</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>011</td>
<td>10</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>011</td>
<td>13</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>011</td>
<td>8</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td></td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>7</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>5</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>100</td>
<td></td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>100</td>
<td>5</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>1</td>
<td>101</td>
<td>6</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>101</td>
<td></td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>101</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>101</td>
<td>5</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>101</td>
<td>3</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>101</td>
<td>2</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>110</td>
<td>3</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>110</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>110</td>
<td></td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>110</td>
<td>5</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>5</td>
<td>110</td>
<td></td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td>110</td>
<td>3</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>111</td>
<td>6</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>111</td>
<td>4</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>111</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>111</td>
<td>8</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>111</td>
<td></td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>111</td>
<td>2</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Parameter</td>
<td>Estimate</td>
<td>Standard Error</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td>----------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\pi_1$</td>
<td>7.8529</td>
<td>0.7053</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau_1$</td>
<td>-3.1470</td>
<td>0.7053</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\pi_2$</td>
<td>8.6116</td>
<td>0.6826</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_{12}$</td>
<td>-2.4616</td>
<td>0.6826</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\pi_3$</td>
<td>8.7645</td>
<td>0.7484</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_{13}$</td>
<td>-2.0638</td>
<td>0.7484</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_{23}$</td>
<td>-1.2346</td>
<td>0.7349</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau_3$</td>
<td>-2.3328</td>
<td>0.7349</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B. Weighted Least Squares Estimates of $\xi^*$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi_1$</td>
<td>7.8623</td>
<td>0.3881</td>
</tr>
<tr>
<td>$\tau_1$</td>
<td>-3.1161</td>
<td>0.3880</td>
</tr>
<tr>
<td>$\pi_2$</td>
<td>8.6217</td>
<td>0.6411</td>
</tr>
<tr>
<td>$\rho_{12}$</td>
<td>-2.3470</td>
<td>0.6411</td>
</tr>
<tr>
<td>$\pi_3$</td>
<td>8.5829</td>
<td>1.0052</td>
</tr>
<tr>
<td>$\rho_{13}$</td>
<td>-1.7562</td>
<td>1.0046</td>
</tr>
<tr>
<td>$\rho_{23}$</td>
<td>-1.2487</td>
<td>0.9490</td>
</tr>
<tr>
<td>$\tau_3$</td>
<td>-2.3832</td>
<td>0.9207</td>
</tr>
</tbody>
</table>

C. Iterative Weighted Least Squares Estimates of $\xi^*$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi_1$</td>
<td>7.8567</td>
<td>0.3861</td>
</tr>
<tr>
<td>$\tau_1$</td>
<td>-3.1147</td>
<td>0.3859</td>
</tr>
<tr>
<td>$\pi_2$</td>
<td>8.6178</td>
<td>0.6412</td>
</tr>
<tr>
<td>$\rho_{12}$</td>
<td>-2.3341</td>
<td>0.6412</td>
</tr>
<tr>
<td>$\pi_3$</td>
<td>8.5500</td>
<td>0.9999</td>
</tr>
<tr>
<td>$\rho_{13}$</td>
<td>-1.7035</td>
<td>0.9992</td>
</tr>
<tr>
<td>$\rho_{23}$</td>
<td>-1.2460</td>
<td>0.9299</td>
</tr>
<tr>
<td>$\tau_3$</td>
<td>-2.3827</td>
<td>0.8974</td>
</tr>
</tbody>
</table>
TABLE 3.3 Sample Variance-Covariance Matrices in MSPC Model

\[
\hat{\Sigma}_{\text{OLS}} = \begin{bmatrix}
6.0413 & 0.2757 & 0.3736 \\
2.8015 & 17.0956 & -0.1848 \\
5.5505 & -4.6198 & 36.5283
\end{bmatrix}
\]

\[
\hat{\Sigma}_{\text{IWLS}} = \begin{bmatrix}
6.0424 & 0.2895 & 0.4269 \\
2.9459 & 17.1319 & -0.1809 \\
6.3559 & -4.5364 & 36.6891
\end{bmatrix}
\]

TABLE 3.4 The Comparisons of Standard Errors of Estimates from MSPC Model

<table>
<thead>
<tr>
<th>Missing %</th>
<th>Parameters</th>
<th>S.E. (OLS)</th>
<th>S.E. (WLS)</th>
<th>S.E. (IWLS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete</td>
<td>π₁</td>
<td>0.3542</td>
<td>0.3542</td>
<td>0.3542</td>
</tr>
<tr>
<td></td>
<td>τ₁</td>
<td>0.3542</td>
<td>0.3542</td>
<td>0.3542</td>
</tr>
<tr>
<td></td>
<td>π₂</td>
<td>0.5798</td>
<td>0.5767</td>
<td>0.5798</td>
</tr>
<tr>
<td></td>
<td>ρ₁₂</td>
<td>0.5798</td>
<td>0.5767</td>
<td>0.5798</td>
</tr>
<tr>
<td></td>
<td>τ₂</td>
<td>0.5798</td>
<td>0.5485</td>
<td>0.5454</td>
</tr>
<tr>
<td></td>
<td>π₃</td>
<td>0.8971</td>
<td>0.8929</td>
<td>0.8971</td>
</tr>
<tr>
<td></td>
<td>ρ₁₃</td>
<td>0.8971</td>
<td>0.8929</td>
<td>0.8971</td>
</tr>
<tr>
<td></td>
<td>ρ₂₃</td>
<td>0.8971</td>
<td>0.8576</td>
<td>0.8534</td>
</tr>
<tr>
<td></td>
<td>τ₃</td>
<td>0.8971</td>
<td>0.8529</td>
<td>0.8479</td>
</tr>
<tr>
<td>10%</td>
<td>π₁</td>
<td>0.3795</td>
<td>0.3749</td>
<td>0.3737</td>
</tr>
<tr>
<td></td>
<td>τ₁</td>
<td>0.3795</td>
<td>0.3746</td>
<td>0.3733</td>
</tr>
<tr>
<td></td>
<td>π₂</td>
<td>0.5884</td>
<td>0.5847</td>
<td>0.5862</td>
</tr>
<tr>
<td></td>
<td>ρ₁₂</td>
<td>0.5884</td>
<td>0.5586</td>
<td>0.5543</td>
</tr>
<tr>
<td></td>
<td>τ₂</td>
<td>0.5884</td>
<td>0.9315</td>
<td>0.9379</td>
</tr>
<tr>
<td></td>
<td>ρ₁₃</td>
<td>0.9493</td>
<td>0.9315</td>
<td>0.9379</td>
</tr>
<tr>
<td></td>
<td>ρ₂₃</td>
<td>0.9493</td>
<td>0.8855</td>
<td>0.8698</td>
</tr>
<tr>
<td></td>
<td>τ₃</td>
<td>0.9493</td>
<td>0.8709</td>
<td>0.8505</td>
</tr>
<tr>
<td>20%</td>
<td>π₁</td>
<td>0.3969</td>
<td>0.3882</td>
<td>0.3862</td>
</tr>
<tr>
<td></td>
<td>τ₁</td>
<td>0.3969</td>
<td>0.3880</td>
<td>0.3859</td>
</tr>
<tr>
<td></td>
<td>π₂</td>
<td>0.6468</td>
<td>0.6412</td>
<td>0.6412</td>
</tr>
<tr>
<td></td>
<td>ρ₁₂</td>
<td>0.6468</td>
<td>0.6412</td>
<td>0.6412</td>
</tr>
<tr>
<td></td>
<td>τ₂</td>
<td>0.6468</td>
<td>0.6221</td>
<td>0.6198</td>
</tr>
<tr>
<td></td>
<td>π₃</td>
<td>1.0377</td>
<td>1.0052</td>
<td>0.9999</td>
</tr>
<tr>
<td></td>
<td>ρ₁₃</td>
<td>1.0377</td>
<td>1.0047</td>
<td>0.9992</td>
</tr>
<tr>
<td></td>
<td>ρ₂₃</td>
<td>1.0191</td>
<td>0.9491</td>
<td>0.9299</td>
</tr>
<tr>
<td></td>
<td>τ₃</td>
<td>1.0191</td>
<td>0.9207</td>
<td>0.8975</td>
</tr>
</tbody>
</table>
TABLE 3.5 Estimates of Secondary Parameters in MSPC Model

<table>
<thead>
<tr>
<th>Parameters</th>
<th>OLS</th>
<th></th>
<th></th>
<th>WLS</th>
<th></th>
<th></th>
<th>IWLS</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{H}_2$</td>
<td>s.e.</td>
<td>$\hat{H}_2$</td>
<td>s.e.</td>
<td>$\hat{H}_2$</td>
<td>s.e.</td>
<td>$\hat{H}_2$</td>
<td>s.e.</td>
<td></td>
</tr>
<tr>
<td>$\pi_2 - \pi_1$</td>
<td>0.7587</td>
<td>0.6718</td>
<td>0.7594</td>
<td>0.6654</td>
<td>0.7611</td>
<td>0.6595</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\pi_3 - \pi_1$</td>
<td>0.9116</td>
<td>0.9871</td>
<td>0.7205</td>
<td>0.9661</td>
<td>0.6933</td>
<td>0.9411</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau_1 - \tau_2$</td>
<td>-2.6854</td>
<td>0.7609</td>
<td>-2.7827</td>
<td>0.7341</td>
<td>-0.7820</td>
<td>0.7309</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau_1 - \tau_3$</td>
<td>-0.8142</td>
<td>1.0901</td>
<td>-0.7328</td>
<td>0.9953</td>
<td>-0.7319</td>
<td>0.9728</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_{12} - \rho_{23}$</td>
<td>-1.2269</td>
<td>1.2051</td>
<td>-1.0983</td>
<td>1.144</td>
<td>-1.088</td>
<td>1.1280</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(\tau_1 + \tau_2 + \tau_3)/3$</td>
<td>-1.9805</td>
<td>0.4227</td>
<td>-1.9442</td>
<td>0.3917</td>
<td>-1.9434</td>
<td>0.3851</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(\rho_{12} + \rho_{23})/2$</td>
<td>-1.8481</td>
<td>0.6044</td>
<td>-1.7979</td>
<td>0.5732</td>
<td>-1.7901</td>
<td>0.5653</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TABLE 3.6 The ANOVA Table in MSPC Model from Incomplete Mungbean Data

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>DF</th>
<th>SS(W)</th>
<th>$t^2$ statistics*</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_1$: First period treatment effect</td>
<td>1</td>
<td>65.1224**</td>
<td>62.88</td>
</tr>
<tr>
<td>$H_2$: Second period treatment effect</td>
<td>1</td>
<td>0.2881</td>
<td>0.504</td>
</tr>
<tr>
<td>$H_3$: Third period treatment effect</td>
<td>1</td>
<td>7.0487**</td>
<td>5.244</td>
</tr>
<tr>
<td>$H_4$: Overall treatment effects</td>
<td>3</td>
<td>72.0216**</td>
<td></td>
</tr>
<tr>
<td>$H_5$: Average treatment effect</td>
<td>1</td>
<td>25.429**</td>
<td></td>
</tr>
<tr>
<td>$H_6$: Second period residual effect</td>
<td>1</td>
<td>13.25**</td>
<td>14.51</td>
</tr>
<tr>
<td>$H_7$: Third period residual effect</td>
<td>1</td>
<td>1.795</td>
<td>1.464</td>
</tr>
<tr>
<td>$H_8$: Third period 2nd-order residual effect</td>
<td>1</td>
<td>2.906</td>
<td>3.96</td>
</tr>
<tr>
<td>$H_9$: Overall residual effects</td>
<td>3</td>
<td>20.636**</td>
<td></td>
</tr>
<tr>
<td>$H_{10}$: Average 1st-order residual effect</td>
<td>1</td>
<td>10.03**</td>
<td></td>
</tr>
<tr>
<td>$H_{11}$: Period effects</td>
<td>2</td>
<td>2.402</td>
<td></td>
</tr>
<tr>
<td>$H_{12}$: Period $\times$ treatment interactions</td>
<td>2</td>
<td>14.485**</td>
<td></td>
</tr>
<tr>
<td>$H_{13}$: Period $\times$ 1st-order residual interactions</td>
<td>2</td>
<td>0.929</td>
<td></td>
</tr>
</tbody>
</table>

* $t^2$ is calculated from separate-period analysis
** significant at $\alpha=0.05$. 
TABLE 3.7 Estimates of Parameters in MCPC Model

<table>
<thead>
<tr>
<th>Parameters</th>
<th>OLS</th>
<th></th>
<th>WLS</th>
<th></th>
<th>IWLS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \hat{\pi}_1 )</td>
<td>( \hat{\pi}_2 )</td>
<td>( \hat{\pi}_3 )</td>
<td>( \hat{\tau} )</td>
<td>( \hat{\rho}_{23} )</td>
<td>( \hat{\pi}_2 - \hat{\pi}_1 )</td>
</tr>
<tr>
<td>( \pi_1 )</td>
<td>8.0044</td>
<td>0.4195</td>
<td>7.9277</td>
<td>0.3877</td>
<td>7.8708</td>
<td>0.4052</td>
</tr>
<tr>
<td>( \pi_2 )</td>
<td>8.5899</td>
<td>0.7314</td>
<td>8.5784</td>
<td>0.6410</td>
<td>8.5767</td>
<td>0.7270</td>
</tr>
<tr>
<td>( \pi_3 )</td>
<td>8.7758</td>
<td>1.0509</td>
<td>8.5902</td>
<td>1.0035</td>
<td>8.5379</td>
<td>1.0100</td>
</tr>
<tr>
<td>( \tau )</td>
<td>-1.8503</td>
<td>0.6176</td>
<td>-1.5083</td>
<td>0.5161</td>
<td>-1.4871</td>
<td>0.5576</td>
</tr>
<tr>
<td>( \rho_{23} )</td>
<td>-1.9657</td>
<td>0.4370</td>
<td>-2.2606</td>
<td>0.3049</td>
<td>-2.3615</td>
<td>0.3244</td>
</tr>
<tr>
<td>( \pi_2 - \pi_1 )</td>
<td>0.5855</td>
<td>0.7592</td>
<td>0.6507</td>
<td>0.6647</td>
<td>0.7059</td>
<td>0.7476</td>
</tr>
<tr>
<td>( \pi_3 - \pi_1 )</td>
<td>0.7714</td>
<td>0.9799</td>
<td>0.6625</td>
<td>0.9638</td>
<td>0.6671</td>
<td>0.9276</td>
</tr>
</tbody>
</table>

TABLE 3.8 Sample Variance-Covariance Matrix/Correlation

\[
\hat{\Sigma}_{\text{IWLS}} = \begin{bmatrix}
6.6883 & 0.2593 & 0.4924 \\
3.1391 & 21.9171 & -0.0950 \\
7.8248 & -2.7328 & 37.7532 \\
\end{bmatrix}
\]

TABLE 3.9 The ANOVA Table in MCPC Model from Incomplete Mungbean Data

<table>
<thead>
<tr>
<th>Sources of variation</th>
<th>DF</th>
<th>SS(Wald)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average treatment effect</td>
<td>1</td>
<td>52.9987**</td>
</tr>
<tr>
<td>First-order residual effect</td>
<td>1</td>
<td>7.1132**</td>
</tr>
<tr>
<td>Second-order residual effect</td>
<td>1</td>
<td>1.0748</td>
</tr>
<tr>
<td>Period effects</td>
<td>2</td>
<td>1.7152</td>
</tr>
</tbody>
</table>

** significant at \( \alpha = 0.05 \)
CHAPTER IV

MULTIVARIATE ANALYSIS OF CROSSOVER DATA
WITH MODELLED COVARIANCE MATRICES

4.1 Introduction

A common characteristic of repeated measures designs is that the correlated data, when taken at equally spaced time intervals, can be observed to have special covariance matrix patterns. Knowledge of these patterns would allow us to represent the covariance matrix with fewer parameters. More importantly, the analysis procedure becomes more powerful when data have the assumed covariance matrix pattern. In this chapter, several time-related models will be selected to analyze the crossover data. These include the variance components model, autoregressive error models, and moving average models. MGLM theory is applied to deal with the missing data situation. Consistent estimates of specified covariance matrices will be sought to test linear hypotheses under asymptotic theory. A numerical example using mungbean data is given to compare the results with the unrestricted covariance model.

4.2 Covariance matrix selection

4.2.1 Models of homogeneous variances

In general, data observed from the same subject during different time periods are assumed to be correlated and have the same variance. Let $\Sigma$ be the covariance matrix from a single subject without missing data; the "var-
iance component" model would assume $\Sigma$ to be

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{bmatrix}$$

(4.2.1)

for the data from three equally spaced time intervals. If there are no concomitant variables, an important feature of this model is that the use of univariate analysis yields the same results as multivariate analysis (Graybill, 1976) in the complete data case. Usually this would not be true for the incomplete data case.

A second covariance matrix commonly seen in the analysis of repeated measures data is the "serial correlation" model. For $p=3$ the structure of $\Sigma$ can be seen as

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{bmatrix}.$$  

(4.2.2)

The error term $\varepsilon_{pk}$ of the serial correlation model can be represented by

$$\varepsilon_{pk} = \rho \varepsilon_{p-1,k} + u_{pk}$$  

(4.2.3)

where $\varepsilon_{pk}$ is the random error of subject $k$ in period $p$ and the $u_{pk}$'s are independent $N(0,\sigma^2_u)$. Potthoff and Roy (1964) suggested using this structure for $\Sigma$ in growth curve analysis. The implication of the serial correlation model is that the correlation coefficients decrease when the time between observations arising from the same subject increases. A more general case of model (4.2.2) can be formed by replacing $\rho^2$ with $\rho_2$, an extra parameter in $\Sigma$, i.e.
\[ \hat{\Sigma} = \sigma^2 \begin{bmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{bmatrix} \]  

(4.2.4)

The structure of (4.2.4) is usually known as the "second-order autocorrelation" matrix with p=3. The random error terms of the model have the relation

\[ \varepsilon_{pk} = \rho_p \varepsilon_{p-1,k} + u_{pk} \]  

(4.2.5)

where the \( u_{pk} \)'s are independent \( N(0, \sigma_u^2) \).

If \( |\rho_2| \) is small, the covariance structure of model (4.2.5) can be assumed to be

\[ \hat{\Sigma} = \sigma^2 \begin{bmatrix} 1 & \rho & 0 \\ \rho & 1 & \rho \\ 0 & \rho & 1 \end{bmatrix} \]  

(4.2.6)

which is usually known as the "successive correlation" matrix. Note that all models defined above have a uniform variance, \( \sigma^2 \), over periods. Data are assumed to be homoscedastic in these cases. In the time series literature, models (4.2.2), (4.2.4) and (4.2.6) are assumed to come from an infinite stationary time series \( \{\varepsilon_{pk}, -\infty < p < \infty\} \). Thus the error terms from the linear model would have arisen from an infinitive sequence \( \{\varepsilon_{pk}\} \).

The analysis of time-related data is presented in standard time series texts (e.g., Fuller, 1976). The general approach to avoid autocorrelation problems in the data analysis is through a linear transformation technique. For instance, to analyze crossover data with the serial correlation model (4.2.2), the correlation can be "taken away" by considering the transformation

\[ \tilde{Z} = AY \]  

(4.2.7)
where \( A = \begin{bmatrix} 1 & -\hat{\rho} & 0 \\ 0 & 1 & -\hat{\rho} \end{bmatrix} \otimes I_n \). \( \hat{\Sigma} \) is the observed data rolled out by rows and \( \hat{\rho} \) is a consistent estimate of \( \rho \). The transformed data \( \tilde{Z} \) are approximately independent and identically distributed, and can be analyzed using the reparametrization technique described in Chapter II. But the analysis can not avoid the action of dropping subjects with missing observations when data are incomplete.

We shall use the MSPC model (3.2.1) to analyze incomplete data with modelled covariance matrices. Assume the \( \Sigma \) is unknown except that it is positive-definite and has one of the patterns described above. Then the WLS procedure can be applied for estimation and hypothesis testing. The estimate \( H \hat{\Sigma}^* \) in (3.3.4) is unbiased provided that \( \hat{\Sigma} \) is unbiased. When an unbiased estimate of \( \Sigma \) is difficult to obtain, a consistent estimate of \( \Sigma \) can be used, in which case \( H \hat{\Sigma}^* \) is not unbiased, but consistent and BAN under the assumption of normality. The Wald statistic (W) can be generated and under \( H_0 \) is asymptotically \( \chi^2 \) distributed. The MGLM theory can be utilized to analyze incomplete crossover data only when a consistent estimate of \( \hat{\Sigma} \) can be found.

Le Vange (1983) shows that a consistent estimate of \( \hat{\Sigma} \) can be generated if there are consistent estimates of \( \sigma^2 \) and the \( \rho \)'s, where \( \sigma^2 \) and the \( \rho \)'s are the parameters defined in \( \hat{\Sigma} \). The basic idea is to use pairwise deletion to estimate \( \hat{\Sigma} \). Let \( \hat{\Sigma}_{u} \) denote the resulting estimator. For the covariance matrix (4.1.1), a consistent estimate of \( \sigma^2 \) can be obtained by averaging the diagonal elements of \( \hat{\Sigma}_{u} \). A consistent estimate of \( \rho \) can be computed by pooling the off-diagonal elements of \( \hat{\Sigma}_{u} \), the correlation matrix generated from \( \hat{\Sigma}_{u} \). The property of consistency is retained in the iterative WLS procedure since the newly generated covariance matrix \( \hat{\Sigma}_{u}(k) \) at the \( k^{th} \) iteration is a consistent estimate of \( \Sigma \) and \( \hat{\sigma}^2(k) \), \( \hat{\rho}(k) \) are consistent because they are the
averages of consistent estimates.

Similarly, we can find consistent estimates of $\sigma^2$ and $\rho$ in the serial correlation model by pooling diagonal elements of $\hat{\sum}_u$ for $\sigma^2$ and by pooling roots of off-diagonal elements of $\hat{\sum}_u$ for $\rho$. Similar methods can be applied to both the second-order autocorrelation model and the successive correlation model. The method becomes difficult when the pattern of $\hat{\sum}$ becomes more complicated.

4.2.2 Models of heterogeneous variances

A common characteristic of the covariance matrices in Section 4.2.1 is the homoscedasticity of variances in $\hat{\sum}$. All elements in $\hat{\sum}$ contain the same parameter $\sigma^2$. Since not all data encountered in practice are homogeneous, we shall introduce several unequal variance matrix patterns of $\hat{\sum}$ as follows.

Consider the first-order nonstationary autoregressive model

$$\varepsilon_{p,k} = \gamma \varepsilon_{p-1,k} + u_{p,k}$$

(4.2.8)

where $\varepsilon_{0,k} = 0$ and the $u_{p,k}$'s are independently distributed with mean zero. The variance of $u_{p,k}$ is $\sigma^2_p$. For $p \neq p'$ we have $\sigma^2_p \neq \sigma^2_{p'}$, in general. The covariance matrix of the $\varepsilon_{p,k}$'s can be represented by

$$\hat{\sum} = \begin{bmatrix}
\sigma^2_1 & \gamma \sigma_1^2 & \gamma^2 \sigma_1^2 \\
\gamma \sigma_1^2 & \sigma^2_2 + \sigma_2^2 & \gamma \sigma_2^2 + \gamma \sigma_1^2 + \gamma \sigma_2^2 \\
\gamma^2 \sigma_1^2 + \gamma \sigma_1^2 + \gamma \sigma_2^2 & \gamma \sigma_2^2 + \gamma \sigma_1^2 + \gamma \sigma_2^2 & \sigma^2_3 + \gamma \sigma_2^2 + \gamma \sigma_3^2 \\
\end{bmatrix}$$

(Symmetrical)

(4.2.9)

The pattern of $\hat{\sum}$ in (4.2.9) implies that the variance increases when the observed period $p$ increases. All elements in $\hat{\sum}$ contain the initial period variance $\sigma^2_1$. We shall call this the first-order nonstationary autoregressive model with multiple variances, abbreviated NSAR1MV.
Another nonstationary first-order autoregressive error model, slightly different from (4.2.8), assumes a common $\sigma^2$ for all periods but permits changes of the parameter $\gamma$ from period to period. Thus, let $\varepsilon_{pk}$ be the error term of subject $k$ in period $p$; then

$$
\varepsilon_{pk} = \gamma_p \varepsilon_{p-1,k} + u_{pk},
$$

where $\varepsilon_{0k} = 0$ and $u_{pk}$'s are independent $N(0, \sigma^2)$. Hence the covariance matrix has the structure:

$$
\Sigma = \sigma^2 \begin{bmatrix}
1 & \gamma_1 & \gamma_1 \gamma_2 \\
1+\gamma_1^2 & \gamma_2(1+\gamma_1^2) \\
1+\gamma_2^2 & \gamma_1 \gamma_2 \\
\gamma_2 & \gamma_1 (1+\gamma_2^2) & 1+\gamma_1^2 \gamma_2 \\
\end{bmatrix}.
$$

This model will be called a first-order nonstationary autoregressive model with multiple covariance parameters, denoted by NSAR1MC. Note that although $\Sigma$ of the NSAR1MC model contains a common quantity $\sigma^2$ for all elements, the values of the $\gamma$'s are not correlations and need not be less than 1 in absolute value.

Moving average models are another kind of time-related model. The general equation of a nonstationary (finite) moving average model for three periods repeated measures is given by

$$
\varepsilon_{pk} = u_{pk} + \gamma_1 u_{p-1,k} + \gamma_2 u_{p-2,k},
$$

where $u_{-1,k} = u_{0k} = 0$ and the other $u_{pk}$'s are independent $N(0, \sigma^2)$. The covariance matrix has the pattern

$$
\Sigma = \sigma^2 \begin{bmatrix}
1 & \gamma_1 & \gamma^2 \\
1+\gamma_1^2 & \gamma_1 (1+\gamma_2^2) \\
\gamma_2 & 1+\gamma_1^2 \gamma_2 \\
\end{bmatrix}.
$$
which is different from (4.2.11) only in the elements corresponding to $\sigma_{13}$ and $\sigma_{22}$. Another moving average model considered permits different variances for different periods, $\text{Var}(\varepsilon_{pk}) = \frac{\sigma^2}{p}$, but has a common covariance parameter $\gamma_p$. The covariance matrix is

$$
\mathbf{\Sigma} = \begin{bmatrix}
\sigma_1^2 & \gamma \sigma_1 \sigma_2 & \gamma \sigma_1 \\
\gamma \sigma_1 \sigma_2 & \sigma_2^2 + \sigma_3^2 & \gamma \sigma_2^2 + \gamma \sigma_2 \\
(\text{Symm}) & \gamma \sigma_1^2 + \gamma \sigma_2^2 + \sigma_3^2 & \end{bmatrix}.
$$

(4.2.14)

The error term model is

$$
\varepsilon_{pk} = u_{pk} + \gamma u_{p+k} + \gamma u_{p-k}.
$$

(4.2.15)

We shall call (4.2.12) the NSMA1MC model and (4.2.15) the NSMA1MV model.

One can use maximum likelihood for estimating the preceding nonstationary models when all data are present (Fuller, 1976). The MLE approach becomes very messy in the incomplete data case and generally there is no closed-form solution for the $\sigma^2$'s and $\gamma$'s when data are assumed to be missing at random. The work of La Vange (1983) using consistent estimation of $\mathbf{\Sigma}$ to obtain BAN estimation on $\xi_k$ is non-trivial since the $\sigma^2$'s and the $\gamma$'s are entangled in $\mathbf{\Sigma}$.

In order to use the IWLS estimation technique, at each iteration one must find an estimate of $\mathbf{\Sigma}$ which satisfies the model for $\mathbf{\xi}$. That is, given an unconstrained estimate, $\hat{\mathbf{\Sigma}}_u$, of $\mathbf{\Sigma}$, (from pairwise deletion, for example), one must estimate the parameters of $\mathbf{\Sigma}$ and compute an estimate, $\hat{\mathbf{\Sigma}}_{M}$, from the model equation (e.g. (4.2.9)). A naive method, based on the ANOVA method for variance component estimation, is to replace parameters by estimators in the covariance matrix model equations and solve for parameter estimates. The NSAR1MV model produces the following equations, where $\mathbf{g} = (s_{ij}) = \hat{\mathbf{\Sigma}}_u$:
Estimation "equations"  

\[ S_{11} = \hat{\sigma}_1^2 \]
\[ S_{12} = \hat{\gamma} \hat{\sigma}_1^2 \]
\[ S_{13} = \hat{\gamma} \hat{\sigma}_2^2 \]
\[ S_{22} = \hat{\gamma} \hat{\sigma}_1^2 + \hat{\sigma}_2^2 \]
\[ S_{23} = \hat{\gamma} \hat{\sigma}_1^2 + \hat{\gamma} \hat{\sigma}_2^2 + \hat{\sigma}_3^2 \]
\[ S_{33} = \hat{\gamma} \hat{\sigma}_1^2 + \hat{\gamma} \hat{\sigma}_2^2 + \hat{\sigma}_3^2 \]
\[ \hat{\sigma}_1^2 = S_{11} \]
\[ \hat{\gamma} = S_{12} / S_{11} \]
\[ \hat{\gamma} = S_{22} - S_{13} \] (assumes \( S_{22} > S_{13} \))
\[ \hat{\sigma}_2^2 = S_{33} - \hat{\rho} S_{23} \] (assumes \( S_{33} > \hat{\rho} S_{23} \))

These nonlinear "equations" are overdetermined and typically have no solution. One approximate solution, shown above, is easily obtained by "solving" for the model parameters \( \hat{\gamma} \), \( \hat{\sigma}_1^2 \). Other approximate solutions can be found easily. For example, one could use nonlinear least squares to find the "best" least squares approximate solution. It is useful to put the approximation in context: "solving" these equations is a part of one iteration in IWLS. One can accept a "quick" approximation, as above, or work very hard to obtain a good approximation, but at the next IWLS iteration the values of the \( s_{ij} \) are likely to make changes which are quite large relative to differences between the quick and refined approximate solutions. The refined approximations are probably not worth the work.

Moreover, the estimates produced by the approximate solution shown above are readily shown to be consistent. That is, the resulting \( \hat{\lambda}_u \) is consistent.

A similar process is easily performed for each of the models described in this section. Once one obtains a consistent estimation of \( \hat{\lambda}_u \), which satisfies the covariance model at each iteration of the IWLS algorithm, one can apply the IWLS method to compute BAN estimates of \( \hat{\xi}^* \) and Wald statistics.
4.3 Numerical example

Using the MSPC model (3.2.2), the incomplete mungbean data can be analyzed with selected patterned covariance matrices. The IWLS procedure converged for all but the NSMA1MC model (4.2.13). Table 4.1 shows the sample covariance/correlation matrix for each model. The homoscedastic models substantially overestimate the first period variance and underestimate the third period variance (compared with the unrestricted estimates in model (3.2.1)). For the heterogeneous variance models, model (4.2.11) obtains good estimates of \( \sigma_{11} \) and \( \sigma_{33} \) but the value of the estimated \( \hat{\sigma}_{22} \) is reduced. Only the NSAR1MV model (4.2.9) and NSMA1MV model (4.2.14), which assume unequal variances during different time periods, obtain fairly close estimates along the diagonal of \( \Sigma \). This indicates the incorrectness of the homoscedasticity assumption in the crossover data. The "between-period" covariances were not fitted well under these models' assumptions. Even the NSAR1MV model with a good estimate of \( \sigma_{23} \) has a negative estimate on \( \sigma_{12} \). This lack-of-fit of \( \hat{\Sigma} \) reveals the difficulty associated with the analysis procedure when there is a "reverse effect" among growth periods in the crossover experiments. 

It should be noted that the dependent variable is the amount of change within the period (post-period length minus pre-period length) and there is a washout period between a post-period measurement and a succeeding pre-period measurement. In many clinical trial crossover designs, the dependent variable is the measurement made at the end of a period, not the change since a preceding measurement. Serial correlation and moving average models would be expected to fit such data more closely than the change data analyzed here. On the other hand, it will be worthwhile to try another example to ensure the "reverse effect" is not due to randomness.
Table 4.2 presents the estimates of $H \xi^*$ under various models. The estimates of models (4.2.9), (4.2.11) and (4.2.14) are based on the non-linear regression estimates. All parameter estimates are similar to the estimates under model (3.2.1) except that the standard errors differ. Variances for models (4.2.1)-(4.2.6) seem to have similar magnitudes over periods due to homoscedasticity assumption in these models. The standard errors of estimates from these models are similar to the standard errors from OLS estimates presented in Table 3.2(A). For model (4.2.11) standard errors are close for estimates in period 1 and period 2 but different from period 3. The most striking change of standard errors is for $\hat{\tau}_3$ from model (4.2.11); the standard error is reduced to only half of the estimates from other models. The standard errors for parameter estimates of models (4.2.9) and (4.2.14), that have similar covariance matrices to model (3.2.1), are reasonably close.

Hypothesis testing yields the results presented in Table 4.3. It appears that levels of significance for the first period treatment effect ($H_4$) in the homogeneous variance models are largely reduced due to the increase of the first period variance estimate. Consequently the overall treatment effects ($H_4$) are less significant. The average treatment effects ($H_5$) and the treatment effects for the other two periods ($H_2$ and $H_3$) are not drastically affected by the change in variance estimators. The Wald statistics corresponding to period by treatment interaction effects ($H_{12}$) are reduced by half for the homogeneous variance models. The period by first-order residual interactions ($H_{13}$) are slightly increased. As for the residual effects, the statistics seem not to change too much among these homogeneous variance models since none of these tests involves period 1 observations. Thus the discrepancy between these models and the unconstrained
model is mainly due to the increase of first period variability in the structured covariance models. The only exception is related to the second-order residual effects which doubles the value of the statistic due to the moderate decrease of the third period variance.

The analysis of heterogeneous variance models results in a better approximation than the homogeneous cases except the NSARIMC model, which shows a strong level of significance for the second-period residual effect (H_3). The change is due to the reduction of second period variance. The levels of significance for the first and third period treatment effects (H_1 and H_3) are similar since there is a decrease in the standard error of \( \hat{\tau}_3 \) resulting from the application of the IWLS procedure. However, the analysis results are quite similar among the unconstrained model, the NSARIMV and NSMA1MV models. The results of using the NSARIMV model are slightly more comparable to model (3.2.1) than using the NSMA1MV model.

The ANOVA results using the approximation (instead of nonlinear estimation) procedure are shown in Table 4.7. Their sample variance-covariance matrices are presented in Table 4.1. There are some differences between the "best" estimates and the approximate estimates of \( \Sigma \), but the Wald statistics produced from both methods seem to be quite consistent. Both methods suggest that the NSARIMV and NSMA1MV models obtain closer results to model (3.2.1) than the other two models.

4.4 Discussion

Since the variance-covariance matrix estimates do not fit as well as expected, it is interesting to investigate the covariance matrix estimates from selected random samples of the data (i.e. 20% missing data in each sample). The results are shown in Table 4.4. Although the possibility
of obtaining an extreme case of $\hat{E}$ is small, we find only one of the six selected random samples yields $S_{23} > 0$; all other $\hat{E}$'s either have $S_{23}$ substantially less than zero or have a small negative correlation that might be ignored. Estimates of the other elements in $\hat{E}$ are fairly consistent among the six samples. In addition, Table 4.5 compares the test statistics based on three different missing patterns: complete data, 10% missing and 20% missing. The results seem to suggest that the incomplete data is still a good representative of the collected data.
### TABLE 4.1  Estimates of Covariance/Correlation Matrices from Selected Covariance Models

<table>
<thead>
<tr>
<th>1. Unrestricted Model 3.2.1</th>
<th>2. Variance Component Model 4.2.1</th>
</tr>
</thead>
</table>
| \[
\begin{bmatrix}
6.0424 & 2.9459 & 6.5559 \\
0.2895 & 17.1319 & -4.5365 \\
0.4269 & -0.1809 & 36.6891 \\
\end{bmatrix}
\] | \[
\begin{bmatrix}
19.8975 & 3.4284 & 3.4284 \\
0.1723 & 19.8975 & 3.4284 \\
0.1723 & 0.1723 & 19.8975 \\
\end{bmatrix}
\] |

<table>
<thead>
<tr>
<th>3. Serial Correlation Model 4.2.2</th>
<th>4. 2nd-order Serial Correlation Model 4.2.4</th>
</tr>
</thead>
</table>
| \[
\begin{bmatrix}
19.8889 & 1.2378 & 0.0770 \\
0.0622 & 19.8893 & 1.2378 \\
0.0039 & 0.0622 & 19.8893 \\
\end{bmatrix}
\] | \[
\begin{bmatrix}
19.9133 & 0.8297 & 8.5423 \\
0.0417 & 19.9133 & 0.8297 \\
0.4290 & 0.0417 & 19.9133 \\
\end{bmatrix}
\] |

<table>
<thead>
<tr>
<th>5. Successive Correlation Model 4.2.6</th>
</tr>
</thead>
</table>
| \[
\begin{bmatrix}
19.8889 & 0.9439 & 0 \\
0.0474 & 19.8889 & 0.9439 \\
0 & 0.0474 & 19.8889 \\
\end{bmatrix}
\] |

<table>
<thead>
<tr>
<th>6. NSAR1MV Model 4.2.9</th>
<th>7. NSAR1MC Model 4.2.11</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(a) Nonlinear Estimate</strong></td>
<td><strong>(b) Approximation</strong></td>
</tr>
</tbody>
</table>
| \[
\begin{bmatrix}
5.5482 & -1.5742 & 0.3531 \\
-0.1432 & 17.2027 & -4.4203 \\
0.0253 & -0.1763 & 36.5525 \\
\end{bmatrix}
\] | \[
\begin{bmatrix}
6.0528 & 3.0023 & 1.4892 \\
0.3464 & 12.4131 & 6.1571 \\
0.0935 & 0.2699 & 41.9123 \\
\end{bmatrix}
\] |

<table>
<thead>
<tr>
<th>8. NSMA1MV Model 4.2.14</th>
<th>9. NSMA1MC Model 4.2.13</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(a) Nonlinear Estimate</strong></td>
<td><strong>(b) Approximation</strong></td>
</tr>
</tbody>
</table>
| \[
\begin{bmatrix}
4.9907 & -0.5142 & -0.5142 \\
-0.0552 & 17.4096 & -1.7354 \\
-0.0381 & -0.0688 & 36.5411 \\
\end{bmatrix}
\] | \[
\begin{bmatrix}
6.0513 & 4.7192 & 4.7192 \\
0.4776 & 16.1339 & 13.3926 \\
0.3062 & 0.5322 & 39.2428 \\
\end{bmatrix}
\] |

* IWLS failed to converge, estimates from WLS procedure are used.
TABLE 4.2 Estimates of Primary and Secondary Parameters from Selected Covariance Models

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Model 3.2.1</th>
<th>Model 4.2.1</th>
<th>Model 4.2.2</th>
<th>Model 4.2.4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>s.e.</td>
<td>Estimate</td>
<td>s.e.</td>
</tr>
<tr>
<td>$\pi_1$</td>
<td>7.8567</td>
<td>0.3861</td>
<td>7.8500</td>
<td>0.7171</td>
</tr>
<tr>
<td>$\tau_1$</td>
<td>-3.1147</td>
<td>0.3859</td>
<td>-3.1194</td>
<td>0.7171</td>
</tr>
<tr>
<td>$\pi_2$</td>
<td>8.6178</td>
<td>0.6412</td>
<td>8.6304</td>
<td>0.6952</td>
</tr>
<tr>
<td>$\rho_{12}$</td>
<td>-2.3341</td>
<td>0.6412</td>
<td>-2.4493</td>
<td>0.6952</td>
</tr>
<tr>
<td>$\tau_2$</td>
<td>-0.3325</td>
<td>0.6198</td>
<td>-0.4217</td>
<td>0.6876</td>
</tr>
<tr>
<td>$\pi_3$</td>
<td>8.5500</td>
<td>0.9999</td>
<td>8.6873</td>
<td>0.7592</td>
</tr>
<tr>
<td>$\rho_{13}$</td>
<td>-1.7035</td>
<td>0.9992</td>
<td>-1.9624</td>
<td>0.7592</td>
</tr>
<tr>
<td>$\rho_{23}$</td>
<td>-1.2460</td>
<td>0.9299</td>
<td>-1.1617</td>
<td>0.7400</td>
</tr>
<tr>
<td>$\tau_3$</td>
<td>-2.3827</td>
<td>0.8974</td>
<td>-2.2678</td>
<td>0.7341</td>
</tr>
<tr>
<td>$\pi_2 - \pi_1$</td>
<td>0.7611</td>
<td>0.6595</td>
<td>0.7804</td>
<td>0.9254</td>
</tr>
<tr>
<td>$\pi_3 - \pi_1$</td>
<td>0.6933</td>
<td>0.9411</td>
<td>0.8373</td>
<td>0.9745</td>
</tr>
<tr>
<td>$\tau_1 - \tau_2$</td>
<td>-2.7820</td>
<td>0.7309</td>
<td>-2.6976</td>
<td>0.9955</td>
</tr>
<tr>
<td>$\tau_1 - \tau_3$</td>
<td>-0.7319</td>
<td>0.9728</td>
<td>-0.8515</td>
<td>1.0249</td>
</tr>
<tr>
<td>$\rho_{12} - \rho_{23}$</td>
<td>-1.0880</td>
<td>1.1280</td>
<td>-1.2876</td>
<td>1.0171</td>
</tr>
<tr>
<td>$(\tau_1 + \tau_2 + \tau_3)/3$</td>
<td>-1.9434</td>
<td>0.3851</td>
<td>-1.9363</td>
<td>0.4127</td>
</tr>
<tr>
<td>$(\rho_{12} + \rho_{23})/2$</td>
<td>-1.7901</td>
<td>0.5653</td>
<td>-1.8055</td>
<td>0.5068</td>
</tr>
<tr>
<td>Parameters</td>
<td>Model 4.2.6</td>
<td>Model 4.2.11</td>
<td>Model 4.2.14</td>
<td>Model 4.2.9</td>
</tr>
<tr>
<td>--------------------</td>
<td>------------</td>
<td>--------------</td>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>$\pi_1$</td>
<td>0.7199</td>
<td>0.7199</td>
<td>0.6692</td>
<td>0.7199</td>
</tr>
<tr>
<td>$\tau_1$</td>
<td>7.8756</td>
<td>7.7566</td>
<td>7.8457</td>
<td>7.8756</td>
</tr>
<tr>
<td>S.E.</td>
<td>0.3727</td>
<td>0.3727</td>
<td>0.3605</td>
<td>0.3727</td>
</tr>
<tr>
<td>$\pi_2$</td>
<td>-3.1440</td>
<td>-3.1520</td>
<td>-3.1520</td>
<td>-3.1440</td>
</tr>
<tr>
<td>$\tau_2$</td>
<td>8.6170</td>
<td>8.6016</td>
<td>8.6016</td>
<td>8.6170</td>
</tr>
<tr>
<td>S.E.</td>
<td>0.3529</td>
<td>0.3605</td>
<td>0.3605</td>
<td>0.3529</td>
</tr>
<tr>
<td>$\pi_3$</td>
<td>8.5852</td>
<td>8.5757</td>
<td>8.5757</td>
<td>8.5852</td>
</tr>
<tr>
<td>$\tau_3$</td>
<td>-2.4964</td>
<td>-2.4729</td>
<td>-2.4729</td>
<td>-2.4964</td>
</tr>
<tr>
<td>S.E.</td>
<td>0.4201</td>
<td>0.4201</td>
<td>0.4201</td>
<td>0.4201</td>
</tr>
<tr>
<td>$\rho_{12}$</td>
<td>-0.5329</td>
<td>-0.4902</td>
<td>-0.4902</td>
<td>-0.5329</td>
</tr>
<tr>
<td>$\rho_{13}$</td>
<td>1.0319</td>
<td>1.0319</td>
<td>1.0319</td>
<td>1.0319</td>
</tr>
<tr>
<td>$\rho_{23}$</td>
<td>-2.4964</td>
<td>-2.4729</td>
<td>-2.4729</td>
<td>-2.4964</td>
</tr>
<tr>
<td>$\tau_3$</td>
<td>8.7509</td>
<td>8.7509</td>
<td>8.7509</td>
<td>8.7509</td>
</tr>
<tr>
<td>S.E.</td>
<td>0.3727</td>
<td>0.3727</td>
<td>0.3727</td>
<td>0.3727</td>
</tr>
<tr>
<td>$\pi_3 - \pi_1$</td>
<td>-0.5329</td>
<td>-0.4902</td>
<td>-0.4902</td>
<td>-0.5329</td>
</tr>
<tr>
<td>$\tau_3 - \tau_1$</td>
<td>1.0319</td>
<td>1.0319</td>
<td>1.0319</td>
<td>1.0319</td>
</tr>
<tr>
<td>$\rho_{12}$</td>
<td>-0.4902</td>
<td>-0.4902</td>
<td>-0.4902</td>
<td>-0.4902</td>
</tr>
<tr>
<td>$\rho_{23}$</td>
<td>-2.4729</td>
<td>-2.4729</td>
<td>-2.4729</td>
<td>-2.4729</td>
</tr>
<tr>
<td>$\tau_3 - \tau_1$</td>
<td>8.7509</td>
<td>8.7509</td>
<td>8.7509</td>
<td>8.7509</td>
</tr>
<tr>
<td>S.E.</td>
<td>0.3727</td>
<td>0.3727</td>
<td>0.3727</td>
<td>0.3727</td>
</tr>
<tr>
<td>$\rho_{12}$</td>
<td>-0.4902</td>
<td>-0.4902</td>
<td>-0.4902</td>
<td>-0.4902</td>
</tr>
<tr>
<td>$\rho_{23}$</td>
<td>-2.4729</td>
<td>-2.4729</td>
<td>-2.4729</td>
<td>-2.4729</td>
</tr>
<tr>
<td>$\tau_3 - \tau_1$</td>
<td>8.7509</td>
<td>8.7509</td>
<td>8.7509</td>
<td>8.7509</td>
</tr>
<tr>
<td>S.E.</td>
<td>0.3727</td>
<td>0.3727</td>
<td>0.3727</td>
<td>0.3727</td>
</tr>
<tr>
<td>$\rho_{12}$</td>
<td>-0.4902</td>
<td>-0.4902</td>
<td>-0.4902</td>
<td>-0.4902</td>
</tr>
<tr>
<td>$\rho_{23}$</td>
<td>-2.4729</td>
<td>-2.4729</td>
<td>-2.4729</td>
<td>-2.4729</td>
</tr>
<tr>
<td>$\tau_3 - \tau_1$</td>
<td>8.7509</td>
<td>8.7509</td>
<td>8.7509</td>
<td>8.7509</td>
</tr>
<tr>
<td>S.E.</td>
<td>0.3727</td>
<td>0.3727</td>
<td>0.3727</td>
<td>0.3727</td>
</tr>
<tr>
<td>$\rho_{12}$</td>
<td>-0.4902</td>
<td>-0.4902</td>
<td>-0.4902</td>
<td>-0.4902</td>
</tr>
<tr>
<td>$\rho_{23}$</td>
<td>-2.4729</td>
<td>-2.4729</td>
<td>-2.4729</td>
<td>-2.4729</td>
</tr>
<tr>
<td>$\tau_3 - \tau_1$</td>
<td>8.7509</td>
<td>8.7509</td>
<td>8.7509</td>
<td>8.7509</td>
</tr>
<tr>
<td>S.E.</td>
<td>0.3727</td>
<td>0.3727</td>
<td>0.3727</td>
<td>0.3727</td>
</tr>
<tr>
<td>$\rho_{12}$</td>
<td>-0.4902</td>
<td>-0.4902</td>
<td>-0.4902</td>
<td>-0.4902</td>
</tr>
<tr>
<td>$\rho_{23}$</td>
<td>-2.4729</td>
<td>-2.4729</td>
<td>-2.4729</td>
<td>-2.4729</td>
</tr>
<tr>
<td>$\tau_3 - \tau_1$</td>
<td>8.7509</td>
<td>8.7509</td>
<td>8.7509</td>
<td>8.7509</td>
</tr>
<tr>
<td>S.E.</td>
<td>0.3727</td>
<td>0.3727</td>
<td>0.3727</td>
<td>0.3727</td>
</tr>
<tr>
<td>$\rho_{12}$</td>
<td>-0.4902</td>
<td>-0.4902</td>
<td>-0.4902</td>
<td>-0.4902</td>
</tr>
<tr>
<td>$\rho_{23}$</td>
<td>-2.4729</td>
<td>-2.4729</td>
<td>-2.4729</td>
<td>-2.4729</td>
</tr>
<tr>
<td>$\tau_3 - \tau_1$</td>
<td>8.7509</td>
<td>8.7509</td>
<td>8.7509</td>
<td>8.7509</td>
</tr>
<tr>
<td>S.E.</td>
<td>0.3727</td>
<td>0.3727</td>
<td>0.3727</td>
<td>0.3727</td>
</tr>
<tr>
<td>$\rho_{12}$</td>
<td>-0.4902</td>
<td>-0.4902</td>
<td>-0.4902</td>
<td>-0.4902</td>
</tr>
<tr>
<td>$\rho_{23}$</td>
<td>-2.4729</td>
<td>-2.4729</td>
<td>-2.4729</td>
<td>-2.4729</td>
</tr>
<tr>
<td>$\tau_3 - \tau_1$</td>
<td>8.7509</td>
<td>8.7509</td>
<td>8.7509</td>
<td>8.7509</td>
</tr>
<tr>
<td>S.E.</td>
<td>0.3727</td>
<td>0.3727</td>
<td>0.3727</td>
<td>0.3727</td>
</tr>
<tr>
<td>$\rho_{12}$</td>
<td>-0.4902</td>
<td>-0.4902</td>
<td>-0.4902</td>
<td>-0.4902</td>
</tr>
<tr>
<td>$\rho_{23}$</td>
<td>-2.4729</td>
<td>-2.4729</td>
<td>-2.4729</td>
<td>-2.4729</td>
</tr>
<tr>
<td>$\tau_3 - \tau_1$</td>
<td>8.7509</td>
<td>8.7509</td>
<td>8.7509</td>
<td>8.7509</td>
</tr>
<tr>
<td>S.E.</td>
<td>0.3727</td>
<td>0.3727</td>
<td>0.3727</td>
<td>0.3727</td>
</tr>
<tr>
<td>$\rho_{12}$</td>
<td>-0.4902</td>
<td>-0.4902</td>
<td>-0.4902</td>
<td>-0.4902</td>
</tr>
<tr>
<td>$\rho_{23}$</td>
<td>-2.4729</td>
<td>-2.4729</td>
<td>-2.4729</td>
<td>-2.4729</td>
</tr>
<tr>
<td>$\tau_3 - \tau_1$</td>
<td>8.7509</td>
<td>8.7509</td>
<td>8.7509</td>
<td>8.7509</td>
</tr>
<tr>
<td>S.E.</td>
<td>0.3727</td>
<td>0.3727</td>
<td>0.3727</td>
<td>0.3727</td>
</tr>
</tbody>
</table>
TABLE 4.3  Test Statistics for Important Hypotheses in the Selected Covariance Models

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>DF</th>
<th>Model (3.2.1)</th>
<th>Model (4.2.1)</th>
<th>Model (4.2.2)</th>
<th>Model (4.2.4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_1$</td>
<td>1</td>
<td>65.1224</td>
<td>18.9209</td>
<td>19.0609</td>
<td>19.0598</td>
</tr>
<tr>
<td>$H_2$</td>
<td>1</td>
<td>0.2881</td>
<td>0.3761</td>
<td>0.4121</td>
<td>0.4251</td>
</tr>
<tr>
<td>$H_3$</td>
<td>1</td>
<td>7.0487</td>
<td>9.5432</td>
<td>9.5122</td>
<td>11.7209</td>
</tr>
<tr>
<td>$H_4$</td>
<td>3</td>
<td>72.0216</td>
<td>28.7612</td>
<td>28.9842</td>
<td>30.9396</td>
</tr>
<tr>
<td>$H_5$</td>
<td>1</td>
<td>25.429</td>
<td>22.0148</td>
<td>22.2446</td>
<td>23.7287</td>
</tr>
<tr>
<td>$H_6$</td>
<td>1</td>
<td>13.25</td>
<td>12.411</td>
<td>12.4247</td>
<td>12.4412</td>
</tr>
<tr>
<td>$H_7$</td>
<td>1</td>
<td>1.795</td>
<td>2.4647</td>
<td>2.6302</td>
<td>3.0265</td>
</tr>
<tr>
<td>$H_8$</td>
<td>1</td>
<td>2.906</td>
<td>6.6808</td>
<td>7.1676</td>
<td>6.5937</td>
</tr>
<tr>
<td>$H_{10}$</td>
<td>1</td>
<td>10.03</td>
<td>12.6921</td>
<td>12.8827</td>
<td>13.9052</td>
</tr>
<tr>
<td>$H_{11}$</td>
<td>2</td>
<td>2.402</td>
<td>0.9713</td>
<td>0.8349</td>
<td>1.8003</td>
</tr>
<tr>
<td>$H_{12}$</td>
<td>2</td>
<td>14.485</td>
<td>7.7463</td>
<td>7.6393</td>
<td>7.5739</td>
</tr>
<tr>
<td>$H_{13}$</td>
<td>2</td>
<td>0.929</td>
<td>1.6024</td>
<td>1.4641</td>
<td>1.6196</td>
</tr>
<tr>
<td>Total SSR</td>
<td></td>
<td>613.142</td>
<td>363.25</td>
<td>429.79</td>
<td>380.587</td>
</tr>
</tbody>
</table>

*Hypotheses refer to Table 3.6.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>DF</th>
<th>Model (4.2.6)</th>
<th>Model (4.2.9)</th>
<th>Model (4.2.11)</th>
<th>Model (4.2.14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_1$</td>
<td>1</td>
<td>19.0624</td>
<td>71.61</td>
<td>47.3622</td>
<td>76.4206</td>
</tr>
<tr>
<td>$H_2$</td>
<td>1</td>
<td>0.4317</td>
<td>0.6896</td>
<td>1.2825</td>
<td>0.5670</td>
</tr>
<tr>
<td>$H_3$</td>
<td>1</td>
<td>9.6142</td>
<td>5.8741</td>
<td>36.484*</td>
<td>5.5008</td>
</tr>
<tr>
<td>$H_4$</td>
<td>3</td>
<td>29.098</td>
<td>78.1857</td>
<td>85.4913</td>
<td>82.527</td>
</tr>
<tr>
<td>$H_5$</td>
<td>1</td>
<td>22.4151</td>
<td>24.186</td>
<td>69.4682*</td>
<td>22.9697</td>
</tr>
<tr>
<td>$H_7$</td>
<td>1</td>
<td>2.6769</td>
<td>1.5951</td>
<td>3.0625</td>
<td>1.5876</td>
</tr>
<tr>
<td>$H_8$</td>
<td>1</td>
<td>7.2201</td>
<td>4.3186</td>
<td>8.0903</td>
<td>4.2783</td>
</tr>
<tr>
<td>$H_9$</td>
<td>3</td>
<td>22.5746</td>
<td>23.9298</td>
<td>218.7330*</td>
<td>21.5435</td>
</tr>
<tr>
<td>$H_{10}$</td>
<td>1</td>
<td>12.9736</td>
<td>9.8368</td>
<td>16.0884</td>
<td>9.6752</td>
</tr>
<tr>
<td>$H_{11}$</td>
<td>2</td>
<td>0.8993</td>
<td>1.7398</td>
<td>5.2699</td>
<td>1.5674</td>
</tr>
<tr>
<td>$H_{12}$</td>
<td>2</td>
<td>7.5819</td>
<td>12.5195</td>
<td>22.4389</td>
<td>12.8258</td>
</tr>
<tr>
<td>$H_{13}$</td>
<td>2</td>
<td>1.4488</td>
<td>1.0265</td>
<td>0.8118</td>
<td>0.9838</td>
</tr>
<tr>
<td>Total SSR</td>
<td></td>
<td>437.159</td>
<td>976.28</td>
<td>3260.16</td>
<td>941.566</td>
</tr>
</tbody>
</table>
TABLE 4.4  The Variance-Covariance Matrices Generated from Six Random Incomplete Samples in Mungbean Experiments

<table>
<thead>
<tr>
<th></th>
<th>Sample Currently Used for Analysis</th>
<th>4.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6.0413  2.8015  5.5505</td>
<td>6.4911  2.2042  8.6512</td>
</tr>
<tr>
<td></td>
<td>2.8015  17.0956  -4.6198</td>
<td>2.2042  12.0944  1.9060</td>
</tr>
<tr>
<td></td>
<td>5.5505  -4.6198  36.5283</td>
<td>8.6512  1.9060  48.5941</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>2.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6.2980  4.7081  6.0662</td>
</tr>
<tr>
<td></td>
<td>4.7081  18.7390  -0.2509</td>
</tr>
<tr>
<td></td>
<td>6.0662  -0.2509  39.5764</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>5.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6.2014  3.4617  7.6078</td>
</tr>
<tr>
<td></td>
<td>3.4617  15.2210  -1.7824</td>
</tr>
<tr>
<td></td>
<td>7.6078  -1.7824  32.0714</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>3.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6.4741  4.0024  6.2421</td>
</tr>
<tr>
<td></td>
<td>4.0024  16.0354  -0.2142</td>
</tr>
<tr>
<td></td>
<td>6.2421  -1.2142  36.5760</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>6.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5.9224  4.1246  3.4913</td>
</tr>
<tr>
<td></td>
<td>4.1246  14.0822  -2.5411</td>
</tr>
<tr>
<td></td>
<td>3.4913  -2.5411  41.2520</td>
</tr>
</tbody>
</table>
### TABLE 4.5 Comparison of Wald Statistics from Three Different Missing Patterns

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Complete (MDM)</th>
<th>10% Missing</th>
<th>20% Missing</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_1$</td>
<td>76.82</td>
<td>72.943</td>
<td>65.1224</td>
</tr>
<tr>
<td>$H_2$</td>
<td>0.811</td>
<td>0.695</td>
<td>0.288</td>
</tr>
<tr>
<td>$H_3$</td>
<td>6.128</td>
<td>9.574</td>
<td>7.049</td>
</tr>
<tr>
<td>$H_4$</td>
<td>83.759</td>
<td>83.652</td>
<td>72.022</td>
</tr>
<tr>
<td>$H_5$</td>
<td>28.396</td>
<td>33.709</td>
<td>25.429</td>
</tr>
<tr>
<td>$H_6$</td>
<td>15.339</td>
<td>12.401</td>
<td>13.25</td>
</tr>
<tr>
<td>$H_7$</td>
<td>0.298</td>
<td>0.088</td>
<td>1.795</td>
</tr>
<tr>
<td>$H_8$</td>
<td>0.529</td>
<td>0.181</td>
<td>2.906</td>
</tr>
<tr>
<td>$H_9$</td>
<td>16.153</td>
<td>12.953</td>
<td>20.636</td>
</tr>
<tr>
<td>$H_{10}$</td>
<td>7.038</td>
<td>4.925</td>
<td>10.03</td>
</tr>
<tr>
<td>$H_{11}$</td>
<td>3.015</td>
<td>3.518</td>
<td>2.402</td>
</tr>
<tr>
<td>$H_{12}$</td>
<td>16.2093</td>
<td>16.926</td>
<td>14.485</td>
</tr>
<tr>
<td>$H_{13}$</td>
<td>3.059</td>
<td>2.951</td>
<td>0.929</td>
</tr>
<tr>
<td>Total SSR</td>
<td>640.528</td>
<td>583.475</td>
<td>613.142</td>
</tr>
</tbody>
</table>

### TABLE 4.6 Test Statistics for Important Hypotheses in the Heteroscedastic Models Using Approximation Method

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>NSAR1MV Model</th>
<th>NSAR1MC Model</th>
<th>NSM1MV Model</th>
<th>NSM1MC Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_1$</td>
<td>63.5805</td>
<td>64.5297</td>
<td>65.0337</td>
<td>102.146</td>
</tr>
<tr>
<td>$H_2$</td>
<td>0.3450</td>
<td>0.6722</td>
<td>0.1425</td>
<td>1.2512</td>
</tr>
<tr>
<td>$H_3$</td>
<td>4.2191</td>
<td>21.4882</td>
<td>5.0259</td>
<td>39.3066</td>
</tr>
<tr>
<td>$H_4$</td>
<td>68.2590</td>
<td>87.0773</td>
<td>70.3626</td>
<td>141.173</td>
</tr>
<tr>
<td>$H_5$</td>
<td>20.3903</td>
<td>60.9150</td>
<td>22.3924</td>
<td>98.9453</td>
</tr>
<tr>
<td>$H_6$</td>
<td>19.3339</td>
<td>34.8568</td>
<td>15.2202</td>
<td>54.1066</td>
</tr>
<tr>
<td>$H_7$</td>
<td>0.8708</td>
<td>2.1323</td>
<td>0.5661</td>
<td>6.3904</td>
</tr>
<tr>
<td>$H_8$</td>
<td>3.0376</td>
<td>9.7450</td>
<td>2.8888</td>
<td>10.8598</td>
</tr>
<tr>
<td>$H_9$</td>
<td>20.9275</td>
<td>37.9821</td>
<td>15.8555</td>
<td>60.9626</td>
</tr>
<tr>
<td>$H_{10}$</td>
<td>7.9729</td>
<td>23.0457</td>
<td>7.3621</td>
<td>45.8897</td>
</tr>
<tr>
<td>$H_{11}$</td>
<td>1.5386</td>
<td>2.1614</td>
<td>1.6163</td>
<td>10.0005</td>
</tr>
<tr>
<td>$H_{12}$</td>
<td>18.7122</td>
<td>26.6755</td>
<td>18.386</td>
<td>40.8464</td>
</tr>
<tr>
<td>$H_{13}$</td>
<td>1.3249</td>
<td>5.9588</td>
<td>2.0049</td>
<td>8.5547</td>
</tr>
<tr>
<td>Total SSR</td>
<td>639.9800</td>
<td>733.5950</td>
<td>559.0690</td>
<td>1181.7000</td>
</tr>
</tbody>
</table>
CHAPTER V

SUMMARY AND SUGGESTIONS FOR FURTHER RESEARCH

5.1 Summary

This study was initiated as an attempt to use a full-rank linear model to analyze crossover data. The traditional less-than-full-rank model may sometimes cause confusion because of the two-fold confounding: from the residual effects and from the overparametrization. Until this work, the analysis of crossover designs has been limited to the use of univariate models with the assumption of equal correlation among the data. It usually ignores the second-order residual effects and sometimes the first-order residual effects in order to obtain a "clean" estimate of the treatment effect.

Chapter II presented the analysis procedures for complete crossover data. Analysis using the traditional combined-period univariate model was discussed. The work of Koch (1972) is extended to the three-period case in which there are actually two analyses performed with transformed data. These transformed data are independent and identically distributed. The intra-subject analysis estimates all parameters except sequence effects. The inter-subject analysis would also allow testing treatment effects and first- and second-order residual effects provided that sequence effects are zero.

No period effect can be estimated in the inter-subject analysis. In contrast, the within-period parameters were clearly defined without reparametrization and transformation in the separate-period analysis. This analysis also indicates the confounding patterns were due to sequence effects. The attempt
to use a standard multivariate model by combining separate-period models together was also presented to explore the difficulties associated with the method.

The idea of analyzing incomplete crossover data with multivariate models was discussed in Chapter III. The application of Kleinbaum's (1973) MGLMM theory to crossover designs was presented. Two incomplete data models were converted from the separate-period univariate model to accommodate the multivariate analysis. It was shown that the tests of overall treatment effects and average treatment effects are possible within the same model. Several solutions for estimating $H \xi^*$ were presented and compared. The values of estimates are similar but their standard errors vary. In general, OLS estimates have more variation than WLS estimates and IWLS estimates in the incomplete crossover data case. The differences between WLS and IWLS estimates are mixed. However, WLS estimates are more likely to have large standard errors than IWLS estimates according to the numerical example.

Chapter IV presents attempts to develop more powerful procedures for data analysis by specifying certain constraints on the variance-covariance matrix in the multivariate model. The incomplete data model with homogeneous variances failed to fit the crossover data from the mungbean experiment. Using selected time-series models might improve the estimates of $\Sigma$ and therefore the estimate of $H \xi^*$. The difference between these constrained heterogeneous models and the unconstrained models is slight in the numerical example. But discrepancy still persisted because of the unusual "reverse effect" between the second period data and the third period data.

The nature of the mungbean experiment conducted for this research is to observe the growth of beansprouts during the period of application. It
was intended to detect the difference between treatment (salt water) and placebo (tap water). When the pattern of treatments sequence becomes the main interest instead of the individual treatment differences, one should modify the model so that parameters corresponding to sequence effects will be the main interest and should not be deleted. This analysis can be specified as the analysis of sequence patterns instead of crossover analysis.

5.2 Suggestions for further research

The separate-period univariate model and multivariate model developed in the preceding chapters are not restricted to eight-sequence crossover designs. Any two treatment crossover designs with no design confounding can utilize these models. A study which investigates three-period and four-period crossover designs that can use separate-period models will be extremely helpful to the experimenter in planning an experiment to which these models applied. In other words, one may be able to choose a "better" design from these selected crossover designs by comparing the variance estimates of the treatment effects.

Another possible extension of this work is related to the loss of efficiency in parameter estimation when missing data occurs. One would rather lose the first period data than the succeeding periods data when the estimates of residual effects are important. Assuming data are missing at random and assuming a formula which shows the relative loss of efficiency in the treatment estimates can be derived, it could help to determine the sample size required to achieve a given power of a statistical test.

Simplification needs to be sought to replace the double iteration procedure suggested in Chapter IV. One alternative is the maximum likelihood method. Although the MLE approach does not necessarily yield a closed form solution for $\hat{\xi}^*$ and $\Sigma(\sigma^2, \rho)$ for incomplete data, it is easy to try with
three-period designs. In particular, one should consider the hierarchical
designs discussed by Trawinski and Bargmann (1964) in the sense that cross-
over designs are more likely to lose information in the succeeding periods
than in preceding periods.

A simulation study may be helpful to determine which of the three esti-
mation procedures (OLS, WLS and IWLS) is most efficient in the incomplete
data case.

Finally, the analysis of crossover designs with censored data should
be studied for the purpose of handling censorings in the crossover designs.
Mathematical proof of using pairwise deletion methods to estimate $\Sigma$ in MSPC model.

For (3.3.2), the result is implied directly from MGLM theory (Kleinbaum, 1973). For (3.3.3), where $p \neq p'$, we have

$$
[I - D_{pp'}^T (D_{pp'}^T D_{pp'})^{-1} D_{pp}] [I - D_{p'p} (D_{p'p}^T D_{p'p})^{-1} D_{p'p}] = I - D_{pp'}^T (D_{pp'}^T D_{pp'})^{-1} D_{pp'}^T - D_{p'p} (D_{p'p}^T D_{p'p})^{-1} D_{p'p}^T
$$
$$
+ D_{pp'} (D_{pp'}^T D_{pp'})^{-1} D_{pp'} D_{p'p} (D_{p'p}^T D_{p'p})^{-1} D_{p'p}
$$

(A.1)

Note that for $p > p'$, $D_{pp'} = [D_{p'p}, Z_{p'p}]$,

$$
D_{pp'}^T D_{p'p} = \begin{bmatrix}
D_{p'p}^T D_{p'p} & D_{p'p}^T Z_{p'p} \\
Z_{p'p}^T D_{p'p} & Z_{p'p}^T Z_{p'p}
\end{bmatrix},
$$

and from

Searle (1967, Section 8.7) we have

$$
[D_{pp'}^T D_{p'p}]^{-1} = [(D_{p'p}^T D_{p'p})^{-1} + (D_{p'p}^T D_{p'p})^{-1} (D_{p'p}^T Z_{p'p}) W (Z_{p'p}^T D_{p'p}) (D_{p'p}^T D_{p'p})^{-1} - W (Z_{p'p}^T D_{p'p}) (D_{p'p}^T D_{p'p})^{-1} (D_{p'p}^T Z_{p'p}) W]
$$

where $W$ is a function of $D_{p'p}$ and $Z_{p'p}$, which implies that

$$
[D_{pp'}^T D_{p'p}]^{-1} = \begin{bmatrix}
I_{pp'} \\
0
\end{bmatrix};
$$

therefore (A.1) becomes

$$
[I - D_{pp'} (D_{p'p}^T D_{p'p})^{-1} D_{p'p}] = \begin{bmatrix}
I_{pp'} \\
0
\end{bmatrix};
$$

and the trace of (A.2) becomes $n_{pp'} - \text{Rank}(D_{p'p})$. 
MACRO PAIRDEL
* USING PAIRWISE DELETION METHOD TO ESTIMATE THE COVARIANCE MATRIX;
* INITIAL VARIANCE-COVARIANCE MATRIX = SIGMA;
SIGMA=J.(NCOL(Y),NCOL(Y),0);
DFMX=SIGMA;
SSE=SIGMA;
* ENTERING DO LOOP TO CALCULATE SS AND DF, THEN SIGMA;
DO T=1 TO NCOL(Y);
  DO S=1 TO T;
    IST=YINDEX(S,T);
    K=0;
    Z=J.(IST(+),NCOL(X),0);
    U=J.(IST(+),2,0);
DO R=1 TO NROW(Y);
   IF IST(R) = 1 THEN DO;
      * SELECT NON-MISSING DATA IN PERIODS S AND T;
      K = K + 1;
      Z(K) = X(R);
      U(K, 1) = Y(R, S);
      U(K, 2) = Y(R, T);
   END;
   END;
XS = Z(1:1:S+1); XT = Z(1:T+1);
DFS = I.(IST(+,)) - XT*INV(XS'*XS)*XS';
DFT = I.(IST(+,)) - XT*INV(XT'*XT)*XT';
MU = DFS*DFT;
DFMX(S, T) = TRACE(MU);
FREE MU;
SSE(S, T) = (DFS*U(1:1:1))*(DFT*U(1:2:1));
SIGMA(S, T) = SSE(S, T)/=DFMX(S, T);
DFMX(T, S) = DFMX(S, T);
SIGMA(T, S) = SIGMA(S, T);
END;
END;

MACRO AE1
*------ NONLINEAR ESTIMATION FOR PARAMETERS OF AUTOREGRESSIVE MODEL -----
S = U(1:1:1); P1 = U(2:1:1); P2 = U(3:1:1);
S11 = SIGMA(1, 1); S12 = SIGMA(1, 2); S13 = SIGMA(1, 3);
S21 = SIGMA(2, 1); S22 = SIGMA(2, 2); S23 = SIGMA(2, 3);
S31 = SIGMA(3, 1); S32 = SIGMA(3, 2); S33 = SIGMA(3, 3);
F = (S-S11)*(P1*P2*S-S13)/((1+P1**2)*S-S22)/
   (P2*(1+P1**2)*S-S23)/((1+P2**2+P1**2*P2**2)*S-S33);
DELTA = 1.0;
DC WHILE (MAX(ABS(Delta)) > 0.01);
J = (1 0 0)/(P1||S||0)/((P1*P2)||P2*S||(P1*S))/((1+P1**2)||(2*P1*S)
   ||(0))/((P2||(1+P1**2)||(2*P1*P2*S)||(1+P1**2)*S))/((1+P2**2+P1**2*P2
   **2)||(2*P1*P2**2*S)||(2*P2*S+2*P2*P1*P1*S));
DELTA = INV(J'*J)*J'*(-F);
U = U+DELTA;
S = U(1:1:1); P1 = U(2:1:1); P2 = U(3:1:1);
F = (S-S11)*(P1*P2*S-S13)/((1+P1**2)*S-S22)/
   (P2*(1+P1**2)*S-S23)/((1+P2**2+P1**2*P2**2)*S-S33);
END;
SIGMA = (((S)/(P1*S))/((P1*P2*S)))||(P1*S)/((1+P1**2)*S)/((P2*(1+P1**2)*S))
   ||((P1*P2*S))/((P2*(1+P1**2)*S))/((1+P2**2+P1**2*P2**2)*S));

MACRO AE1
*------ NONLINEAR ESTIMATION FOR PARAMETERS OF AUTOREGRESSIVE MODEL -----
S = U(1:1:1); P1 = U(2:1:1); P2 = U(3:1:1);
S11 = SIGMA(1, 1); S12 = SIGMA(1, 2); S13 = SIGMA(1, 3);
S21 = SIGMA(2, 1); S22 = SIGMA(2, 2); S23 = SIGMA(2, 3);
S31 = SIGMA(3, 1); S32 = SIGMA(3, 2); S33 = SIGMA(3, 3);
F = (S-S11)*(P1*P2*S-S13)/((1+P1**2)*S-S22)/
   (P2*(1+P1**2)*S-S23)/((1+P2**2+P1**2*P2**2)*S-S33);
DELTA = 1.0;
DC WHILE (MAX(ABS(Delta)) > 0.01);
J = (1 0 0)/(P1||S||0)/((P1*P2)||P2*S||(P1*S))/((1+P1**2)||(2*P1*S)
   ||(0))/((P2||(1+P1**2)||(2*P1*P2*S)||(1+P1**2)*S))/((1+P2**2+P1**2*P2
   **2)||(2*P1*P2**2*S)||(2*P2*S+2*P2*P1*P1*S));
DELTA = INV(J'*J)*J'*(-F);
U = U+DELTA;
S = U(1:1:1); P1 = U(2:1:1); P2 = U(3:1:1);
F = (S-S11)*(P1*P2*S-S13)/((1+P1**2)*S-S22)/
   (P2*(1+P1**2)*S-S23)/((1+P2**2+P1**2*P2**2)*S-S33);
END;
SIGMA = (((S)/(P1*S))/((P1*P2*S)))||(P1*S)/((1+P1**2)*S)/((P2*(1+P1**2)*S))
   ||((P1*P2*S))/((P2*(1+P1**2)*S))/((1+P2**2+P1**2*P2**2)*S));
MACRO MA1

*------- NONLINEAR ESTIMATION FOR COVARIANCE OF MOVING AVERAGE MODEL -------*

S=U(1,); P1=U(2,); P2=U(3,);
S11=SIGMA(1,1); S12=SIGMA(1,2); S13=SIGMA(1,3);
S22=SIGMA(2,2); S23=SIGMA(2,3);
S33=SIGMA(3,3);

F=(S-S11)//(P1*P1*S-S12)//(S2*S-S13)//((1+P1*P1)*S-S22)//
(1*(1+P1*P1)*P1*P2*P2)*S-S33);
DELTA=1.0;
DO WHILE (MAX(ABS(DELTA))>0.01);

J=(1 0 0)//(P1||S||O)//(P2||O||S)//((1+P1*P1)||(2*P1*P1)||(0))
//((P1*(1+P2))/((1+P2)*S)||S33))//((1+P1*P1+P2*P2)||2*P1*P1)
||(2*P2*P2));
DELTA=INV(J'*J)*J*(-F);
U=U+DELTA;
S=U+DELTA;
S11=SIGMA(1,1); S12=SIGMA(1,2); S13=SIGMA(1,3);
S22=SIGMA(2,2); S23=SIGMA(2,3);
S33=SIGMA(3,3);

END;

MACRO IGLS

* LIMIT = 0.001;
EPOCH=1.0;
MAXITEA=10;
DO N=1 TO MAXITEA WHILE (EPSILON>_LIMIT_);
HBOLD=HBETA;
GLES;
RSC=J.(NCOY, NCOY, 0);

DO R=1 TO IRW(Y);
RSQ=RSC+(Y(R,) * BLOCK(X(R,C1) * YINDEX(R,1), X(R,C2) * YINDEX(R,2), X(R,C3) * YINDEX(R,3)) * BETA) * (Y(R,) * BLOCK(X(R,C1) * YINDEX(R,1), X(R,C2) * YINDEX(R,2), X(R,C3) * YINDEX(R,3)) * BETA);

END;

SIGMA=ESQ#/DFMX;
IF N=1 THEN DO;
BETA=HBETA;
XIV1=XIVX;
SIGMA=SIGMA;
END;
DEB=ABS(HBETA-HBOLD);
EPSILON=MAX(DEB);
END;

% END OF MACRO ITERATION;
MACRO GLES
* USING GENERALIZED LEAST SQUARE METHOD TO ESTIMATE PARAMETERS - BETA;
* THE MACRO SHOULD NOT BE USED IF DATA WERE NOT SORTED BY MISSING PATTERNS;
XIVX=J(NPARK,NPARK,0);
XIVY=J(NPARK,1,0);
XVV=XIVX;
**** CHANGE THE INITIAL VALUES OF XIVX & XIVY IF THE DESIGN MATRICES DIFFER;
DO R=1 TO NROW(Y);
  K=1;
  A=J(YINDEX(R,1),NCOL(Y),0);        * INCIDENCE MATRIX FOR REDUCING THE
  * DIMENSION OF SIGMA;
  YVEC=Y(R,:);
  * XVEC SHOULD BE RE-DEFINED FOR COMBINED PERIOD MULTIVARIATE MODEL *;
  XVEC=BLOCK(X(R,C1),X(R,C2),X(R,C3));
  DO C=1 TO NCOL(Y);
    IF YINDEX(R,C)=1 THEN DO;
      A(K,C)=1;
      K=K+1;
    ENDDO;
  ENDDO;
END;
BETA=SOLVE(XIVX,XIVY);
HBETA=H*BETA;
* END OF MACRO GLES - GENERALIZED LEAST SQUARES METHOD;
%

MACRO INITIAL
NOTE INITIAL ESTIMATE OF SIGMA - COVARIANCE MATRIX - USING PAIRWISE DELETION;
PFOINT SIGMA COLNAME=ZNAME ROWNAME=ZNAME;
NOTE INITIAL ESTIMATE OF BETA - USING ORDINARY LEAST SQUARE METHOD;
PFOINT BETA COLNAME=YNAME ROWNAME=BETNAME;
NOTE STANDARD DEVIATIONS OF HBETA;
PFOINT STDHB COLNAME=YNAME ROWNAME=BETANAME;
* END OF MACRO INITIAL ESTIMATION;
MACRO TESTGLH
HEETA0=H*BETAO;
STDHE0=(VSE*VECDIAG(H*INV(XIVX0)*H'))##0.5;
STDHE3=(VECDIAG(H*INV(XIVX0)*XIV*INV(XIVX0)*H'))##0.5;
HEETA=HEETA0||STDHE0||STDHE3;
NOTE SKIP=2 OLS ESTIMATE OF HBETA;
PRINT HEETA COLNAME=YNAME ROWNAME=BETANAME;
HEETA1=H*BETAO1;
STDHE1=(VECDIAG(H*INV(XIVX1)*H'))##0.5;
HEETA1=HEETA1||STDHE1;
NOTE SKIP=5 NON-ITERATIVE WLS ESTIMATE OF HBETA;
PRINT HEETA1 COLNAME=YNAME ROWNAME=BETANAME;
NOTE SKIP=2 NON-ITERATIVE WLS ESTIMATE OF SIGMA;
PRINT SIGMA1 COLNAME=ZNAME ROWNAME=ZNAME;
HEETA=H*BETAA;
STDHE=(VECDIAG(H*INV(XIVX)*H'))##0.5;
W=HEETA**INV(H*INV(XIVX)*H')*HBETA;
W=W/||NROW(H);
HBETA=HEETA1||STDHE;
NCTE SKIP=5 ********************************************;
NOTE * TOTAL NUMBER OF ITERATIONS *
NCTE ********************************************;
PRINT N;
NOTE SKIP=5 ********************************************;
NOTE * FINAL ESTIMATE OF PARAMETERS - HBETA *
NOTE * ********************************************;
PRINT HBETA COLNAME=YNAME ROWNAME=BETANAME;
NCTE SKIP=5 ********************************************;
NOTE * FINAL ESTIMATE OF COVARIANCE - SIGMA *
NCTE ********************************************;
PRINT SIGMA COLNAME=ZNAME ROWNAME=ZNAME;
NCTE SNAP W ********************************************;
NOTE * CHI-SQUARES STATISTICS - W *
NOTE * ********************************************;
PRINT W;
* END OF MACRO TESTGLH - TEST GENERAL LINEAR HYPOTHESIS;*
* MAIN PROGRAM STARTS HERE;
*;
*;
PROC MATRIX FUZZ;
FETCH X DATA=DISK.CROSS00F(KEEP=X1-X4);
FETCH Y DATA=DISK.CROSS00F(KEEP=PERIOD1-PERIOD3);
FETCH YINDEX DATA=DISK.CROSS00F(KEEP=INDEX1-INDEX3);
UNAME='BETA' 'OLSSTD' 'IWLSTD';
YNAME='BETA' 'STDBETA';
ZNAME='PERIOD1' 'PERIOD2' 'PERIOD3';
BETANAME='PERIOD1' 'TREAT1' 'PERIOD2' 'TREAT2' 'PERIOD3' 'CARRY12' 'CARRY23' 'TREAT3';
* ESTIMATE INITIAL VALUE OF BETA AND SIGMA WITH ORDINARY
* LEAST SQUARE METHOD;
DF=YINDEX*YINDEX;
C1=1:2; C2=1:3; C3=1:4;
NPARK=NCOL(C1)+NCOL(C2)+NCOL(C3);
H=I(NPARK);
BETA=J(NPARK,1,0);
SIGMA=I(NCOL(Y));
GLES;
PAIRDEL;
U=(7.9 / -.25 / -1.7);
BETA0=BETA;
XIVX=XIVX;
MSE=SUM(VECDIAG(SSE))#/SUM(VECDIAG(DFMX));
STDHB=(MSE*VECDIAG(H*INV(XIVX)*H'))##0.5;
INITIAL;
IGLS;
NOTE PAGE TEST OVERALL PARAMETERS IN THE MODEL;
TESTGLH;
NOTE PAGE TEST PERIOD BY TREATMENT INTERACTION EFFECTS;
H=(0 1 0 0 -1 0 0 0 0 / 0 1 0 0 0 0 0 0 -1);
BETANAME='TRT1_2' 'TRT1_3';
TESTGLH;
NOTE PAGE TEST PERIOD BY FIRST ORDER RESIDUAL INTERACTION EFFECTS;
H=(0 0 0 1 0 0 0 -1 0);
BETANAME='CARY12_23';
TESTGLH;
NOTE PAGE TEST OVERALL TREATMENT EFFECTS;
H=(0 1 0 0 0 0 0 0 / 0 0 0 0 1 0 0 0 0 / 0 0 0 0 0 0 0 1);
BETANAME='TREAT1' 'TREAT2' 'TREAT3';
TESTGLH;
NOTE PAGE TEST OVERALL RESIDUAL EFFECTS;
H=(0 0 0 1 0 0 0 0 0 / 0 0 0 0 0 1 0 0 / 0 0 0 0 0 0 0 1);
BETANAME='CARRY12' 'CARRY13' 'CARRY23';
TESTGLH;
NOTE PAGE TEST PERIOD EFFECTS;
H=(-1 0 1 0 0 0 0 0 0 / -1 0 0 0 0 1 0 0 0);
BETANAME='P2_P1' 'P3_P1';
TESTGLH;
NOTE PAGE TEST AVERAGE TREATMENT EFFECTS;
H=(0 1 0 0 1 0 0 0 1)/3;
BETANAME='AVGTREAT';
    TESTGLH;
NOTE PAGE TEST AVERAGE FIRST ORDER RESIDUAL EFFECTS;
H=(0 0 0 0 0.5 0 0 0 0.5 0);
BETANAME='FRSTRESD';
    TESTGLH;
NOTE PAGE TEST SECOND ORDER RESIDUAL EFFECTS;
H=(0 0 0 0 0 0 1 0 0 0);
BETANAME='2NDRSID';
    TESTGLH;
NOTE PAGE TEST FIRST PERIOD TREATMENT EFFECTS;
H=(0 1 0 0 0 0 0 0 0);
BETANAME='TREAT1';
    TESTGLH;
NOTE PAGE TEST SECOND PERIOD TREATMENT EFFECTS;
H=(0 0 0 0 1 0 0 0 0);
BETANAME='TREAT2';
    TESTGLH;
NOTE PAGE TEST THIRD PERIOD TREATMENT EFFECTS;
H=(0 0 0 0 0 0 0 0 1);
BETANAME='TREAT3';
    TESTGLH;
NOTE PAGE TEST 2-ND PERIOD FIRST ORDER RESIDUAL EFFECTS;
H=(0 0 0 1 0 0 0 0 0);
BETANAME='CARRY12';
    TESTGLH;
NOTE PAGE TEST THIRD PERIOD FIRST ORDER RESIDUAL EFFECTS;
H=(0 0 0 0 0 0 1 0 0);
BETANAME='CARRY23';
    TESTGLH;
BIBLIOGRAPHY


Kershner, R. P. (1980b) Some results on estimation and hypothesis testing in repeated measurement designs when residual effects are present. Unpublished paper.


Patterson, H. D. and Lucas, H. L. (1962) Change-over designs. NC Agriculture Experimental Station Bulletin No. 147, 1-52.


