*This research was supported in part by the National Science Foundation under Grant GP 42325.*

DESIGN AND ANALYSIS OF SERIAL EXPERIMENTS*

Lynn Weidman

Department of Statistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 997

April, 1975
Several cases of the following situation are considered. Each experimental unit receives a prescribed series of treatment applications, and a particular characteristic of interest is measured in association with each application. The experimenter is not only interested in the effect a treatment has on the observation associated with its application, called direct effect, but also in the effects it has on succeeding observations, called residual effects, and its combined effect with following applications on their associated observations, called interaction effects. Comparing the treatment effects to select the most effective overall treatment is the purpose of these experiments.

The design problem is to select the series in such a manner that variances for all the effects in a particular class of effect, e.g., direct effects, are as equal as possible and covariances between any two members of different classes are the same.

The following cases are considered:

i) Each application has a direct, two residual, and two interaction effects. Every combination of three consecutive treatments, where no treatment is repeated, occurs equally often.

ii) No restrictions on treatment order and any specified number, say n-1, of residual or interaction effects may be associated with all applications. Every combination of n consecutive treatments occurs equally often, and designs are called serial arrays. Simplified analysis results when interactions are not present. Method of analysis using orthogonal
polynomials is available when the treatments are different levels of a single factor.

iii) Several series of dependent random variables are given, where the observations for each series follow a set of conditional density functions. Experimenter wants to estimate and compare parameters of the densities for the different series. If the conditional densities satisfy certain conditions the following usual maximum likelihood results follow: (a) parameter estimates are consistent and asymptotically normal; (b) likelihood ratio tests are asymptotically \( \chi^2 \). Designs possibly useful for data collection in this case are constructed.

iv) Consider \( 2^\omega \) factorials, \( \omega = 2,3,4,5,6 \), where the factor combinations are applied in series in order to estimate the direct and one residual effect for the two levels of each factor. Serial factorial designs are derived for the given values of \( \omega \), where \( 2^\omega \) experimental units each receive 4 different combinations. Construction consists of defining fractions and their order of application for each unit.

In cases (i), (ii), and (iv) the response model is linear and analysis is by least squares.
# TABLE OF CONTENTS

Acknowledgements---------------------------------------------------------iv  
Notation---------------------------------------------------------------v  

I. Introduction and Summary--------------------------------------------1  

II. Triple Designs and Their Analysis  
2.1 Properties and Construction----------------------------------------5  
2.2 Response Model and Estimation of Parameters------------------------9  
2.3 Variances and Covariances of the Estimates-------------------------17  
2.4 Tests of Hypotheses-----------------------------------------------23  

III. Serial Arrays and Their Analysis  
3.1 Description and Construction----------------------------------------26  
3.2 General Analysis---------------------------------------------------30  
3.3 Analysis and Hypothesis Testing When Interactions 
are Excluded------------------------------------------------------------38  
3.4 Blocking of Units--------------------------------------------------41  
3.5 Levels of a Single Factor-------------------------------------------43  

IV. Maximum Likelihood for Series of Dependent Random Variables  
4.1 The Problem of Comparing Dependent Observations-------------------50  
4.2 Preliminary Results------------------------------------------------51  
4.3 Notation and Assumptions-------------------------------------------53  
4.4 Properties of Maximum Likelihood Estimates and 
Likelihood Ratio Tests------------------------------------------------57  
4.5 Summary------------------------------------------------------------66  

V. Construction of Some 2- and 3-Repeating Designs  
5.1 n-Repeating Designs-----------------------------------------------68  
5.2 Construction of 2-Repeating Designs--------------------------------71  
5.3 Construction of 3-Repeating Designs--------------------------------78  

VI. Serial Factorial Designs  
6.1 The Serial Factorial Model and Notation-----------------------------82  
6.2 $2^3$ Designs-------------------------------------------------------87
6.3 $2^4$ Designs------------------------------------------89
6.4 $2^5$ Designs------------------------------------------91
6.5 $2^6$ Designs------------------------------------------96
6.6 General Analysis of $2^\omega$ Designs-----------------98
6.7 Simplified Analysis for $\omega$ Even-------------------105

Bibliography-----------------------------------------------109
ACKNOWLEDGEMENTS

To the members of the faculty of the Department of Statistics I wish to express my gratitude for their contributions to my graduate training. I especially acknowledge the suggestions and encouragement given to me by my advisor, Professor I.M. Chakravarti, during the course of this investigation.

I wish to thank Professor N.L. Johnson for his careful reading of the manuscript and the other members of my doctoral committee, Professors D.G. Kelly, K.J.C. Smith, and T. Brylawski, for their review of my work. Additional thanks go to Dr. H.D. Tolley for suggesting the topic of Chapter 4.

I am indebted to the National Aeronautics and Space Administration and the Department of Statistics for their financial support.

For her excellent typing of the manuscript I thank Ms. June Maxwell.

I would like to extend special appreciation to my parents who continually encouraged my work.

Finally, I thank my fellow graduate students and other friends on this campus who have helped to make my non-statistical hours pleasant and entertaining.
\textbf{NOTATION}

$\mathbf{a}$ is a column vector, and $\mathbf{a}'$ is the corresponding row vector.

$A = (a_{ij})$ is a matrix with $a_{ij}$ as the element in the $i$-th row and $j$-th column.

$A'$ is the transpose of $A$.

$A^{-1}$ is any generalized inverse of $A$ and is defined by $AA^{-1}A = A$.

$A \otimes B$ is the Kronecker product of the matrices $A$ and $B$, where the $(i,j)$-th partition of $A \otimes B$ is defined as $a_{ij}B$.

$I_n$ is the identity matrix of order $n$.

$J_n$ is the $(n \times n)$ matrix with all entries one.

$O_{m,n}$ is the $(m \times n)$ matrix with all entries zero.

$0_n$ is the $(n \times 1)$ vector of zeros.

$1_n$ is the $(n \times 1)$ vector of ones.

For $\mathbf{x} = (x_1, \ldots, x_m)'$ and $\mathbf{y} = (y_1, \ldots, y_n)'$ random vectors,

$E(\mathbf{x})$ is the $(m \times 1)$ vector with $i$-th element $E(x_i)$;

$\text{Cov}(\mathbf{x}, \mathbf{y})$ is the $(m \times n)$ matrix with $\text{Cov}(x_i, y_j)$ as the $(i,j)$-th element;

$\text{Var}(\mathbf{x})$ is the $(m \times m)$ matrix $\text{Cov}(\mathbf{x}, \mathbf{x})$.

$x \sim N(\mu, \sigma^2)$ means the random variable $x$ has the univariate normal distribution with mean $\mu$ and variance $\sigma^2$.

$x \sim N_r(\mu, \Sigma)$ means that the random vector $x$ has the $r$-variate multivariate normal distribution with mean vector $\mu$ and variance matrix $\Sigma$.

$x \sim \chi^2_r$ means that the random variable $x$ has the central chi-square distribution with $r$ degrees of freedom.

$x \sim t_r$ means that the random variable $x$ has the central Student's $t$ distribution with $r$ degrees of freedom.
$x \sim F_{r,s}$ means that the random variable $x$ has the F distribution with $r$ and $s$ degrees of freedom.

$L_{m \to} x$ means that the random vector $x_m$ converges in law to the random vector $x$.

$L_{m \to} G$ means that the random vector $x_m$ converges in law to the random vector $x$ which has distribution function $G$.

$p \lim_{m \to} x = c$ means that $x_m$ converges in probability to the constant $c$.

$\lim_{m \to} x_a = c$ means that $x_m$ converges almost surely to the constant $c$.

$p \lim_{m \to} X = C$ means that each element of the matrix $X_m$ converges in probability to the corresponding element of $C$.

\[\square\] denotes the conclusion of a proof or example.
CHAPTER I
INTRODUCTION AND SUMMARY

Comparison of the effectiveness of several treatments under certain assumptions commonly referred to as the general linear model has supplied much of the impetus for developing the subject area of experimental design. Among the situations which have been considered is that in which each experimental unit receives a prescribed series of treatment applications, and a particular characteristic of interest is measured in association with each application. In this case the experimenter is not only interested in the effect a treatment has on the observation associated with its application, but also in the effects it has on succeeding observations.

Definition 1.1: The effect a treatment has on the observation associated with its application is called its direct effect. For the first observation following that one its influence is the first residual effect, for the next observation it is the second residual effect, etc. If treatment h applied at time i combines with treatment k applied at time i+j to influence the observation associated with the application at time i+j, this is called the j-th interaction effect of treatments h and k.

Designs used in this situation to estimate direct, residual, and interaction effects are commonly called change-over designs. For a summary of designs of this type see [20] and [22].
Some constructions for change-over designs which will be important here were given by Williams [28],[29]. He assumed that each unit was to receive each treatment exactly once and that the designs were to be balanced either for a single residual effect or for pairs of residual effects. Balance means that all estimates in the same class (e.g., first residuals, direct effects) have equal variances and that covariances between the members of two classes are all equal. We use the following definition in describing these designs.

Definition 1.2: In a design with a series of treatments applied to each unit, each \( n \) consecutive applications in which no treatment can occur more than once is called an \( n \)-plet. If a treatment can be repeated in any \( n \) consecutive applications they are called \( n \)-tuples. If \( v \) treatments are used, there are \( v(v-1)\ldots(v-n+1) \) possible \( n \)-plets and \( v^n \) possible \( n \)-tuples.

The designs for a single residual effect given in [28] consist of either one or two latin squares, where the number of times each 2-plet occurs is equal to the number of latin squares. Designs balanced for pairs of residual effects as given in [28] and [29] consist of a set of latin squares. In such a set each 3-plet occurs once and each 2-plet occurs at the beginning of some row and at the end of some other row exactly once. Each row of a latin square is the series of treatments for a single experimental unit in these designs. Finney and Outhwaite [12], [13] considered applying a long series of treatments to a single unit in order to estimate the same effects. Further designs along this line were given by Sampford [26].

A combination of these two approaches was used by Federer and Atkinson [11] under the title tied-double-change-over designs. For these
designs each unit is assigned several series of treatments which are applied successively and are balanced for the estimation of a single residual effect. Chapter 2 presents designs and analysis for what we call triple designs, which assume the presence of first and second residual effects. The method of construction involves combining in a specified manner the Latin squares used as designs balanced for pairs of residual effects by Williams.

In general, let \( n \) represent the maximum number of observations on which any given treatment application has an effect. We now drop the restriction that no treatment may occur more than once in any \( n \) consecutive applications, and in Chapter 3 we introduce serial arrays for estimating in this case all treatment effects in a balanced manner. Thus we will consider \( n \)-tuples rather than \( n \)-plets as we did in Chapter 2. The primary property of serial arrays is that each \( n \)-tuple occurs equally often. This along with the other properties of serial arrays allows us to derive general formulas for estimating all the effects being considered. Serial arrays are also applicable to the case of examining \( v \) different levels of a single treatment, under assumptions that the treatment effects are polynomials. This case is discussed in Section 3.5.

When each unit in an experiment repeatedly receives a single treatment and the response pattern is assumed to fit a particular linear model, we have what is commonly called a growth curve. An early examination of this case was given by Wishart [31],[32], and extensive work has been done since that time. (See, e.g., [8],[10],[16],[17],[23],[24],[25].) In Chapters 4 and 5 we allow a more complicated response pattern in which each response is dependent on previous responses in addition to
the direct, residual, and interaction effects of the treatments. Chapter 4 considers the problem of estimating and comparing parameters included in several independent series of dependent random variables. The case of primary interest to us is when each series of dependent variables is the set of observations for a single unit, and the series for each unit is independent of all others. For each series the conditional probability density functions must satisfy the conditions specified by C1 through C5. This work is an extension of results given by Bhat [4] and Billingsley [6]. Chapter 5 presents some designs which may be useful in establishing the verification of the conditions on the conditional density functions.

In the final chapter we consider using series of treatment applications in the factorial case. Factorial designs consist of treatments which are combinations of several factors, and in fractional factorials a particular subset of the total set of possible combinations is used, each combination being applied to a different unit. Main effects of and interactions between factors are of interest in the designs. If we apply several of the combinations in a fraction, in a given manner, to each unit available, we will perhaps be able to estimate the direct and residual effects of the individual factors. Such information would be useful in deciding which factor combination should be used over some future extended period of application. This general situation was discussed by Patterson [21] under the title of serial factorial designs. For the special case of $2^\omega$ factorials, $\omega = 2,3,4,5,6$, some designs on fewest possible units, and their analysis, are derived in Chapter 6.
CHAPTER II
TRIPLE DESIGNS AND THEIR ANALYSIS

2.1 Properties and Construction

In 1955 Federer and Atkinson [11] presented what they called tied-double-change-over designs which are used to compare effects of \( v \) different treatments.

**Definition 2.1:** Let \( v \) and \( s \) be integers and let \( I \) be the number of experimental units being treated. A *tied-double-change-over design* has the following properties:

(i) each experimental unit receives a series of \( sv \) treatment applications;

(ii) no treatment occurs in two consecutive applications;

(iii) all 2-plets occur equally often in the whole design; and

(iv) at each time of application every treatment is applied to the same number of units, i.e., \( I/v \) is an integer.

Since there are \( v(v-1) \) 2-plets, the total number of applications in such a design is \( dv(v-1) + I \), where \( d \) is some integer. Each observation is affected by one direct and one residual treatment effect and, possibly, by a single interaction effect. Variances for all estimates in each group of effects are equal. Also, all covariances between estimates from two given groups are equal.

In this chapter we extend these assumptions to include second residual effects and interactions for each 3-plet of consecutive treatments. Analogous to the above case we require that all of the \( v(v-1)(v-2) \)
3-plets occur equally often in consecutive applications. Each unit will be assigned a series of the \( v \) treatments which is to be repeated \( s \) times. The last two treatments of this series are to be applied to the unit in the first two application periods. No observations are made until after the third application, so these first two may be called "conditioning" and are used only in estimating residual and interaction effects.

Before proceeding we introduce some terminology.

**Definition 2.2:** Let \( V = \{ \tau_1, \tau_2, \ldots, \tau_v \} \) be a set of \( v \) symbols. A **rectangle** is a matrix with \( r \) rows and \( c \) columns whose entries are members of \( V \). An \( r \times r \) rectangle is a **square**. If each member of \( V \) occurs exactly once in each row and column of a square it is called a **latin square**.

The designs examined in this chapter are constructed from Williams' [29] designs balanced for pairs of residual effects. Williams' designs for \( v \) treatments consist of sets of \( v-1 \) latin squares of size \( v \), as described in Chapter 1. Each 3-plet occurs exactly once and each 2-plet occurs once at the beginning of some row and once at the end of some other row in those designs.

**Definition 2.3:** Suppose we want to compare \( v \) treatments, and \( r \) and \( s \) are integers. A triple design has the following properties:

(i') there are \( r(v-1)(v-2) \) experimental units;

(ii') \( sv \) is the number of observations made on each unit;

(iii') observations are made on all the units in the same time periods;

(iv') each treatment is applied \( rs(v-1)(v-2) \) times;

(v') all 3-plets occur \( rs \) times in the whole design;

(vi') the \( v \)-th treatment occurs \( r(v-1) \) times at application period \( j \), \( j=nt-1 \) or \( nt \), \( n=1,2,\ldots,s \), and 0 times at periods \( nt-1 \) and \( nt \); and
(vii') all other treatments occur \( r(v-3) \) times at period \( j \),
\( j \neq nt-1 \) or \( nt \), and \( r(v-2) \) times at periods \( nt-1 \) and \( nt \).

We give a procedure for constructing triple designs for any integral
\( (r,s,t) \), where the set of treatments is \( V = \{0,1,\ldots,v-1\} \).

1. Take one of Williams' designs described above for \( v-1 \) treat-
ments, where the first row of each latin square ends with a
different number. Write down these squares one below the
other.

2. Place a \( v-1 \) at the beginning of each of the \( (v-1)(v-2) \) rows.
We now have \( v-2 \ (v-1) \times v \) rectangles.

3. In the \( i \)-th rectangle, \( i=2,3,\ldots,v-2 \), move column \( j \),
\( j=1,2,\ldots,v-i+1 \), to column \( (j+i-1) \) and the last \( i-1 \) columns
to the first \( i-1 \), retaining their order. This cyclically
rearranges the columns so that treatment \( v-1 \) now occurs in
column \( i \) in rectangle \( i \), \( i=1,2,3,\ldots,v-2 \), and never occurs in
columns \( v-1 \) and \( v \).

4. Repeat each row after itself \( s-1 \) additional times and place
the two treatments at the end of each row also at its begin-
ning. These are the conditioning treatments. Repeat the
entire resulting rectangle \( r-1 \) additional times.

5. Let each row of this \( r(v-1)(v-2) \times (sv+2) \) rectangle represent
the treatment application series for a single unit and each
column represent an application period.

All of the properties of a triple design, except perhaps \( (v') \), are
clearly included in these designs. We show that \( (v') \) is also satisfied.

**Lemma 2.1:** All 3-plets occur exactly \( rs \) times for the rectangles gener-
ated as above.

**Proof:** We need consider only the upper left quadrant of size \( (v-1)(v-2) \)
\( \times (2+v) \) from our rectangle. If each 3-plet is included once then the whole
rectangle will include each \( rs \) times because of the method of repeating
rows and rectangles in its construction.

From the properties of the original design for $v-1$ treatments each 3-plet without $v-1$ occurs exactly once. We show that every 3-plet including $v-1$ occurs exactly once.

(a) Since every 2-plet without $v-1$ occurs once at the end of a row in the original design, every 3-plet of the form $a \ b \ v-1$ occurs exactly once in the new design. Similarly, every 3-plet of the form $v-1 \ a \ b$ occurs once.

(b) Each initial row of a square in the original design begins with 0 and ends with one of the other treatments $1, 2, \ldots, v-2$, such that each different possible last element occurs once. Our construction method has placed $v-1$ between the first and last elements of the first row of each square so that all 3-plets of the form $0 \ v-1 \ a$, $a=1, 2, \ldots, v-2$, occur exactly once. A similar result holds for all the second rows, third rows, etc. In this way every 3-plet with $v-1$ as the middle element occurs exactly once.

Thus every 3-plet occurs exactly once in the quadrant of interest. \hfill \Box

**Example 2.1:** We demonstrate the construction of a triple design for $r=1, s=2, v=4$. The original design balanced for pairs of residual effects consists of two latin squares of size $v-1 = 3$.

\[
\begin{array}{ccc}
0 & 1 & 2 \\
1 & 2 & 0 \\
2 & 0 & 1 \\
\end{array} \hspace{1cm} \begin{array}{ccc}
0 & 2 & 1 \\
1 & 0 & 2 \\
2 & 1 & 0 \\
\end{array} \]  

(2.1.1)

The derived triple design is then
where (a) represents the conditioning periods, (b) is the design derived directly from the squares (2.1.1), and (c) is a repetition of the rows of (b).

Note that in (b) the treatment v-1 = 3 occurs in the first column for the first 3 units and in the second column for the last 3.

2.2 Response Model and Estimation of Parameters

As mentioned above, these triple designs enable us to estimate all the direct, first residual, and second residual effects for all v treatments, as well as interactions between treatments occurring within any 3 consecutive periods. Because of the special position given to treatment v-1 in the construction of these designs, the variances for and covariances between classes of estimates will not be the same for all treatments. However, they will be the same for all treatments except v-1. This will allow us to derive simplified equations for estimating the treatment effects of all the other treatments. Also, the experimenter should decide how best to handle this special relationship of treatment v-1 before using a triple design. One appropriate method is to use a control or standard treatment as v-1 and randomly assign the others to the remaining numbers.

For the model presented here we consider any interactions to be zero, so that our interest is centered on deriving estimates of the
direct and residual treatment effects. Assuming a linear response model we have

\[ Y_{ijh} = N_{ijh} \left( \mu + \gamma_i + \beta_j + \delta_h + \sum_{p=0}^{v-1} N_{i(j-1)p} \rho_p + \sum_{q=0}^{v-1} N_{i(j-2)q} \eta_q + \epsilon_{ijh} \right), \]  

(2.2.1)

\( i=1,2,\ldots, r(v-1)(v-2), \ j=1,2,\ldots, sv, \ h=0,1,\ldots,v-1, \) where

\( Y_{ijh} \) is the observation made following the application of treatment \( h \) to unit \( i \) at time \( j \),

\[ N_{ijh} = \begin{cases} 
1 & \text{if treatment } h \text{ was applied to unit } i \text{ at time } j \\
0 & \text{otherwise} \end{cases} \]

\( \mu \) is the overall mean,

\( \gamma_i \) is the effect of unit \( i \),

\( \beta_j \) is the effect of time period \( j \),

\( \delta_h, \rho_h, \eta_h \) are respectively the direct, first residual, and second residual effects of treatment \( h \), and

\( \epsilon_{ijh} \) is a random error term for observation \( Y_{ijh} \).

Note that for each \( i,j \) only one \( N_{ijh} \) is non-zero. Thus the number of observations we have is \( r(v-1)(v-2) \times v \). We consider only these responses in our analysis, ignoring all \( Y_{ijh} \) which are identically zero.

We make the additional assumptions that the \( \epsilon_{ijh} \) are independent and identically distributed as \( N(0, \sigma^2) \) and

\[ \sum_{i=1}^{r(v-1)(v-2)} \gamma_i = \sum_{j=1}^{sv} \beta_j = \sum_{h=0}^{v-1} \delta_h = \sum_{p=0}^{v-1} \rho_p = \sum_{q=0}^{v-1} \eta_q = 0. \]

(2.2.2)

The number of independent effects to be estimated is then \( r(v-1)(v-2) + v(s+3) - 4. \)
Due to the distributional assumptions on the $e_{ijh}$ we may estimate the effects by the method of maximum likelihood by calculating the estimates $\hat{y}_{ijh}$ which minimize the sum of squares

$$G = \frac{r(v-1)(v-2) sv}{i=1} \sum_{j} \sum_{h=0}^{v-1} (y_{ijh} - \hat{y}_{ijh})^2$$

(2.2.3)

under the restrictions (2.2.2). We do this in the usual manner by taking partial derivatives with respect to the effects and setting them equal to 0.

(a) \[-\frac{1}{2} \frac{\partial G}{\partial \mu} = \sum_{i} \sum_{j} \sum_{h} N_{ijh} y_{ijh} - rsv(v-1)(v-2)\mu - sv \sum_{i} Y_{i} - r(v-1)(v-2)\sum_{j} b_{j} \]

\[-rsv(v-1)(v-2)\sum_{h} (\delta_{h} + \rho_{h} + \eta_{h}) = 0. \quad (2.2.4a)\]

(The summations range over all possible values of the given subscript unless otherwise indicated.) Using (2.2.2) this reduces to

$$rsv(v-1)(v-2)\mu = y_{...} \quad (2.2.4b)$$

where $y_{...}$ is the sum of all observations.

(b) \[-\frac{1}{2} \frac{\partial G}{\partial y_{i}} = \sum_{j} \sum_{h} N_{ijh} y_{ijh} - sv\mu - svY_{i} - \sum_{j} b_{j} - s \sum_{h} (\delta_{h} + \rho_{h} + \eta_{h}) = 0 \quad (2.2.5a)\]

which reduces to

$$sv(\mu + \gamma_{i}) = y_{i...} \quad (2.2.5b)$$

where $y_{i...}$ is the sum of observations for unit $i$.

Before taking derivatives with respect to the $b_{j}$ we note that the effects for treatment v-1 do not occur in all periods. (Recall the properties of these triple designs.) In fact, if $n=1,2,\ldots,s$, then

(i) $\delta_{v-1}$ does not occur with $b_{j}$, $j=nv-1$ and $nv$, (ii) $\rho_{v-1}$ does not occur with $b_{j}$, $j=nv$, $nv+1$, and 1, and (iii) $\eta_{v-1}$ does not occur with $b_{j}$, $j=nv+1$, $nv+2$, 1, and 2.
(c) \[-\frac{1}{2} \frac{\partial G}{\partial \delta_h}\]:

If \(j = nv-1, n=1,2,\ldots,s\), then we have

\[
\sum_{i} \sum_{h} N_{ijh} Y_{ijh} - s(v-1)(v-2)\mu - \sum_{i} Y_{i} - r(v-2) \sum_{h \neq v-1} \delta_{h} - r(v-1) (\rho_{v-1} + \eta_{v-1}) - r(v-3) \sum_{p \neq v-1} (\rho_{p} + \eta_{p}) - s(v-1)(v-2)\beta_{j} = 0 .
\]

(2.2.6a)

Using (2.2.2) this becomes

\[
s(v-1)(v-2)(\mu + \beta_{j}) - r(v-2)\delta_{v-1} + 2r(\rho_{v-1} + \eta_{v-1}) = y_{v,j} .
\]

(2.2.6b)

where \(y_{v,j}\) is the sum of observations for period \(j\). Similarly we have 4 other equations for different values of \(j\).

\[
s(v-1)(v-2)(\mu + \beta_{j}) - r(v-2)(\delta_{v-1} + \rho_{v-1}) + 2r\eta_{v-1} = y_{v,j}\]

\(j = nv\)

(2.2.6c)

\[
s(v-1)(v-2)(\mu + \beta_{j}) + 2r\delta_{v-1} - r(v-2)(\rho_{v-1} + \eta_{v-1}) = y_{v,j}\]

\(j = nv + 1\)

(2.2.6d)

\[
s(v-1)(v-2)(\mu + \beta_{j}) + 2r(\delta_{v-1} + \rho_{v-1}) - r(v-2)\eta_{v-1} = y_{v,j}\]

\(j = nv + 2\)

(2.2.6e)

\[
s(v-1)(v-2)(\mu + \beta_{j}) + 2r(\delta_{v-1} + \rho_{v-1} + \eta_{v-1}) = y_{v,j}\]

all other \(j\).

(2.2.6f)

(d) \[-\frac{1}{2} \frac{\partial G}{\partial \delta_h}\]:

For \(h \neq v-1\) our equation is

\[
\sum_{i} \sum_{j} N_{ijh} Y_{ijh} - rs(v-1)(v-2)(\mu + \delta_{h}) - s \sum_{i} Y_{i} - r(v-3) \sum_{j \neq nv-1, } \beta_{j} - r(v-2) \sum_{j = nv-1, } \beta_{j} - rs \sum_{p \neq h} N_{ijh} N_{i(j-1)p} + \rho_{p} - rs \sum_{q \neq h} N_{ijh} N_{i(j-2)q} = 0 .
\]

(2.2.7a)
Using (2.2.2) and the property that any two treatment pairs occur in
designated positions in the same 3-tuple rs(v-2) times, we can re-
write this equation as

$$\text{rs}(v-1)(v-2)(\mu + \delta) + r \sum_{j=\text{nv}, 1}^{\text{nv}} \beta_j - \text{rs}(v-2)(\rho_1 + \eta_1) = y_{v=1}, \quad (2.2.7b)$$

where $y_{v=1}$ is the sum of all observations where the direct effect of
treatment $h$ occurs.

For the direct effect of treatment $v-1$ we have a slightly different

equation:

$$\sum_{i,j} N_{ij} y_{ij} - \text{rs}(v-1)(v-2)(\mu + \delta) - s \sum_{i} y_i - r(v-1) \sum_{j \neq \text{nv}, 1}^{\text{nv}} \beta_j - \sum_{p=v-1}^{\text{nv}} N_{ij} y_{ij} \sum_{j \neq \text{nv}, 1}^{\text{nv}} N_{ij} \sum_{q \neq v-1}^{\text{nv}} N_{ij} (\mu + \delta) = 0. \quad (2.2.7c)$$

Again we can reduce this equation to a simpler form.

$$\text{rs}(v-1)(v-2)(\mu + \delta) - r(v-1) \sum_{j=\text{nv}, 1}^{\text{nv}} \beta_j - \text{rs}(v-2)(\rho_1 + \eta_1)$$

$$= y_{v=1}. \quad (2.2.7d)$$

The equations for residual effects are derived in the same manner.

(e) $- \frac{\partial G}{\partial \rho_p}$:

$$\text{rs}(v-1)(v-2)(\rho_1 + \mu) + r \sum_{j=\text{nv}, 1}^{\text{nv+1, 1}} \beta_j - s(v-2)(\delta + \eta_p) = y_{v=1}, \quad (2.2.8a)$$

$$\text{rs}(v-1)(v-2)(\rho_1 + \mu) - r(v-1) \sum_{j=\text{nv}, 1}^{\text{nv+1, 1}} \beta_j - s(v-2)(\delta + \eta_1) = y_{1, v=1}, \quad (2.2.8b)$$

where $y_{v=1}$ is the sum of observations where the first residual effect
of treatment $p$ occurs.
(f) \[ \frac{1}{2} \sum_{q} \frac{\partial G}{\partial \eta_q} : \]

\[ rs(v-1)(v-2)(\eta_q + \mu) + r \sum_{j=\nu v+1} \beta_j - s(v-2)(\delta_q \rho_q) = \gamma_\circ(2)q, \quad (2.2.9a) \]

\[ q \neq v-1, \]

\[ rs(v-1)(v-2)(\eta_{v-1} + \mu) - r(v-1) \sum_{j=\nu v+1} \beta_j - s(v-2)(\delta_{v-1} \rho_{v-1}) \]

\[ = \gamma_\circ(2)v-1, \quad q = v-1, \quad (2.2.9b) \]

where \( \gamma_\circ(2)q \) is the sum of observations where the second residual treatment of \( q \) occurs.

The system of normal equations for model (2.2.1) is given by (2.2.4b), (2.2.5b), (2.2.6b-f), (2.2.7b), (2.2.7d), (2.2.8a-b), and (2.2.9a-b). Direct solution of (2.2.4b), (2.2.5b), and (2.2.6b-f) gives the following set of equations, where the effects of all treatments except \( v-1 \) have been eliminated.

\[ \hat{\mu} = \bar{y}_{..} \]

\[ \gamma_i = \bar{y}_{i..} - \bar{y}_{..} \quad (2.2.10a) \]

\[ \hat{\beta}_j = (\bar{y}_{.j} - \bar{y}_{..}) + \frac{v-2}{sv}\delta_{v-1} - \frac{2}{sv}(\rho_{v-1} + \eta_{v-1}), \quad j=nt-1 \]

\[ \hat{\beta}_j = (\bar{y}_{.j} - \bar{y}_{..}) + \frac{v-2}{sv}(\delta_{v-1} \rho_{v-1}) - \frac{2}{sv} \eta_{v-1}, \quad j=nt \quad (2.2.10c) \]

\[ \hat{\beta}_j = (\bar{y}_{.j} - \bar{y}_{..}) - \frac{2}{sv} \delta_{v-1} + \frac{v-2}{sv} (\rho_{v-1} \eta_{v-1}), \quad j=nt+1,1 \quad (2.2.10d) \]

\[ \hat{\beta}_j = (\bar{y}_{.j} - \bar{y}_{..}) - \frac{2}{sv} (\delta_{v-1} \rho_{v-1}) + \frac{v-2}{sv} \eta_{v-1}, \quad j=nt+2,2 \quad (2.2.10e) \]

\[ \hat{\beta}_j = (\bar{y}_{.j} - \bar{y}_{..}) - \frac{2}{sv} (\delta_{v-1} \rho_{v-1} \eta_{v-1}), \text{ all other } j. \quad (2.2.10f) \]

A \( \bar{y} \) indicates the mean of the given \( y \) and a \( \hat{\quad} \) over a parameter indicates that it is the maximum likelihood estimate.

Substitution of these results into equations (2.2.7d), (2.2.8b) and (2.2.9b) gives a set of 3 equations for estimating the effects of treatment \( v-1 \).
\[
\frac{sv-2}{sv} \delta_{v-1} + \frac{(v-4)(v-1) - sv(v-2)}{sv(v-2)(v-1)} \rho_{v-1} + \frac{4(v-1) - sv(v-2)}{sv(v-2)(v-1)} \eta_{v-1} \\
= (\overline{y}_{v-1} - \overline{y}_{\ldots \ldots}) + \frac{1}{s(v-2)} \sum_{j=nv-1, n} (\overline{y}_{j} - \overline{y}_{\ldots \ldots}) \quad (2.2.11a)
\]

\[
\frac{(v-1)(v-4) - sv(v-2)}{sv(v-2)(v-1)} \delta_{v-1} + \frac{sv-2}{sv} \rho_{v-1} + \frac{(v-1)(v-4) - sv(v-2)}{sv(v-2)(v-1)} \eta_{v-1} \\
= (\overline{y}_{(1)v-1} - \overline{y}_{\ldots \ldots}) + \frac{1}{s(v-2)} \sum_{j=nv, nv+1, 1} (\overline{y}_{j} - \overline{y}_{\ldots \ldots}) \quad (2.2.11b)
\]

\[
\frac{4(v-1) - sv(v-2)}{sv(v-2)(v-1)} \delta_{v-1} + \frac{(v-1)(v-4) - sv(v-2)}{sv(v-2)(v-1)} \rho_{v-1} + \frac{sv-2}{sv} \eta_{v-1} \\
= (\overline{y}_{(2)v-1} - \overline{y}_{\ldots \ldots}) + \frac{1}{s(v-2)} \sum_{j=1, 2, nv+1, nv+2} (\overline{y}_{j} - \overline{y}_{\ldots \ldots}) \quad (2.2.11c)
\]

Denote the 3 coefficients on the left hand side (LHS) of (2.2.11a) by \(c_{v1}, c_{v2}, c_{v3}\), respectively, and denote the RHS of the 3 equations by \(y_{0v}, y_{1v}, \) and \(y_{2v}\), respectively. Then the estimated effects for treatment \(v-1\) are given in matrix form by

\[
\begin{pmatrix}
\hat{\delta}_{v-1} \\
\hat{\rho}_{v-1} \\
\hat{\eta}_{v-1}
\end{pmatrix} =
\begin{pmatrix}
c_{v1} & c_{v2} & c_{v3}
c_{v2} & c_{v1} & c_{v2}
c_{v3} & c_{v2} & c_{v1}
\end{pmatrix}
\begin{pmatrix}
y_{0v} \\
y_{1v} \\
y_{2v}
\end{pmatrix} \quad (2.2.12a)
\]

or, with the obvious notational changes,

\[
\hat{\delta}_{v-1} = C_{v}^{\top} X_{v} \quad (2.2.12b)
\]

Once we have obtained \(\hat{\delta}_{v-1}\) we can substitute its entries into (2.2.10a-g) to give us the \(\hat{\beta}_{j}\)’s. More importantly, we can now substitute into (2.2.7b), (2.2.8a), and (2.2.9a) to derive the estimates of the remaining treatment effects. These equations reduce to the
following 3.

\[
\delta_h - \frac{1}{v-1} (\rho_h + \eta_h) + \frac{2}{sv(v-1)} \hat{\delta}_{v-1} + \frac{(v-4)}{sv(v-1)(v-2)} \hat{\rho}_{v-1} - \frac{4}{sv(v-1)(v-2)} \hat{\eta}_{v-1} \\
= (\bar{y}_{*h} - \bar{y}_{**}) - \frac{1}{s(v-2)(v-1)} \sum_{j=nt, \text{even}}^{nv-1} (\bar{y}_{*j} - \bar{y}_{**}), \ h \neq v-1
\] (2.2.13a)

\[
- \frac{1}{v-1} (\delta_p + \eta_p) + \rho_p + \frac{(v-4)}{sv(v-1)(v-2)} \hat{\delta}_{v-1} + \frac{2}{sv(v-1)} \hat{\rho}_{v-1} + \frac{(v-4)}{sv(v-1)(v-2)} \hat{\eta}_{v-1} \\
= (\bar{y}_{(1)h} - \bar{y}_{**}) - \frac{1}{s(v-2)(v-1)} \sum_{j=nt+1}^{nv-1} (\bar{y}_{*j} - \bar{y}_{**}), \ p \neq v-1
\] (2.2.13b)

\[
- \frac{1}{v-1} (\delta_q + \eta_q) - \frac{4}{sv(v-1)(v-2)} \hat{\delta}_{v-1} + \frac{(v-4)}{sv(v-1)(v-2)} \hat{\rho}_{v-1} + \frac{2}{sv(v-1)} \hat{\eta}_{v-1} \\
= (\bar{y}_{(2)q} - \bar{y}_{**}) - \frac{1}{s(v-2)(v-1)} \sum_{j=nt+1}^{nv+2} (\bar{y}_{*j} - \bar{y}_{**}), \ q \neq v-1
\] (2.2.13c)

In (2.2.13a) denote the coefficients of \( \hat{\delta}_{v-1}, \hat{\rho}_{v-1}, \) and \( \hat{\eta}_{v-1} \) by \( c_{h1}, c_{h2}, \) and \( c_{h3}, \) respectively. Let \( y_{0h}, y_{1h}, \) and \( y_{2h} \) denote the RHS of (2.2.13a-c), respectively. Then the estimates are given by

\[
\begin{pmatrix}
\hat{c}_h \\
\hat{\rho}_h \\
\hat{\eta}_h
\end{pmatrix} = \begin{pmatrix}
1 & -\frac{1}{v-1} & -\frac{1}{v-1} \\
-\frac{1}{v-1} & 1 & -\frac{1}{v-1} \\
-\frac{1}{v-1} & -\frac{1}{v-1} & 1
\end{pmatrix}^{-1} \begin{pmatrix}
y_{0h} \\
y_{1h} \\
y_{2h}
\end{pmatrix} - \begin{pmatrix}
c_{h1} & c_{h2} & c_{h3}
\end{pmatrix} \begin{pmatrix}
\hat{\delta}_{v-1} \\
\hat{\rho}_{v-1} \\
\hat{\eta}_{v-1}
\end{pmatrix}
\] (2.2.14a)

or, with the obvious notational changes,

\[
\hat{\delta} = T^{-1}[y_h - C_h \hat{\delta}_{v-1}],
\]

where

\[
T^{-1} = \frac{v(v-1)}{(v-1)(v-4)-2} \begin{pmatrix}
v-2 & 1 & 1 \\
1 & v-2 & 1 \\
1 & 1 & v-2
\end{pmatrix}.
\]
2.3 Variances and Covariances of the Estimates

For the purpose of making comparisons among the effects estimated in Section 2 we need to calculate the matrices \( \text{Var}(\hat{\theta}_{v-1}) \), \( \text{Var}(\hat{\theta}_h) \), \( \text{Cov}(\hat{\theta}_h, \hat{\theta}_{v-1}) \), and \( \text{Cov}(\hat{\theta}_h, \hat{\theta}_k), k \neq h \) and \( h, k \neq v-1 \). We will calculate each of the elements of these matrices individually.

First note that if \( \bar{y}_x \) is any of the means that appears as an element of \( y_h \) or \( y_v \), and \( \bar{e}_x \) is the corresponding mean of \( \epsilon_{ijh} \)'s, then \( \bar{y}_x - E(\bar{y}_x) = \bar{e}_x \), where \( E \) denotes expectation. This is a consequence of the assumptions that all of the effects in (2.2.11) are fixed and \( E(\epsilon_{ijh}) = 0 \), all \( (i,j,h) \). Now let \( a_x \) denote the number of summands in \( \bar{e}_x \) and \( a_w \) the same for \( \bar{e}_w \). Then \( E(\bar{e}_x \bar{e}_y) = \sigma^2 \cdot \frac{a_x \cdot a_w}{(\text{number of summands common to } \bar{e}_x \text{ and } \bar{e}_w)} \). We give an example of a covariance calculation using this method.

**Example 2.2:** Calculate \( \text{Cov}(y_{0v}, y_{1h}) \)

\[
\text{Cov}(y_{0v}, y_{1h}) = E[(y_{0v} - E(y_{0v}))(y_{1h} - E(y_{1h}))]
\]

\[
= E[(\bar{e}_{0v} - \bar{e}_{0o}) + \frac{1}{s(v-2)} \sum_{j=1}^{nv} (\bar{e}_{jv} - \bar{e}_{oo})] \\
\times [(\bar{e}_{(1)h} - \bar{e}_{oo}) - \frac{1}{s(v-1)(v-2)} \sum_{j=1}^{nv+1} (\bar{e}_{jv} - \bar{e}_{oo})].
\]

(For simplicity of notation in this example we will take \( \sigma^2 = 1 \).

This does not affect the calculations in any way and \( \sigma^2 \) is placed in (2.3.6).)

The number of summands in 

\[
\begin{cases}
\bar{e}_{0v} \text{ and } \bar{e}_{(1)h} \text{ is } sr(v-2)(v-1) \\
\bar{e}_{jv} \text{ is } r(v-1)(v-2) \\
\bar{e}_{oo} \text{ is } srv(v-1)(v-2)
\end{cases}
\]
Every $\xi_{ijh}$ is a summand for $\xi_{oooo}$.

Now we calculate the expectations term by term.

$$E(\xi_{v-1}^{o}-\xi_{oooo})(\xi_{v-1}^{o})h^{0} = \frac{r(v-2)}{[sr(v-1)(v-2)]^2} - \frac{2sr(v-1)(v-2)}{[sr(v-1)(v-2)]^2v}$$

$$+ \frac{sr(v-1)(v-2)}{[sr(v-1)(v-2)]^2}$$

$$= \frac{r_{v-1}^2(v-2) - rsv(v-1)(v-2)}{[sr(v-1)(v-2)]^2} \quad (2.3.2)$$

$$E(\xi_{v-1}^{o}-\xi_{oooo})\left[ -\frac{1}{3}(v-1)(v-2) \right] \sum_{j=1,nv, \atop nv+1} (\xi_{v-1}^{o,j} - \xi_{oooo}) \right]$$

$$= \frac{sr(v-1)(v-2)}{[sr(v-1)(v-2)]^2} \quad (2.3.3a)$$

depends on the different values of $j$. For $j = nv$, $\xi_{v-1}^{o,v}$ and $\xi_{v-1}^{o,j}$ have no summands in common, and for $j = nv+1,1$, they have $r(v-2)$ in common. Thus if $j = nv$,

$$E(\xi_{v-1}^{o}-\xi_{oooo})(\xi_{v-1}^{o,j} - \xi_{oooo}) = 0 - \frac{sr(v-1)(v-2)}{[sr(v-1)(v-2)]^2v} - \frac{r(v-1)(v-2)}{[r(v-1)(v-2)]^2sv}$$

$$+ \frac{sr(v-1)(v-2)}{[sr(v-1)(v-2)]^2}$$

$$= - \frac{sr(v-1)(v-2)}{[sr(v-1)(v-2)]^2} \quad (2.3.3b)$$

and if $j = 1$ or $nv+1$,

$$E(\xi_{v-1}^{o}-\xi_{oooo})(\xi_{v-1}^{o,j} - \xi_{oooo}) = \frac{r(v-1)}{[r(v-1)(v-2)]^2s} - \frac{sr(v-1)(v-2)}{[sr(v-1)(v-2)]^2} \quad ,$$

where the last term follows from $(2.3.3b)$. This expected value simplifies to

$$\frac{rsr(v-1)(v-2) - rsv(v-1)(v-2)}{[sr(v-1)(v-2)]^2} \quad . \quad (2.3.3c)$$
There are \( s \) of each of the terms (2.3.3b) and (2.3.3c) in (2.3.3a), so that this last becomes

\[
- \frac{s}{s(v-1)(v-2)} \left\{ \frac{-\text{srv}(v-1)(v-2) + \text{rv}(v-1)}{[\text{rv}(v-1)(v-2)]^2} \right\}
\]

\[
= \frac{2\text{rv}(v-2) - \text{rv}^2}{(v-2)[\text{rv}(v-1)(v-2)]^2} .
\tag{2.3.3d}
\]

Next we consider

\[
E(\overline{c}_o(1h) - \overline{c}_{o...}) \left[ \frac{1}{s(v-2)} \sum_{j=nv-1}^{nv} (\overline{c}_{o:j} - \overline{c}_{o...}) \right] .
\tag{2.3.4a}
\]

If \( j = nv-1 \), \( \overline{c}_o(1h) \) and \( \overline{c}_{o:j} \), have \( r(v-3) \) summands in common, and if \( j = nv \), they have \( r(v-2) \). Thus if \( j = nv-1 \),

\[
E(\overline{c}_o(1h) - \overline{c}_{o...})(\overline{c}_{o:j} - \overline{c}_{o...}) = \frac{\text{rv}^2(v-3) - \text{rv}(v-1)(v-2)}{[\text{rv}(v-1)(v-2)]^2} ,
\tag{2.3.4b}
\]

and if \( j = nv \),

\[
E(\overline{c}_o(1h) - \overline{c}_{o...})(\overline{c}_{o:j} - \overline{c}_{o...}) = \frac{r(v-2)}{[r(v-1)(v-2)]^2s} - \frac{\text{rv}(v-1)(v-2)}{[\text{rv}(v-2)(v-1)]^2} 
\]

\[
= \frac{\text{rv}(v-2)}{[\text{rv}(v-1)(v-2)]^2} .
\tag{2.3.4c}
\]

There are \( s \) terms each of the form (2.3.4b) and (2.3.4c) in (2.3.4a), so it becomes

\[
\frac{s}{s(v-2)} \left\{ \frac{\text{rv}^2(v-3) - \text{rv}(v-1)(v-2) + \text{rv}(v-2)}{[\text{rv}(v-1)(v-2)]^2} \right\}
\]

\[
= -\frac{\text{rv}(v-2)^2 + \text{rv}^2(v-3)}{(v-2)[\text{rv}(v-1)(v-2)]^2} .
\tag{2.3.4d}
\]

Next look at
\[
E \left[ \frac{1}{s(v-2)} \sum_{j=\text{nv},\text{nv-1}} \left( \varepsilon_{j} - \bar{\varepsilon}_{\text{o..o..}} \right) \right] \left[ \frac{1}{s(v-1)(v-2)} \sum_{j=\text{nv, nv+1}} \left( \varepsilon_{j} - \bar{\varepsilon}_{\text{o..o..}} \right) \right].
\]  

(2.3.5a)

When looking at the products of the given sums we see that there are $s$ terms with $E(\varepsilon_{j} - \bar{\varepsilon}_{\text{o..o..}})^2$ and $4s^2 - s$ terms with $E(\varepsilon_{j} - \bar{\varepsilon}_{\text{o..o..}})(\varepsilon_{k} - \bar{\varepsilon}_{\text{o..o..}})$, $j \neq k$.

\[
E(\varepsilon_{j} - \bar{\varepsilon}_{\text{o..o..}})^2 = \frac{r(v-1)(v-2)}{[r(v-1)(v-2)]^2} - \frac{2r(v-1)(v-2)}{[r(v-1)(v-2)]^2\text{sv}} + \frac{rs(v-1)(v-2)}{[rs(v-1)(v-2)]^2}
\]

\[
= \frac{rsv^2 - rsv^2(v-1)(v-2) - rsv(v-1)(v-2)}{[rs(v-1)(v-2)]^2},
\]  

(2.3.5b)

\[
E(\varepsilon_{j} - \bar{\varepsilon}_{\text{o..o..}})(\varepsilon_{k} - \bar{\varepsilon}_{\text{o..o..}}) = 0 - \frac{rs(v-1)(v-2)}{[rs(v-1)(v-2)]^2}.
\]  

(2.3.5c)

Putting these last two results in (2.3.5a) we have

\[
- \frac{1}{s^2(v-1)(v-2)^2} \left\{ s[rs^2v^2(v-1)(v-2) - rs(v-1)(v-2)] - (4s^2 - s)rs(v-1)(v-2) \right\}
\]

\[
= \frac{-rs^2 + 4rs}{[rs(v-2)(v-1)]^2(v-2)}.
\]  

(2.3.5d)

Adding (2.3.2), (2.3.3d), (2.3.4d), and (2.3.5d) we get

\[
\text{Cov}(y_{0v}, y_{1h}) = \frac{1}{[rs(v-2)(v-1)]^2(v-2)}
\]

\[
x \{ r^2(v-2)^2 - rs[(v-1)(v-2)^2 + v(v-4) - 2(v-2) + (v-1)(v-2) + (v-4)] \}
\]

\[
= \frac{r^2(v-2)^2 - rs(v^3 - 3v^2 - 2)}{[rs(v-2)(v-1)]^2(v-2)} \sigma^2.
\]  

(2.3.6)
Denote the value (2.3.6) by $w_{vh2}$. We list the rest of the needed covariances. Let $d = \left[ rsv(v-1)(v-2) \right]^2$. For $\text{Cov}(y_v, y_h)$, the remaining entries are as follows:

\[
\text{Cov}(y_{0v}, y_{0h}) = -\frac{rsv(v^3+6v^2-13v+10)}{d(v-2)} \sigma^2 = w_{vh1} \sigma^2 \quad (2.3.7a)
\]

\[
\text{Cov}(y_{1v}, y_{0h}) = \frac{rv^2[(v-2)^2 - s(v^2-5v+9)]}{d(v-2)} \sigma^2 = w_{vh3} \sigma^2 \quad (2.3.7b)
\]

\[
\text{Cov}(y_{0v}, y_{2h}) = \frac{rv[(v-2)^2 - s(v^3-5v^2+8v+8)]}{d(v-2)} \sigma^2 = w_{vh4} \sigma^2 \quad (2.3.7c)
\]

Using these results we have the matrix

\[
\text{Cov}(y_v, y_h) = \begin{pmatrix}
    w_{vh1} & w_{vh2} & w_{vh4} \\
    w_{vh3} & w_{vh1} & w_{vh2} \\
    w_{vh4} & w_{vh3} & w_{vh1}
\end{pmatrix} \sigma^2 = w_{vh} \sigma^2 \quad (2.3.8)
\]

For $\text{Var}(y_v)$ we use the following entries

\[
\text{Var}(y_{0v}) = \frac{rsv^2(v-1)(v-2)(v-3)}{d(v-2)} \sigma^2 = w_{v1} \sigma^2 \quad (2.3.9a)
\]

\[
\text{Cov}(y_{0v}, y_{1v}) = -\frac{rsv^2(v-1)(v-3)}{d(v-2)} \sigma^2 = w_{v2} \sigma^2 \quad (2.3.9b)
\]

\[
\text{Cov}(y_{0v}, y_{2v}) = -\frac{rsv^2(v-1)(v-4)}{d(v-2)} \sigma^2 = w_{v3} \sigma^2 \quad (2.3.9c)
\]
Then

\[
\text{Var}(\mathbf{y}_v) = \begin{bmatrix}
  \nu_1 & \nu_2 & \nu_3 \\
  \nu_2 & \nu_1 & \nu_2 \\
  \nu_3 & \nu_2 & \nu_1 \\
\end{bmatrix} \sigma^2 \\
= \mathbf{w}_v \sigma^2 .
\] (2.3.10)

For \(\text{Var}(\mathbf{y}_h)\), \(h \neq v\), we use the following entries.

\[
\text{Var}(y_{0h}) = \frac{\text{rsy}(v-2)[(v-1)^3(v-2)-2]}{d(v-1)(v-2)} \sigma^2 = \mathbf{w}_h \sigma^2
\] (2.3.11a)

\[
\text{Cov}(y_{0h}, y_{1h}) = -\frac{\text{rsy}^2(v^3-6v^2+15v-11)}{d(v-1)(v-2)} \sigma^2 = \mathbf{w}_h \sigma^2
\] (2.3.11b)

\[
\text{Cov}(y_{0h}, y_{2h}) = -\frac{\text{rsy}^2(v^3-6v^2+15v-12)}{d(v-1)(v-2)} \sigma^2 = \mathbf{w}_h \sigma^2.
\] (2.3.11c)

Then

\[
\text{Var}(\mathbf{y}_h) = \begin{bmatrix}
  \mathbf{w}_1 & \mathbf{w}_2 & \mathbf{w}_3 \\
  \mathbf{w}_2 & \mathbf{w}_1 & \mathbf{w}_2 \\
  \mathbf{w}_3 & \mathbf{w}_2 & \mathbf{w}_1 \\
\end{bmatrix} \sigma^2 \\
= \mathbf{w}_h \sigma^2 .
\] (2.3.12)

For \(\text{Cov}(\mathbf{y}_h, \mathbf{y}_k)\), \(h \neq k\), we use the following entries.

\[
\text{Cov}(y_{0h}, y_{0k}) = -\frac{\text{rsy}[(v-1)^2(v-2)^2+4]}{d(v-1)(v-2)} \sigma^2 = \mathbf{w}_{hk} \sigma^2
\] (2.3.13a)

\[
\text{Cov}(y_{0h}, y_{1k}) = \frac{\text{rv}[(v-1)[(v-1)^2+4(v-1)]](v-2)}{d(v-1)(v-2)} \sigma^2 = \mathbf{w}_{hk} \sigma^2
\] (2.3.13b)

\[
\text{Cov}(y_{0h}, y_{2k}) = \frac{\text{rv}[(v-1)^2(v-2)^2-20]}{d(v-1)(v-2)} \sigma^2 = \mathbf{w}_{hk} \sigma^2 .
\] (2.3.13c)
Then
\[
\text{Cov}(\gamma_h, \gamma_k) = \begin{pmatrix}
\hat{w}_{hk1} & \hat{w}_{hk2} & \hat{w}_{hk3} \\
\hat{w}_{hk2} & \hat{w}_{hk1} & \hat{w}_{hk2} \\
\hat{w}_{hk3} & \hat{w}_{hk2} & \hat{w}_{hk1}
\end{pmatrix} \sigma^2
\]
\[= \hat{w}_{hk} \sigma^2. \tag{2.3.14}\]

Using the matrices \( \hat{w}_{vh}, \hat{w}_v, \hat{w}_h, \) and \( \hat{w}_{hk} \) with the estimates given by (2.2.12b) and (2.2.14b), we can calculate the covariances between effects.

\[
\text{Var}(\hat{\delta}_{v-1}) = C_v^V \hat{w}_v C_v^V \sigma^2 = Z_v \sigma^2 \tag{2.3.15a}
\]

\[
\text{Var}(\hat{\delta}_h) = T^{-1} [\hat{w}_h - \hat{w}_h^T C_h C_h - C_h \hat{w}_h C_h + C_h Z_h C_h] T^{-1} \sigma^2 = Z_h \sigma^2 \tag{2.3.15b}
\]

\[
\text{Cov}(\hat{\delta}_{v-1}, \hat{\delta}_h) = [C_v^V \hat{w}_v - Z_v C_h] T^{-1} \sigma^2 = Z_{vh} \sigma^2 \tag{2.3.15c}
\]

\[
\text{Cov}(\hat{\delta}_h, \hat{\delta}_k) = T^{-1} [\hat{w}_{hk} - \hat{w}_{hk}^T C_h C_h - C_h \hat{w}_{hk} C_h + C_h Z_h C_h] T^{-1} \sigma^2 = Z_{hk} \sigma^2. \tag{2.3.15d}
\]

### 2.4 Tests of Hypotheses

Once the estimates of all the treatment effects have been made as in Section 2 and all the variances and covariances have been calculated as in Section 3, we will be able to test for the superiority of certain treatments over the others. If the treatment we select as best will be used in the future for repeated application to a single unit, we are not interested specifically in the individual effects, but in the sum of the direct and all residual effects for each treatment.

Thus we want to find the treatment with the largest value of \( \hat{\delta}_h + \hat{\rho}_h + \hat{\eta}_h \), if it is significantly larger than those of all other treatments. If there is a group of treatments with their sums of effects significantly larger than all other treatments but not significantly
differing among themselves, then the experimenter must use some other means of choosing a treatment from this group for future use. Criteria for this selection might include cost, availability, ease of application, etc.

The number of degrees of freedom available for error estimation in model (2.2.1) with restrictions (2.2.2) is

$$n_e = rsv(v-1)(v-2) - r(v-1)(v-2) - v(s+3)+4. \quad (2.4.1)$$

Also, the maximum likelihood estimates \( \hat{y}_{ijh} \) calculated by minimizing the sum of squares \( G \) of (2.2.3) are

$$\hat{y}_{ijh} = N_{ijh}(\hat{\mu} + \hat{\gamma}_i + \hat{\beta}_j + \hat{\delta}_h + \sum_{p=0}^{v-1} N_{i(j-1)h} \hat{\varrho}_p + \sum_{q=0}^{v-1} N_{i(j-2)q} \hat{\eta}_q). \quad (2.4.2)$$

Using these two results we get the usual maximum likelihood estimate of \( \sigma^2 \) given by

$$\hat{\sigma}^2 = \frac{1}{n_e} \sum_{i=1}^{r} \sum_{j=1}^{v-1} \sum_{h=0}^{v-1} (y_{ijh} - \hat{y}_{ijh})^2. \quad (2.4.3)$$

We are now ready to test the hypotheses of interest. In order to compare sums of treatment effects for two different treatments we use the usual analysis of variance statistic, the exact form of the statistic depending on whether or not one of the treatments being compared is the \((v-1)\)-th. Let the hypothesis be \( H: (\delta_h + \rho_h + \eta_h) = (\delta_k + \rho_k + \eta_k) \), \( h \neq k \). Then if \( H \) is true, the statistic

$$t_{hk} = \frac{(\hat{\delta}_h + \hat{\rho}_h + \hat{\eta}_h) - (\hat{\delta}_k + \hat{\rho}_k + \hat{\eta}_k)}{\text{Estimated}\{\text{Var}(\hat{\delta}_h + \hat{\rho}_h + \hat{\eta}_h) + \text{Var}(\hat{\delta}_k + \hat{\rho}_k + \hat{\eta}_k) \}}$$

$$- 2 \text{cov}[(\hat{\delta}_h + \hat{\rho}_h + \hat{\eta}_h), (\hat{\delta}_k + \hat{\rho}_k + \hat{\eta}_k)] \quad (2.4.4)$$
has the \( t_{\eta_e} \) distribution.

To test \( H_0: (\delta_h + \rho_h + \eta_h) = (\delta_k + \rho_k + \eta_k), \ h \neq k, \ h, \ k \neq v-1 \), we make use of the matrices \( Z_v, Z_h, Z_{vh}, \) and \( Z_{hk} \) in order to derive the denominator of (2.4.4). The estimated variances and covariances are obtained by replacing \( \sigma^2 \) with \( \hat{\sigma}^2 \).

If we want to make a size \( \alpha \) test, \( 0 < \alpha < 1 \), for \( H_0 \), then we reject the hypothesis of interest if \(|t_{hk}| \) is greater than \( t_{\eta_e,1-\alpha/2} \), where \( P(t_{\eta_e} > t_{\eta_e,1-\alpha/2}) = \alpha/2 \). Other tests for relations between treatment effects are carried out according to the usual analysis of variance methods, making use of the results of Sections 2 and 3.
CHAPTER 11

SERIAL ARRAYS AND THEIR ANALYSIS

3.1 Description and Construction

For the triple designs of the previous chapter we assumed that no


treatment was applied more than once to a given unit in any 3 consecu-


tive periods, where 2 was the number of residual effects being consid-

ered for each application. Here we consider n-1, n=2,3,..., residuals


and allow a treatment to be repeated up to n times in any n consecu-


tive periods. Because of this we will also be able to estimate inter-


actions between the same treatment applied at different periods. Thus


these designs are an extension of those commonly referred to as change-


over designs. (See, e.g., [11],[20],[22],[28],[29]). It is also a gen-


eralization of Sampford's [26] type I designs as those use only one


experimental unit. The method we use requires us to sacrifice the pro-


perty of the designs of the previous chapter that each treatment is


applied equally often to every unit. Note also that the estimates of


interaction effects will not be balanced with respect to the effects of


application periods. Section 5 considers a special application of the


designs given in this section.


Before giving the method of design construction we introduce some

terminology. Let V = {0,1,...,v-1} (modulo v) be a set of symbols as


used in definition 2.2.


Definition 3.1: An ordered collection D = (d_1,d_2,...,d_w) of w


symbols from V is called a difference set of size w (in short,
difference set) if it is used to generate a \( r \times (w+1) \) rectangle \( R = (r_{ij}) \) with the following entries.

\[
\begin{align*}
r_{11} &= 0 \\
r_{ij} &= \sum_{k=1}^{j-1} d_k \pmod{v} & \quad j=2,3,\ldots,w+1 \\
r_{ij} &= r_{1j} + (i-1) \pmod{v} & \quad i=2,3,\ldots,v \\
& \quad j=1,2,\ldots,w+1.
\end{align*}
\] (3.1.1a, 3.1.1b, 3.1.1c)

In words, the \( j \)-th element of the first row of \( R \) is the sum \( \pmod{v} \) of the first \( j-1 \) elements of \( D \), and the \( i \)-th row is derived from the first by adding \( i-1 \) \( \pmod{v} \) to each of its elements.

**Example 3.1:** Let \( V = \{0,1,2\} \) and \( D = (2,0,1) \). The rectangle generated by \( D \) is

\[
\begin{array}{cccc}
0 & 2 & 2 & 0 \\
1 & 0 & 0 & 1 \\
2 & 1 & 1 & 2
\end{array}
\] (3.1.2)

We shall use this same example to exhibit later results.

**Definition 3.2:** (Bose [7]). Consider an \( r \times c \) rectangle \( O \) from the set \( V \). If each \( r \times n \) sub-rectangle of \( O \) contains all possible \( n \)-tuples with the same frequency \( \lambda \), the rectangle is called an orthogonal array of strength \( v \), size \( r \), \( c \) constraints, and \( v \) levels. The array may be denoted by \( (r,c,v,n) \) and \( \lambda \) is the index of the array. Clearly \( r = \lambda v^n \), since there are \( v^n \) different \( n \)-tuples.

**Definition 3.3:** Consider an \( r \times c \) rectangle \( S \) with entries from \( V \). \( S \) is called a serial array (SA) of strength \( n \), \( r \) rows, \( c \) columns, \( v \) levels, and index \( \lambda \) if (i) all of the \( v \) symbols occur \( r/v \) times in every column and (ii) all \( n \)-tuples of the \( v \) symbols occur \( \lambda \) times. Such an SA is denoted by \( (r,c,v,n,\lambda) \).
Note from definitions 3.2 and 3.3 that an \((r,c,v,n)\) orthogonal array with index \(\lambda\) is also a \((r,c,v,n,\lambda(c-n+1))\) SA, so that the class of orthogonal arrays is a subclass of the class of serial arrays.

There are \(v^n\) different \(n\)-tuples, \((c-n+1)\) of which occur in each row of a \((r,c,v,n,\lambda)\) SA. Thus the parameters must satisfy the relation

\[ r(c-n+1) = \lambda v^n. \]  

(3.1.3)

One way of choosing \(c\) and \(r\) is to let \(r = v^{n-1}\) and \(c = \lambda v^n + n - 1\). Only SA's of this type will be investigated here, and we shall give a recursive method of construction.

When generating a rectangle from a difference set each \((n-1)\)-tuple of the set gives rise to \(v\) \(n\)-tuples, where the differences between consecutive members of the \(n\)-tuples are simply the elements of the related \((n-1)\)-tuple of the difference set. There is one of these \(n\)-tuples in each row of the rectangle. If every \((n-1)\)-tuple occurs exactly once in a difference set, then every \(n\)-tuple will occur exactly once in the generated rectangle because each different \((n-1)\)-tuple of the difference set gives rise to a unique set of \(v\) \(n\)-tuples in the rectangle. In other words, the rectangle will have \(v \cdot v^{n-1} = v^n\) different, or all possible, \(n\)-tuples. Similarly, if we can find \(v^{n-2}\) difference sets of size \(c-1\) such that every \((n-1)\)-tuple occurs \(\lambda\) times, then by generating a \(v 	imes c\) rectangle from each difference set we will have each \(n\)-tuple occurring a total of \(\lambda\) times in these rectangles. The total number of rows in these rectangles is \(v \cdot v^{n-2} = v^{n-1}\). Now take these \(v^{n-1}\) rows and place them one below the other to form a \(v^{n-1} \times c\) rectangle. This rectangle is a \((v^{n-1},c,v,n,\lambda)\) SA and from (3.1.3) \(c = \lambda v^n + n - 1\). We have proved the following.
Lemma 3.1: If we can find $v^{n-2}$ difference sets of size $\lambda v+n-2$ such that every $(n-1)$-tuple occurs $\lambda$ times, then we can construct a $(v^{n-1},\lambda v+n-1,v,n,\lambda)$ serial array.

Notice that if we consider the $v^{n-2}$ difference sets of the above lemma as the rows of a rectangle then they form a $(v^{n-2},\lambda v+n-2,v,n-1,\lambda)$ SA. By continuing this reverse process of finding difference sets to use in generating SA's, it is seen finally that all we need to construct is a difference set of size $\lambda v$. Then by successive generation of rectangles, using the rows of each SA as difference sets, a $(v^{n-1},\lambda v+n-1,v,n,\lambda)$ SA results after $n-1$ generations. We have proved the following.

Lemma 3.2: Let $V = \{0,1,\ldots,v-1\}$. Write down a difference set of size $\lambda v$ such that each member of $v$ occurs $\lambda$ times. From it we can generate a $(v^{n-1},\lambda v+n-1,v,n,\lambda)$ serial array, $n$ any integer $\geq 2$.

There are $(\lambda v)!/(\lambda!)^v$ such difference sets.

Once we have an SA with specified parameters we can consider the rows as experimental units, symbols as treatments, and columns as order of treatment application in a design for comparing effects of different treatments. With these designs we can estimate direct and residual treatment effects balanced with respect to period effects. We can also estimate interaction effects, but they will not be balanced with respect to period effects. The designs given by Williams [28] as balanced for one residual effect with $v$ even are $(v,v,v,1,1)$ SA's and for $v$ odd are $(2v,v,v,1,2)$ SA's, when the second latin square is written below the first.
Example 3.1a: Rectangle (3.1.2) is a (3,4,3,2,1) SA. We generate a (3^2, 4+1, 3, 2+1, 1) = (9, 5, 3, 3, 1) SA by using the rows of that SA as difference sets.

\[
\begin{array}{cccc}
0 & 2 & 2 & 0 \\
0 & 0 & 2 & 1 \\
1 & 1 & 0 & 2 \\
2 & 2 & 1 & 0 \\
\hline
1 & 0 & 0 & 1 \\
0 & 1 & 1 & 1 \\
1 & 2 & 2 & 2 \\
2 & 0 & 0 & 0 \\
\hline
2 & 1 & 1 & 2 \\
0 & 2 & 0 & 1 \\
0 & 1 & 2 & 1 \\
1 & 2 & 1 & 0 \\
\hline
1 & 0 & 0 & 1 \\
2 & 1 & 2 & 0 \\
2 & 1 & 2 & 0 \\
\end{array}
\]

(3.1.4)

Now place the second rectangle under the first and the third under the second to get the desired SA. □

3.2 General Analysis

Before using an SA to compare effects of different treatments, the experimenter should have a good idea of the maximum number of residual or interaction effects to be included in his model. If this number is n-1, an SA of strength n is needed. For each complete replication of such a design \( v^{n-1} \) experimental units are required, each receiving a series of \( \lambda v + n - 1 \) treatments. Assume that we are to use \( m \) replications of an SA, \( m \) integral, and that we have available \( mv^{n-1} \) homogeneous experimental units. Then we need not allow for unit effects in our response model and it can be written as

\[
y_{ijh} = N_{ijh}(\mu + \beta_j + \sum_{k=0}^{n-1} \sum_{q=0}^{v-1} n_{i(q-j)q} \delta_{kq} + \sum_{k=1}^{n-1} \sum_{q=0}^{v-1} n_{i(j-k)q} \hat{\kappa}_{qh} + \epsilon_{ij})
\]

(3.2.1)

\( i=1,2,\ldots, mv^{n-1}, j=1,2,\ldots, \lambda v, h=0,1,\ldots, v-1, \) where
$y_{ijh}$ is the response for unit $i$ at time $j$ receiving treatment $h$, 
$\mu$ is the overall mean response, 
$\beta_j$ is the effect of time period $j$, 

$$N_i(j-k)q = \begin{cases} 1 \text{ if treatment } q \text{ is applied to unit } i \text{ at time } (j-k) \\ 0 \text{ otherwise} \end{cases},$$

$\delta_{kq}$ is the $k$-th residual effect of treatment $q$ and the 0-th residual is the direct effect, 
$\xi_{kqh}$ is the interaction effect when $h$ is the currently applied treatment and $q$ was applied $k$ periods previously, and the 
$\varepsilon_{ij}$ are independent, identically distributed $N(0,\sigma^2)$.

Note that as in model (2.2.1) a large number of the $y_{ijh}$'s are identically zero. Here there are $mn^2$ observations of interest.

In order to simplify the analysis, we place the following restrictions on the parameters.

$$\sum_{j=1}^{\lambda v} \beta_j = \sum_{q=0}^{v-1} \delta_{kq} = \sum_{h=0}^{v-1} \xi_{kqh} = \sum_{q=0}^{v-1} \xi_{kqh} = 0,$$ (3.2.2)

where the subscripts not included in the summation are held fixed. The number of independent parameters, excluding $\sigma^2$, in (3.2.1) is then

$$1 + (\lambda v - 1) + n(v-1) + 2(n-1)(v-1)^2 = v^2(n-1) + v(\lambda+2-n) - 1.$$ (3.2.5)

Since we are considering the first $n-1$ applications in the design as conditioning and make $\lambda v$ observations on each unit, the total number of observations is $m\lambda v n$. The design we use must have $m\lambda v n$ large enough to allow for the estimation of the $v^2(n-1) + v(\lambda+2-n)-1$ independent parameters and still have sufficient degrees of freedom available for the estimation of $\sigma^2$. This can be accomplished by choos-
ing m or \( \lambda \) large enough, and which of these we choose to increase depends upon the number of treatments being used and the number of homogenous units and time periods available. In the following analysis we assume that a sufficient number of observations are available for estimation of all parameters.

For the model (3.2.1) we will calculate the maximum likelihood estimates of the parameters. This will be accomplished by minimizing the sum of squares

\[
G = \sum_{i=1}^{m} \sum_{j=1}^{v} \sum_{h=0}^{n-1} y_{ijh} - N_{ijh} (\mu + \beta_j + \sum_{k=0}^{n-1} \sum_{q=0}^{v-1} N_{i(j-k)q} \delta_{kq}) + \sum_{k=1}^{n-1} \sum_{q=0}^{v-1} N_{i(j-k)q} \zeta_{kq} \right)^2.
\]

We will derive the normal equations by taking the partial derivatives of \( G \) with respect to the effects. We use \( r = mv^{n-1} \).

(a) \(- \frac{\partial G}{\partial \beta_j} :\)

\[
\sum_{i} \sum_{h} N_{ijh} y_{ijh} - r\mu - r\beta_j - \sum_{i} \sum_{h} \sum_{k=0}^{n-1} \sum_{q} N_{ijh} N_{i(j-k)q} \delta_{kq} - \sum_{i} \sum_{k=1}^{n-1} \sum_{q} N_{ijh} N_{i(j-k)q} \zeta_{kq} = 0.
\]

(The summations range over all possible values of the given subscript unless otherwise indicated.) Looking at the fourth term, for fixed \( k=k' \) we have the sum \( \sum_{i} \sum_{h} N_{ijh} N_{i(j-k')q} \delta_{k'q} \), which is the sum of \( k' \)-th residuals occurring in period \( j \). But in an SA every treatment occurs equally often in every column, including the \( (j-k') \)-th. Thus this sum becomes \( \frac{r}{v} \sum_{q} \delta_{k'q} = 0 \) by (3.2.2), and the fourth term vanishes. We
write (3.2.5a) as

\[ r(\mu + \beta_j) + \sum_{i} \sum_{k} \sum_{q} \sum_{q} N_{ijh} N_{i(j-k)q} kq = y_{ijh}. \]  

(3.2.5b)

(b) \[ \frac{1}{k} \frac{\partial G}{\partial k'q'} : \]

\[ \sum_{i} \sum_{j} \sum_{h} N_{ijh} N_{i(j-k)q} q_{ijh} - \frac{r}{v} \sum_{j} \beta_j \]

\[ - \sum_{i} \sum_{j} \sum_{k} \sum_{q} N_{ijh} N_{i(j-k)q} N_{i(j-k')q} \delta_{kq} \]

\[ - \sum_{i} \sum_{j} \sum_{k=1}^{n-1} \sum_{q} N_{ijh} N_{i(j-k)} p_{i(j-k')q} \zeta_{kq} = 0. \]

(3.2.6a)

The simplified form of the second and third terms arises from properties of SA's and by (3.2.2) the latter term is 0. By rearranging the order of summation in the fourth term it becomes

\[ \frac{r}{v} \delta_{k'q'} + \sum_{k \neq k'} \sum_{n=1}^{n-1} \sum_{q} \sum_{i} \sum_{j} \sum_{h} N_{ijh} N_{i(j-k)q} N_{i(j-k')q} \delta_{kh} = 0. \]

(3.2.6b)

since every \( k' \)-th residual occurs equally often. For fixed \( k \), the \( k \)-th residual effect of each treatment occurs equally often with the \( k' \)-th residual of treatment \( q' \) because every \( n \)-tuple occurs exactly \( m \lambda \) times in our design. Combining this with (3.2.2) the second term of (3.2.6b) vanishes. Similarly, the fifth term of (3.2.6a) is 0 and we have

\[ r(\mu + \delta_{k'q'}) = y_{i(k')q'} . \]

(3.2.6c)

(c) \[ \frac{1}{\gamma} \frac{\partial G}{\partial \gamma_{k'q'}h'} : \]

\[ \sum_{i} \sum_{j} N_{ijh} N_{i(j-k')q} q_{ijh} - \frac{r}{v^2} \mu - \sum_{i} \sum_{j} N_{ijh} N_{i(j-k')q} \beta_j \]
\begin{align}
- \sum_{i} \sum_{j} \sum_{k=0}^{n-1} \sum_{q} ^{N_i(j-k)q} N_{ijh'} i(j-k')q' \delta_{qk} \\
- \sum_{i} \sum_{j} \sum_{k=1}^{n-1} \sum_{q} \sum_{h} ^{N_{ijh} i(j-k)q} N_{ijh'} i(j-k')q' \zeta_{qh'} = 0 \tag{3.2.7a}
\end{align}

First look at the fourth term. When \( k = k' \),

\[ N_{i(j-k')q'} i(j-k')q' = \begin{cases} 1 & \text{if } q = q' \\ 0 & \text{otherwise} \end{cases}, \]

and the sum is \( \sum_{i} \sum_{j} N_{ijh'} \delta_{k'q'} = \frac{r \lambda v}{v^2} \delta_{k'q'} \), since each treatment occurs equally often as the \( k' \)-th residual with the direct effect of \( h' \). When \( k \neq k' \), the sum is 0 because each \( n \)-tuple occurs equally often, as described previously. In the fifth term we have \( N_{ijh} N_{ijh'} = 0 \) if \( h \neq h' \) and \( N_{i(j-k')q} i(j-k')q' = 0 \) if \( q \neq q' \). Thus we can split it into two further terms:

\[ \sum_{i} \sum_{j} N_{ijh'} i(j-k')q' \zeta_{k'q'h'} = \frac{r \lambda v}{v^2} \zeta_{k'q'h'} \tag{3.2.7b} \]

and

\[ \sum_{i} \sum_{j} \sum_{k=1}^{n-1} \sum_{q} \sum_{h} N_{ijh'} i(j-k)q' i(j-k')q' \zeta_{qh'} = 0 \tag{3.2.7c} \]

where (3.2.7b) follows as above and (3.2.7c) as for the fifth term of (3.2.6a). Thus we have

\[ \frac{r \lambda}{v} (\mu + \delta_{k'q'} + \zeta_{k'q'h'}) + \sum_{i} \sum_{j} N_{ijh'} i(j-k')q' \beta_{j} = y(k')q'h' \tag{3.2.7d} \]

where \( y(k')q'h' \) is the sum of all observations where the direct effect of \( h' \) and the \( k' \)-th residual of \( q' \) are present.
(d) \(-\frac{\partial G}{\partial \mu}\):

\[
\sum_{i} \sum_{j} \sum_{h} N_{ijh} y_{ijh} \cdot r \lambda \nu \mu - r \sum_{j} \beta_{j} - \sum_{i} \sum_{j} \sum_{k=0}^{n-1} \sum_{q} N_{ijh} N_{i(j-k)q} \xi_{qk}h = 0.
\]

(3.2.8a)

Using (3.2.2) and the fact that each n-tuple occurs equally often this reduces to

\[
r \lambda \nu \mu = y_{\ast \ast}.
\]

(3.2.8b)

(3.2.5b), (3.2.6c), (3.2.7d), and (3.2.8b) form the set of normal equations. The solutions of these equations depend on the particular SA used for the design because of the summations appearing in (3.2.5b) and (3.2.7d). We will give a solution in matrix form for which we use the following vectors.

\[
\beta' = (\beta_{1}, \beta_{2}, \ldots, \beta_{\lambda v})
\]

(3.2.9a)

\[
\delta' = (\delta_{00}, \delta_{01}, \ldots, \delta_{0(v-1)}, \delta_{10}, \ldots, \delta_{(n-1)0}, \ldots, \delta_{(n-1)(v-1)})
\]

(3.2.9b)

\[
\xi' = (\xi_{100}, \xi_{110}, \ldots, \xi_{1(v-1)0}, \xi_{101}, \ldots, \xi_{(n-1)0(v-1)}, \ldots, \xi_{(n-1)(v-1)(v-1)})
\]

(3.2.9c)

\[
\theta' = (\mu, \beta', \delta', \xi')
\]

(3.2.9d)

\[
Y_{\theta} = (Y_{\ast 1}, Y_{\ast 2}, \ldots, Y_{\ast \lambda v})
\]

(3.2.9e)

\[
Y_{\delta} = (Y_{\ast (k)q})
\]

(3.2.9f)

where the subscripts are ordered as those in \(\delta'\).

\[
Y_{\xi} = (Y_{(k)qh})
\]

(3.2.9g)

where the subscripts are ordered as in \(\xi'\).

\[
Y_{\theta} = (Y_{\ast \ast}, Y_{\beta'}, Y_{\delta'}, Y_{\xi'})
\]

(3.2.9h)
Let $A$ be the matrix of coefficients for the parameters of the normal equations arranged so that

$$Y_\theta = A\theta \quad .$$

(3.2.10a)

Then the usual least squares estimate of $\theta$ is given by

$$\hat{\theta} = A^{-1}Y_\theta \quad ,$$

(3.2.10b)

where $A^{-1}$ is any generalized inverse of $A$.

**Example 3.1b:** Consider the analysis of the $(9,5,3,3,1)$ SA of the previous example. From (3.1.3) we note two things: (i) every 2-tuple occurs in each pair of adjacent columns and (ii) every 2-tuple occurs in columns separated by one other column. As a result the sums in equations (3.2.5b) and (3.2.7d) are both 0, where we also use restrictions (3.2.2).

The vectors (3.2.9a) - (3.2.9h) are

$$\beta^t = (\beta_1, \beta_2, \beta_3) \quad ,$$

(3.2.11a)

$$\delta^t = (\delta_{00}, \delta_{01}, \delta_{02}, \delta_{10}, \delta_{11}, \delta_{12}, \delta_{20}, \delta_{21}, \delta_{22}) \quad ,$$

(3.2.11b)

$$\xi^t = (\xi_{100}, \xi_{110}, \xi_{120}, \xi_{101}, \xi_{111}, \xi_{121}, \xi_{202}, \xi_{212}, \xi_{222}) \quad ,$$

(3.2.11c)

$$\theta^t = (\mu, \beta^t, \delta^t, \xi^t) \quad ,$$

(3.2.11d)

$$\gamma_\theta^t = (y_{s10}, y_{s11}, y_{s12}, y_{s20}, y_{s21}, y_{s22}) \quad ,$$

(3.2.11e)

$$\gamma_\delta^t = (y_{s(0)0}y_{s(0)1}y_{s(0)2}, y_{s(1)0}, y_{s(1)1}),$$

(3.2.11f)

$$\gamma_\xi^t = (y_{s(1)0}y_{s(2)}y_{s(1)1}y_{s(1)2}, y_{s(2)0}, y_{s(2)1}) \quad ,$$

(3.2.11g)

and

$$\gamma_\theta^t = (y_{s10}, y_{s11}, y_{s12}, y_{s20}, y_{s21}, y_{s22}) \quad .$$

(3.2.11h)

The numbers of elements in (3.2.11a) - (3.2.11d) are, respectively, 3, 9, 18, and 30, and the same is true for (3.2.11e) - (3.2.11h).
A has the following form:

\[
\begin{bmatrix}
1 & 3 & 9 & 18 \\
27 & 0.1_3 & 0.9 & 0.1_8 \\
9.1_3 & 9.1_3 & 0.3 \times 9 & 0.3 \times 18 \\
9.1_9 & 0.9 \times 3 & 9.1_9 & 0.9 \times 18 \\
3.1_{18} & 3.1_3 & 3.1_3 & 0.9 \times 3 \\
0.18 \times 3 & 0.18 \times 3 & 3.1_{18} \\
3.1_3 & 3.1_3 & 3.1_3 & 3.1_3 \\
\end{bmatrix}
\]

(3.2.12)

The numbers at the side and top indicate the numbers of rows and columns, respectively, in the given partitions.

The variance matrix of the estimated parameters is

\[
\text{Var}(\hat{\theta}) = A^{-} \sigma^2.
\]

(3.2.13)

Although the generalized inverse \( A^{-} \) is not unique unless \( A \) is non-singular, any estimable function \( b' \hat{\theta} \) of the parameters has a unique variance given by \( b' A^{-} b \sigma^2 \). The maximum likelihood estimate \( \hat{\sigma}^2 \) of \( \sigma^2 \) is calculated by replacing the parameters of (3.2.4) by their maximum likelihood estimates and dividing the sum \( G \) by

\[
n_e = m \lambda v - v^2(n-1) - v(\lambda+2-n) + 1.
\]

(3.2.14)
Tests of hypotheses are carried out as in the previous chapter and are discussed for a special case in the next section.

3.3 Analysis and Hypothesis Testing When Interactions are Excluded

In many cases the experimenter will test for the absence of interaction effects in an attempt to simplify the form of his response model. For this test he will need to estimate the parameters of the model

$$y_{ijh} = N_{ijh}(\mu + \beta_j + \sum_{k=0}^{n-1} \sum_{q=0}^{v-1} N_{i(j-k)q} \delta_{kq} + \epsilon_{ij})$$  \hspace{1cm} (3.3.1)$$

where the symbols and subscripts have the same meaning as in (3.2.1) and the relevant restrictions from (3.2.2) hold. The normal equations are (3.2.5b), (3.2.6c), and (3.2.8b), where the summation of interactions in (3.2.5b) is no longer present.

$$r\lambda\nu\mu = y_{o..} \hspace{1cm} (3.3.2a)$$
$$r(\mu + \beta_j) = y_{oj..} \hspace{1cm} (3.3.2b)$$
$$r\lambda(\mu + \delta_{kq}) = y_{o(k)q} \hspace{1cm} (3.3.2c)$$

These equations have the solutions

$$\hat{\mu} = \bar{y}_{o..} \hspace{1cm} (3.3.3a)$$
$$\hat{\beta}_j = \bar{y}_{o..} - \bar{y}_{o..} \hspace{1cm} (3.3.3b)$$

and

$$\hat{\delta}_{kq} = \bar{y}_{o(k)q} - \bar{y}_{o..} \hspace{1cm} (3.3.3c)$$

where $\bar{y}$ indicates the mean of the sum of observations $y$. From these equations we can determine the variances of and covariances between estimates in a manner similar to that used in Section 2.3. We make use of the fact that there are $r$ summands in $y_{o..}$ and $\frac{r\lambda\nu}{\nu} = r\lambda$ summands in $y_{o(k)q}$.
\[
\text{var}(\hat{\beta}_j) = \left( \frac{1}{r} - \frac{2}{r \lambda v} + \frac{1}{r \lambda \nu} \right) \sigma^2 = \frac{\lambda v - 1}{r \lambda \nu} \sigma^2 \\
\text{var}(\hat{\delta}_{kq'}) = \left( \frac{1}{r \lambda} - \frac{2}{r \lambda v} + \frac{1}{r \lambda \nu} \right) \sigma^2 = \frac{v - 1}{r \lambda \nu} \sigma^2
\] (3.3.4a)

\[
\text{cov}(\hat{\beta}_j, \hat{\delta}_{kq'}) = 0
\] (3.3.4c)

since by construction the estimates are orthogonal,

\[
\text{cov}(\hat{\beta}_j, \hat{\beta}_j) = -\frac{1}{r \lambda \nu} \sigma^2
\] (3.3.4d)

\[
\text{cov}(\hat{\delta}_{kq'}, \hat{\delta}_{k'q'}) = \begin{cases} 
-\frac{1}{r \lambda \nu} \sigma^2 & \text{if } k=k', \ q\neq q' \\
\left( \frac{m \lambda v n - 2}{r^2 \lambda^2} - \frac{2}{r \lambda v} + \frac{1}{r \lambda \nu} \right) \sigma^2 & \\
\frac{m \lambda v n - 1 - r \lambda}{r^2 \lambda^2 \nu} \sigma^2 & \text{if } k\neq k'.
\end{cases}
\] (3.3.4e)

The number of independent parameters, excluding \( \sigma^2 \) in (3.3.1) is \( 1 + (\lambda v - 1) + n(v - 1) = v(n+\lambda) - n \), which leaves

\[
n_{\text{ew}} = m \lambda v n - v(n+\lambda) + n
\] (3.3.5)

degrees of freedom for estimating \( \sigma^2 \). This estimate is

\[
\hat{\sigma}_{w}^2 = \frac{1}{n_{\text{ew}}} \sum_{i=1}^{r} \sum_{j=1}^{\lambda v} \sum_{h=0}^{v-1} \left[ y_{ijh} - N_{ijh}(\hat{\mu} + \hat{\beta}_j + \sum_{k=0}^{n-1} \sum_{q=0}^{v-1} N_{i(j-k)q} \hat{\delta}_{kq}) \right]^2.
\] (3.3.6)

It is then clear that the usual F-test, which is a function of the likelihood ratio test, for the hypothesis \( H_0 : \tau_{kqh} = 0 \), all \((k,q,h)\), is based on the statistic

\[
F = \frac{(n_{\text{ew}} \hat{\sigma}_{w}^2 - n_{e} \hat{\sigma}_{w}^2) / (n_{\text{ew}} - n_{e})}{\hat{\sigma}_{w}^2}.
\] (3.3.7)

If \( H_0 \) is true, \( F \) has the \( F_{n_{\text{ew}} - n_{e}, n_{e}} \) distribution. Thus for a size
a test, \(0 < \alpha < 1\), we reject \(H_0\) if \(F > F_{\alpha}\), where \(P(F_{n_{ew} - n, n_{ew} > F_{\alpha}}) = \alpha\).

Now to compare treatment effects when interactions are taken to be zero, either after performance of the above test or by assumption, we can apply directly the method of Section 2.4. The test statistic for the hypothesis \(H_1: \delta_{kq} = \delta_{kq'}, q \neq q'\), is

\[
\begin{align*}
t_{kqq'} &= \frac{\delta_{kq} - \delta_{kq'}}{2\hat{\sigma}^{2}/r\lambda},
\end{align*}
\]

which under \(H_1\) has the \(t_{n_{ew}}\) distribution. \(H_1\) is the hypothesis that two different treatments have equal k-th residual effects, \(k=0,1,\ldots,n-1\). Again we may want to select a "best" treatment, i.e., one which has the greatest overall effect as measured by \(\sum_{k=0}^{n-1} \delta_{kq}\). For comparison of these sums of effects for treatments \(q\) and \(q'\) we use the statistic analogous to (2.4.4),

\[
\begin{align*}
t_{qq'} &= \frac{\sum_{k=0}^{n-1} (\delta_{kq} - \hat{\delta}_{kq'})}{\text{Estimated variance of } \sum_{k=0}^{n-1} (\delta_{kq} - \hat{\delta}_{kq'})},
\end{align*}
\]

which under \(H_1\) has the \(t_{n_{ew}}\) distribution. We now calculate the required variance.

\[
\begin{align*}
\text{var}(\sum_{k=0}^{n-1} \delta_{kq}) &= \frac{n(\lambda v - 1)}{r\lambda v} \sigma^2 + n(n-1)\frac{(m\lambda v)^{n-1} - r\lambda}{r^2 \lambda^2 v} \sigma^2 \\
\text{cov}(\sum_{k=0}^{n-1} \delta_{kq}, \sum_{k=0}^{n-1} \delta_{kq'}) &= \frac{n}{r\lambda v} \sigma^2 + n(n-1)\frac{(m\lambda v)^{n-1} - r\lambda}{r^2 \lambda^2 v} \sigma^2 \\
\text{var}(\sum_{k=0}^{n-1} \delta_{kq} - \sum_{k=0}^{n-1} \hat{\delta}_{kq'}) &= 2 \left[ \text{var}(\sum_{k=0}^{n-1} \delta_{kq}) - \text{cov}(\sum_{k=0}^{n-1} \delta_{kq}, \sum_{k=0}^{n-1} \hat{\delta}_{kq'}) \right] \\
&= \frac{2n}{T} \sigma^2.
\end{align*}
\]
Using this result (3.3.9) becomes
\[
t_{qq'} = \frac{\sum_{k=0}^{n-1} (\hat{\delta}_{kq} - \hat{\delta}_{kq'})}{2nc^2/r}.
\]

(3.3.11)

3.4 Blocking of Units

For the type of serial arrays we are considering in this chapter, a total of \(mv^{n-1}\) experimental units are required. As \(m, v,\) or \(n\) increases in size the difficulty of finding such a large number of homogeneous units also grows, and we would like to have some method for mitigating this problem. A clue to such a method is given by the recursive nature of the SA construction which we have given. Recall that each column includes all \(v\) treatments in the first \(v\) rows, second \(v\) rows, and each succeeding \(v\) rows. Let \(u\) be any integer such that \(mv^{n-1}/uv = s\) is an integer. Now we need only find \(s\) groups of \(uv\) homogeneous units each and consider each successive group of \(uv\) rows in our SA as a block. Assume that the differences in responses between blocks may be accounted for by independent random variables with mean 0 and common variance. That is, for the case of no interactions, (3.3.1) becomes

\[
Y_{ijh} = N_{ijh}(\mu + \beta_j + \sum_{k=0}^{n-1} \sum_{q=0}^{v-1} N_i(j-k)q \delta_{kq} + \sum_{g=1}^{s} M_{ig} \alpha_g + \epsilon_{ij}),
\]

(3.4.1)

where
\[
M_{ig} = \begin{cases} 
1 & \text{if unit } i \text{ is in block } g \\
0 & \text{otherwise} 
\end{cases}
\]

\(\alpha_g \sim \mathcal{N}(0, \sigma^2_\alpha)\) is the random effect of block \(g\)

\(\text{cov}(\alpha_g, \alpha_{g'}) = 0, g \neq g',\) and

\(\text{cov}(\epsilon_{ij}, \alpha_g) = 0, \) all \((i, j, g)\).
This model is analogous to the one used for inter- and intra-block analysis in randomized block designs. If we carry out that analysis here we will see that no information about treatment effects is gained by using the totals of observations for the blocks, due to the orthogonality of block totals and these effects. Recalling from the previous paragraph that each treatment occurs exactly \( u \) times in each column of a block in our SA, the total of all observations in block \( g \), denoted by \( y_g \) is

\[
y_g = u\lambda \nu + \sum_{i,j, \text{block } g} \epsilon_{ij} + u\lambda \nu g
\]

(3.4.2a)

and

\[
E y_g = u\lambda \nu,
\]

(3.4.2b)

\[
E(y_g - u\lambda \nu)^2 = u\lambda \nu (\sigma^2 + u\lambda \nu \alpha^2),
\]

(3.4.2c)

where \( \hat{\mu} \) is the maximum likelihood estimate of \( \mu \) derived previously. Thus, in the case where we assume no interactions, the expected values of the block totals give us no additional information about the treatment effects. Only the variance structure of the model is changed, and the estimates of the variances are derived from (3.3.6) and (3.4.2c).

In the case where interactions are included in the model, these results no longer hold. Because the interactions do not occur in any particular pattern in an SA, in general, the block totals will now include the sum of all interaction effects occurring in the given block. (3.4.2b) is now

\[
E y_g = u\lambda \nu + \sum_{i, \text{block } g} \lambda \nu \sum_{j=1}^{n-1} \sum_{k=0}^{v-1} \sum_{q=0}^{1} \sum_{h=0}^{n} N_{ijh} N_{i(j-k)q} \xi_{k,q,h}.
\]

(3.4.3)

We will not further consider the analysis of this case.
3.5 Levels of a Single Factor

In this section we consider the adaptation of serial arrays to a different experimental situation. Instead of \( v \) different treatments, suppose we have \( v \) equally spaced levels of a single factor whose direct and residual effects can be represented as polynomials. Let \( x_{ijq} \) denote level \( q \) being applied at time \( j \) to unit \( i \), where the lowest level is denoted by 0, the next by 1, etc., the highest being denoted by \( v-1 \). The application order of these levels is determined by the SA being used for the design, the rows representing units and the columns order of application. If we again assume that there are no interaction effects, we write the response model as

\[
y_{ij} = \mu + \beta_j + \sum_{k=0}^{n-1} \sum_{q=0}^{v-1} N_i(j-k)q p_k(x_{ijq}) + \varepsilon_{ij}, \quad (3.5.1a)
\]

where \( p_k \) represents the polynomial for the \( k \)-th residual effect, and the \( \varepsilon_{ij} \) are independent and identically distributed \( N(0, \sigma^2) \). The restriction

\[
\sum_{j=1}^{\lambda v} \beta_j = 0 \quad (3.5.1b)
\]

is still assumed to hold.

Orthogonal polynomials will be used to simplify estimation of the unknown coefficients in the \( p_k \). Let \( \xi_{dij} \) denote the \( d \)-th degree orthogonal polynomial for the level applied to unit \( i \) at time \( j \). Then

\[
\sum_{q=0}^{v-1} N_i(j-k)q p_k(x_{ijq}) = a_{1i}^{k} \xi_{1ij} + a_{2i}^{k} \xi_{2ij} + \cdots + a_{d_k}^{k} \xi_{d_k i (j-k)} \quad (3.5.2)
\]

where \( d_k \) is the degree of the polynomial for the \( k \)-th residual effect.
The number of parameters, excluding $\sigma^2$, to be estimated is
$$\lambda + \sum_{k=0}^{n-1} d_k \ldots$$
In the SA we choose for our design we must have $m\lambda n$
sufficiently large to allow for estimation of all these parameters.

For model (3.5.1) we assumed that there were no interaction effects.
Actually, because of (3.5.2), if we assume a certain additive form for
interaction effects, they will fit into the given model. Let the $k$-th
interaction effect be of the form
$$b^k_1 (\xi_{ij} + \xi_{i(j-k)}) + \ldots + b^k_d (\xi_{dij} + \xi_{d(i-j-k)})$$
(3.5.3)
a sum of the respective direct and residual orthogonal polynomials, $d_k$
determined by $k$. Then by rearrangement of terms these effects will fit
into the form (3.5.2), $b^k_d$ contributing to both $a^0_d$ and $a^k_d$, all $k$
and $d$.

Now we derive the normal equations from the sum of squares
$$G_1 = \sum_{i=1}^{r} \sum_{j=1}^{n} (y_{ij} - \mu - \beta_j - \sum_{k=0}^{n-1} \sum_{d=1}^{d_k} a^k_d \xi_{d(i-k)})^2$$
(3.5.4)
where we have made use of (3.5.2). We introduce the further notation
that $\xi_{dh}$ is the $d$-th degree orthogonal polynomial for level $h$.

(a) $\frac{\partial G_1}{\partial \beta_j} = 0$

$$\sum_{i} y_{ij} - \mu - \beta_j - \sum_{k} \sum_{d} a^k_d \xi_{d(i-k)} = 0$$
(3.5.5a)
The summations are over the entire range of the indices unless otherwise
noted. For fixed $k$ and $d$ the fourth term may be written
$$\sum_{i} a^k_d \xi_{d(i-k)} = a^k_d \frac{r}{v} \sum_{h} \xi_{dh}$$
(3.5.5b)
since every level occurs $\frac{r}{v}$ times in each column of an SA.
By construction of sets of orthogonal polynomials this sum is 0 for
all k and d, so we have
\[ r(\beta_j + \mu) = \gamma_0 \] (3.5.5c)
(Note that the final subscript on \( \xi_{di(j-k)} \) may be 0 or negative.
Such values indicate application during the conditioning periods before
observations were taken. A value of -(n-2) represents the first con-
ditioning period and 0 the last.)
(b) \( - \frac{L}{2} \sum \frac{\partial G}{\partial a_d} \):
\[ \sum_{i, j} \sum_{d} \xi_{di(j-k)} \gamma_{ij}^{-\mu - \beta_j} \sum_{k, d} a_d^{\beta_j} \xi_{di(j-k)} = 0 . \] (3.5.6a)
For fixed j, the terms containing \( \mu \) and \( \beta_j \) have the form
\[ (\mu + \beta_j) i \sum \xi_{di(j-k')} = \frac{\xi}{v} (\mu + \beta_j) i \sum \xi_{d'h} = 0 . \] (3.5.6b)
For fixed d and \( k \neq k' \) the quadruple sum is
\[ a_d^k \sum_{i, j} \sum \xi_{di(j-k')} \xi_{di(j-k)} = \frac{r \lambda}{v} a_d^k \sum_{h, q} \xi_{d'h} \xi_{d'q} . \] (3.5.6c)
which holds because every n-tuple occurring equally often implies that
every pair of levels occur together as \( k' \)-th and \( k \)-th residuals, respec-
tively, equally often. Fixing \( h \) in the RHS sum we have
\[ \sum_{q} \xi_{d'h} \xi_{dq} = 0 , \] (3.5.6d)
again by construction of orthogonal polynomials. Thus for \( k \neq k' \) all sum-
mands in the last term of (3.5.6a) reduce to 0. If \( k = k' \), the sum becomes
\[ \sum_{i, j} \sum_{d} a_d^{k'} \xi_{di(j-k')} \xi_{di(j-k')} = \frac{r \lambda a_d^{k'}}{h} \sum \xi_{d'h} \xi_{d'k'}/d' , \] (3.5.6e)
since only one level is applied at any particular time \((j-k')\) on unit
i, and each level is applied equally often as the k'-th residual
in the design. Thus (3.5.6a) may be written
\[ r\lambda a_{d'}^{k'} \sum_{h} \xi_{d'h}^2 = \sum_{i \in j} \xi_{d'i}^{(j-k')}y_{ij}. \]  
(3.5.6f)

(c) \[ -\frac{1}{2} \frac{\partial G_i}{\partial \mu} : \]
\[ \sum_{i \in j} y_{ij} - r\lambda \nu \mu - r \sum_{j} \beta_j - \sum_{i \in j} \sum_{k \in d} a_{d'}^{k'} \xi_{d'i}(j-k) = 0. \]  
(3.5.7a)

For fixed k=k' and d=d' the quadruple sum is
\[ a_{d'}^{k'} \sum_{i \in j} \xi_{d'i}(j-k') = r\lambda \sum_{h} \xi_{d'h} = 0. \]  
(3.5.7b)

Using this and (3.5.1b) we have
\[ r\lambda \nu \mu = y_{..}. \]  
(3.5.7c)

The normal equations are given by (3.5.5c), (3.5.6f) and (3.5.7c)
and their solution is
\[ \hat{\mu} = \bar{y}_{..} \]  
(3.5.8a)
\[ \hat{\beta}_j = \bar{y}_{..j} - \bar{y}_{..}. \]  
(3.5.8b)
\[ \hat{a}_d^{k} = \frac{r}{\sum_{i=1}^{r} \xi_{di}(j-k) y_{ij} / r\lambda \sum_{h=0}^{2} \xi_{dh}}. \]  
(3.5.8c)

In order to show that these estimates are uncorrelated and to find the
variance matrix for them we write the normal equations in matrix form.

We use the following two vectors.
\[ \chi = (y_{11}, y_{12}, \ldots, y_{1,\nu}, y_{21}, \ldots, y_{r,\nu}) \]  
(3.5.9a)
\[ \phi = (\mu, \beta_1, \beta_2, \ldots, \beta_{\nu}; a_1^0, a_2^0, \ldots, a_d^0, a_1^1, \ldots, a_d^{n-1}). \]  
(3.5.9b)

The design matrix \( B' \) is of the form
\[ B' = \begin{bmatrix} \frac{1}{\lambda^v} & I_{\lambda^v} & \Xi_1 \\ \vdots & \vdots & \vdots \\ \frac{1}{\lambda^v} & I_{\lambda^v} & \Xi_r \end{bmatrix} \lambda^v \]

where the partitions are as indicated, its size is \( r\lambda^v \times \)

\[ (1+\lambda^v\sum_{k=0}^{n-1} d_k, \text{ and} \]

\[ \Xi_i = \begin{bmatrix} \xi_{i1i} & \xi_{2i1} & \cdots & \xi_{d_0i1} & \xi_{i10} & \cdots & \xi_{d_{n-1}i1}(2-n) \\ \xi_{i2i} & \xi_{2i2} & \cdots & \xi_{d_0i2} & \xi_{i11} & \cdots & \xi_{d_{n-1}i1}(3-n) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \xi_{i1i}, \lambda^v & \xi_{2i}, \lambda^v & \cdots & \xi_{d_0i}, \lambda^v & \xi_{i1},(\lambda^v-1) & \cdots & \xi_{d_{n-1}i},(\lambda^v-(n-1)) \end{bmatrix} \]

is the set of orthogonal polynomials for unit \( i \). Using this notation the model (3.5.1) can be written

\[ \chi = B'\phi + \Xi \]

where \( \Xi \) is the \( r\lambda^v \)-vector of errors \( \varepsilon_{ij} \). Then the normal equations (3.5.5c), (3.5.6f), and (3.5.7c) are also

\[ BB'\phi = B\chi \]

whose solutions (3.5.8a), (3.5.8b), and (3.5.8c) are

\[ \hat{\phi} = (BB')^{-1}B\chi \]
But

\[
\Xi_1 \Xi_1' = \\
\begin{bmatrix}
\sum_{j=1}^{\lambda} 2^{v-1} \xi_{1ij} \\
\sum_{j=0}^{\lambda} \xi_{d,0}^{v-1} \xi_{1ij} \\
\vdots \\
\sum_{j=0}^{\lambda} \xi_{d,1}^{v-1} \xi_{1ij} \\
\sum_{j=2-n}^{\lambda} \xi_{d,n-1}^{v-1} \xi_{1ij}
\end{bmatrix}
\]

(3.5.13)

So letting

\[
D_k =
\begin{bmatrix}
\sum_{h=0}^{v-1} 2^{r} \xi_{1i}^{2h} \\
\sum_{h=0}^{v-1} 2^{r} \xi_{2h}^{2h} \\
\vdots \\
\sum_{h=0}^{v-1} 2^{r} \xi_{d,2h}^{2h}
\end{bmatrix}
\]

we have

\[
BB' = \\
\begin{bmatrix}
r\lambda v & r1' \lambda v & 0' \Sigma_{d,2} \\
\frac{-r_{1} \lambda v}{r_{1} \lambda v} & \frac{-r_{1} \lambda v}{r_{1} \lambda v} & \frac{D_{0}}{D_{0}} & 0 \\
\frac{-r_{1} \lambda v}{r_{1} \lambda v} & \frac{-r_{1} \lambda v}{r_{1} \lambda v} & \frac{D_{1}}{D_{1}} & \ddots \\
\frac{0' \Sigma_{d,2}}{D_{0,2}} & \frac{0' \Sigma_{d,2}}{D_{0,2}} & \frac{D_{n-1}}{D_{n-1}}
\end{bmatrix}
\]

(3.5.15)

Because of the block diagonal form of \(BB'\), its generalized inverse is obtained by replacing each of its non-zero matrices on the diagonal with its own generalized inverse. Each \(D_k, k=0,1,\ldots,n-1\), is non-
singular and has a true inverse obtained by replacing each of its
diagonal elements with its inverse. In the upper left of $BB'$ we
have
\[
\begin{pmatrix}
  r_{\lambda\lambda} & r_{1\lambda}^{\prime} \\
  r_{1\lambda} & r_{1\lambda}^{\prime}
\end{pmatrix}
\]
which has as a generalized inverse
\[
\begin{pmatrix}
  0 & 0^{\prime} \\
  0 & 0
\end{pmatrix}
\]
Thus, as discussed on page 37, $\text{var}(\hat{\phi}) = (BB')^{-1}\sigma^2$
\[
= \begin{pmatrix}
  0 & 0^{\prime} + \Sigma_k d_k \\
  0 & r_{\lambda\lambda}^{-1}
\end{pmatrix}
\]
\[
\begin{pmatrix}
  0^{\prime} + \Sigma_k d_k & 0 \\
  r_{\lambda\lambda}^{-1} & D_0^{-1}
\end{pmatrix}
\]
\[
\begin{pmatrix}
  0 & D_1^{-1} \\
  0 & \ddots
\end{pmatrix}
\]
\[
\begin{pmatrix}
  0 & \cdots & 0
\end{pmatrix}
\]
Thus the period effect and coefficient estimates are orthogonal, and
tests of hypothesis can be made using (3.5.8a), (3.5.8b), (3.5.8c) and
(3.5.17).

Again we note that we have $\nu$ factor levels, $n$ effects for each
level, and a polynomial of degree $d_k$ for the $k$-th residual effect,
k=0,1,...,n-1. We must select a $(\nu^{n-1},\lambda\nu+n-1,\nu,n,\lambda)$ SA and repeat it
$m$ times for our design, where $m$ and $\lambda$ are chosen so that the number
of observations $m\lambda\nu^n$ is sufficiently larger than the number of fixed
parameters, $\lambda\nu + \sum_{k=0}^{n-1} d_k$, to allow for estimation of all of them and $\sigma^2$. 
CHAPTER IV

MAXIMUM LIKELIHOOD FOR SERIES OF DEPENDENT RANDOM VARIABLES

4.1 The Problem of Comparing Dependent Observations

In Chapters 2 and 3 we considered designs that were appropriate for estimating direct, residual, and interaction effects of treatments under certain assumptions about the covariances between observations on a single experimental unit. For the cases discussed we assumed that all observations were independent. Now we want to consider the case where the observations on a unit are dependent, and the conditional distribution of an observation, given all previous observations, is determined in terms of unknown population parameters. Observations on different units are assumed to be independent.

There are two specific experimental situations for which this type of analysis would seem to be appropriate. In the first situation each experimental unit receives a single treatment which is applied repeatedly at equally spaced time points. The number of units must exceed the number of treatments under consideration, so that each treatment is applied to at least one unit. Interactions between treatments cannot be estimated in this case, but direct and residual effects can. The experimental units are divided into groups in the second situation. Each unit in a group receives the same series of treatment applications, and each group is assigned a particular series. The series are like those of the previous chapters, the exact treatment order depending on the number of residuals
and interactions under consideration and the specific design selected. In the next chapter we will consider some designs to be used in this situation. A particular case is that every unit receives a different series of treatments.

If we are dealing with series of independent observations, we can use the method of maximum likelihood to estimate all unknown parameters, as we have done previously. For comparing parameters in that case the likelihood ratio test has certain desirable statistical properties. In Section 4 certain results are proved which generalize the work of Billingsley [6] and Bhat [4] and allow us to apply these likelihood methods to the case of dependent observations. Theorems 4.4.1, 4.4.2, and 4.4.3 give the results of interest. Known statistical results which are used in the proofs of these theorems are given in Section 2, and in Section 3 the assumptions concerning the conditional distributions are presented.

4.2 Preliminary Results

Theorem 4.2.1: (Loeve [18; p. 387]).

Let \( \{x_i\} \) be a sequence of dependent random variables. If

\[
\lim_{m \to \infty} \sum_{i=1}^{m} \frac{\text{var}(x_i)}{b_i^2} < \infty \quad \text{as} \quad b_i \to \infty ,
\]  

(4.2.1)

then

\[
\lim_{m \to \infty} \frac{1}{b_m} \sum_{k=1}^{m} \{x_k - E(x_k|x_{k-1}, \ldots, x_1)\} \overset{a.s.}{\to} 0 .
\]  

(4.2.2)

\( (E(x_k|x_{k-1}, \ldots, x_1) \) denotes the conditional expectation of \( x_k \) given \( x_{k-1}, x_{k-2}, \ldots, x_1 )\).

The next theorem follows from Theorem 9.1 and its extension to prove Theorem 1.2 as given in Billingsley [6; chapter 9].
Theorem 4.2.2: (Billingsley). Let \( \{x_i\} \) be a sequence of \( r \)-dimensional random vectors with moments of order \( 2+\delta \), for some \( \delta > 0 \), and let \( F_0, F_1, \ldots \) be a non-decreasing sequence of Borel fields such that \( \mathbb{E}\{x_j^k|F_{m-1}\} = 0 \) almost surely, \( m=1,2,\ldots \), where \( x_j^k \) is the \( j \)-th component of \( x^k \), \( j=1,2,\ldots,v \). Suppose that

\[
\lim_{m \to \infty} \frac{1}{m} \sum_{k=1}^{m} \mathbb{E}(x_k^i x_k^j | F_{k-1}) \overset{a.s.}{\to} \beta_{ij} \tag{4.2.3}
\]

and

\[
\lim_{m \to \infty} \frac{1}{m^{1+\delta/2}} \sum_{i=1}^{m} \mathbb{E}(|x_k^i|^2 + \delta |F_{k-1}|) \overset{a.s.}{\to} 0 \tag{4.2.4}
\]

for all \( i,j=1,2,\ldots,r \), where the matrix \( B = (\beta_{ij}) \) is positive definite. Then

\[
m^{-\frac{\delta}{2}} \sum_{k=1}^{m} x_k^i \overset{L}{\to} N_\mathbb{R}(0,B). \tag{4.2.5}
\]

Theorem 4.2.3: (Billingsley [6; p. 62]). Suppose \( x_m \) is a random vector in \( \mathbb{E}^r \) and \( \mu \) is a probability measure in \( \mathbb{E}^r \) such that \( x_m \overset{L}{\to} \mu \).

Also suppose that \( y_m \) is a second random vector in \( \mathbb{E}^r \) satisfying either

\[
|x_m - y_m| \leq \varepsilon_m |x_m|, \quad \text{p lim} \varepsilon_m = 0 \tag{4.2.6}
\]

or

\[
|x_m - y_m| \leq \varepsilon'_m |y_m|, \quad \text{p lim} \varepsilon'_m = 0. \tag{4.2.7}
\]

In either case \( \text{p lim}(x_m - y_m) = 0 \), so that \( y_m \overset{L}{\to} \mu \). \(|\cdot|\) denotes the usual Euclidean length in \( \mathbb{E}^r \).

We will use a result on sums of quadratic forms (see Graybill [14; p. 88]) to give a small extension to a theorem of Billingsley [6; p. 64].

Theorem 4.2.4: Let \( y \sim N_\mathbb{R}(0,V) \). Suppose that \( y'A_y = \sum_{i=1}^{p} y'A_i y \) and \( y'A_y, y'A_1 y, \ldots, y'A_{p-1} y \) have \( \chi^2 \) distributions with \( n_n, n_1, \ldots, n_{p-1} \) degrees of freedom respectively. If \( y'A_p y \) is non-negative, it is
\( \chi^2_{n_p}, n_p = n - \sum_{i=1}^{p-1} n_i \), and the \( \chi'^A \chi_i, i=1,2,\ldots,p \), are independent.

**Theorem 4.2.5:** (Billingsley). Let \( 1 \leq r_1 \leq r_2 \leq \ldots \leq r_k \). For \( i=1,2,\ldots,k-1 \) let \( H_i: E^r_i \rightarrow E^{r_{i+1}} \) be a linear mapping of rank \( r_i \). Define \( M_k = I \) and \( M_i = H_{k-1} \ldots H_i, i=1,2,\ldots,k-1 \). Suppose \( x_m \) and \( x \) are random vectors in \( E^r_k \), \( x_m \overset{L}{\rightarrow} x \), and \( x \sim N_{r_k}(0,V) \). Then \( M_i' VM_i \) is non-singular and

\[
z_i(m) = x_m' M_i(M_i' VM_i)^{-1} M_i' x \overset{L}{\rightarrow} \chi^2_{r_i}
\]

(4.2.8)

and the \( k-1 \) random variables \( z_{i+1}(m) - z_i(m), i=1,2,\ldots,k-1 \), are asymptotically independent.

**Corollary 4.2.1:** In Theorem 4.2.5 \( z_1(m) \) is asymptotically independent of \( z_{i+1}(m) - z_i(m), i=1,2,\ldots,k-1 \).

**Proof:** Consider \( z_i(m) = (z_i(m) - z_{i-1}(m)) + z_{i-1}(m) \), \( i=2,3,\ldots,k \).

Since \( E^{r_i-1} \subset E^{r_i} \), \( z_i(m) - z_{i-1}(m) \) is non-negative. From Theorem 4.2.5, \( z_i(m) \overset{L}{\rightarrow} \chi^2_{r_i} \) and \( z_{i-1}(m) \overset{L}{\rightarrow} \chi^2_{r_{i-1}} \). Denote the limiting forms of the random variables by \( x'A_i x \) and \( x'A_{i-1} x \), respectively. Applying Theorem 4.2.4 with \( p=2 \), we get \( z_i(m) - z_{i-1}(m) \overset{L}{\rightarrow} \chi^2_{r_i-r_{i-1}} \) and asymptotically independent of \( z_{i-1}(m) \).

Now let \( z_k(m) = \sum_{i=2}^{k} (z_i(m) - z_{i-1}(m)) + z_1(m) \). By the above discussion and Theorem 4.2.4 we have that \( z_1(m) \) and \( z_i(m) - z_{i-1}(m), i=2,3,\ldots,k \), are asymptotically independent. \( \square \)

### 4.3 Notation and Assumptions

Let \( x_i = (x_{i1}, x_{i2}, \ldots, x_{im_i}) \), \( i=1,2,\ldots,s \), be \( s \) independent sequences of possibly dependent random variables, each having its own joint probability density function (p.d.f.) with respect to a \( \sigma \)-finite
measure $\mu^i_1$, given by

$$f^i(x_1^i, x_2^i, \ldots, x^i_{m^i_1}, \vartheta^i) = f^i(x_{m^i_1}^i, \vartheta^i)$$  \hspace{1cm} (4.3.1)

depending on a set of real parameters $\vartheta^i \in \Theta \subseteq \mathbb{R}^{r}$, for some $r$. The conditional p.d.f. of $x_k^i$ given $x_{k-1}^i$ is

$$f^{ik}_{\vartheta^i} = \frac{f^i(x_k^i, \vartheta^i)}{f^i(x_{k-1}^i, \vartheta^i)}$$  \hspace{1cm} (4.3.2)

so that

$$f^i(x_{m^i_1}^i, \vartheta^i) = \prod_{k=1}^{m^i_1} f^{ik}_{\vartheta^i}$$  \hspace{1cm} (4.3.3)

which we assume to be continuous in $\vartheta^i$. Let $\vartheta^{i0}$ be the true value of $\vartheta^i$ and $E_k^i$ denote the conditional expectation of $x_k^i$ given $x_{k-1}^i$.

For clarity in presenting the assumptions below we will drop for the moment the index $i$ in the above notation so that $f^{ik}_{\vartheta^i} = f^k$, $\vartheta^i = \vartheta$, etc.

Partial derivatives with respect to components of $\vartheta$ will be denoted by subscripts. E.g., $f^k_u = \frac{\partial f^k}{\partial \vartheta_u}$, $f^k_{uv} = \frac{\partial^2 f^k}{\partial \vartheta_u \partial \vartheta_v}$, and $f^k_{uvw} = \frac{\partial^3 f^k}{\partial \vartheta_u \partial \vartheta_v \partial \vartheta_w}$.

**Condition 1**: The range of $x_k^i$ given $x_{k-1}^i$ does not depend on $\vartheta$. Differentiation with respect to $\vartheta$ through the third order may be carried out under the integral sign, integration being with respect to $x_k^i$.

These derivatives are also continuous in $\vartheta$.

This condition ensures that $\ln f^k = g^k$ is well defined and that its first three derivatives exist and are continuous with respect to $\vartheta$.

Since $\int f^k d\mu = 1$, where $x$ is the range of $x_k^i$, we have

$$\int_x g^k f^k d\mu = \int_x u f^k d\mu = \int_x f^k d\mu = 0.$$  \hspace{1cm} (4.3.4)
Thus
\[ E_k g_k^u = 0 \quad (4.3.5) \]
and
\[- E_k g_{uv}^k = E_k g_u^k g_v^k = c_{uv}^k(\theta). \quad (4.3.6)\]

To ensure independence of the parameters we assume that the matrix
\[ \Sigma^k_\theta = (c_{uv}^k(\theta)) \]
is nonsingular. Relation (4.3.5) also implies that the
\[ W_m^u = \sum_{k=1}^{m} g_u^k \quad u = 1, 2, \ldots, r \quad (4.3.7) \]
are martingale sequences.

**Condition 2:**
\[ p \lim_{m \to \infty} \frac{1}{m} \sum_{k=1}^{m} |g_{uvw}^k| < \infty \quad (4.3.8) \]
for some neighborhood of \( \theta^0 \) and all \( u, v, w \).

**Condition 3:**
\[ \lim_{m \to \infty} \frac{1}{m} \sum_{k=1}^{m} L_{k} \theta_0 \overset{a.s.}{\to} \Sigma_{\theta_0} = (\sigma_{uv}(\theta_0)), \quad (4.3.9) \]
where \( \Sigma_{\theta_0} \) is a positive definite matrix of constants. Under condition 1 (Cl) this condition holds trivially for \( g_k^k \) identically distributed, \( k=1, 2, \ldots, m \).

**Condition 4:**
\[ \lim_{m \to \infty} \frac{1}{m^{1+\delta/2}} \sum_{k=1}^{m} E_k |g_u^k|^{2+\delta} \overset{a.s.}{\to} 0 \quad (4.3.10) \]
for some \( \delta > 0 \). If the \( f^k \) possess some additional properties this condition may be relaxed or omitted. (See e.g., Billingsley [5], Silvey [27].)
Condition 5:
\[
\lim_{m \to \infty} \frac{1}{m^2} \sum_{k=1}^{m} \text{var}(g_{uv}^k) < \infty \quad (4.3.11)
\]
for all u,v.

Suppose that each of the sequences \( x_i^i, i=1,2,\ldots,s \), and their respective conditional p.d.f.'s \( f^{i_k}_i = f^{i_k}_i \) satisfy C1-C5. Suppose further that the number of observations \( m = \sum_{i=1}^{s} m_i \to \infty \) in such a way that
\[
\lim_{m \to \infty} \frac{m_i^i}{m} = \lambda_i, \quad i=1,2,\ldots,s, \lambda_i \neq 0. \quad (4.3.12)
\]

Let \( \Theta \times \Theta \times \ldots \times \Theta = \Theta^{rs} \in \Theta^{rs} \) be the total parameter space and \( q \leq rs \) be the number of distinct parameters occurring in \( \theta^1 \times \theta^2 \times \ldots \times \theta^s \in \Theta^{rs} \).

Using C1, C3, C4, and the 1-dimensional version of Theorem 4.2.2, we have that \( \frac{1}{\sqrt{m_i^i}} \sum_{k=1}^{m_i^i} g_{ik}^{i_k} \bigg|_{\widehat{\Theta}^i = \widehat{\Theta}^{10}} \) converges in distribution to a normal distribution with mean 0. Thus
\[
p \lim_{m_i \to \infty} \frac{1}{m_i^i} \sum_{k=1}^{m_i^i} g_{ik}^{i_k} \bigg|_{\widehat{\Theta}^i = \widehat{\Theta}^{10}} = 0. \quad (4.3.13a)
\]

Using C1, C3, C5, and Theorem 4.2.1 we have
\[
\lim_{m_i \to \infty} \frac{1}{m_i^i} \sum_{k=1}^{m_i^i} g_{uv}^{i_k} \bigg|_{\widehat{\Theta}^i = \widehat{\Theta}^{10}} a_i^{10} - \sigma_i^{10} (\theta^{10}), \quad (4.3.13b)
\]
i=1,2,\ldots,s, and u,v=1,2,\ldots,r. If some of the parameters are common to more than one \( \widehat{\Theta}^i \), i=1,2,\ldots,s, then \( q < rs \). In any case, let \( \widehat{\Theta} = (\widehat{\Theta}_1, \widehat{\Theta}_2, \ldots, \widehat{\Theta}_q) \) be the vector of the q distinct parameters which we want to estimate. Now considering the derivatives of \( g_{ik}^{i_k} \) with respect to the components of \( \widehat{\Theta} \), if a partial derivative is taken with respect to \( \widehat{\Theta}_u \), if a partial derivative is taken with respect to \( \widehat{\Theta}_u \), then its value is 0. Thus (4.3.13a) and (4.3.13b) hold for
u, v = 1, 2, ..., q, differentiation being with respect to the components of \( \theta \), and let \( \sigma_{uv}(\theta^0) \) be the resulting limit in (4.3.13b). Let
\[
\Sigma^i_0 = (\sigma_{uv}(\theta^0)), \quad \sigma_{uv}(\theta^0) = \sum_{i=1}^{S} \lambda_i \sigma_{uv}^i(\theta^0), \quad \Sigma_0 = \sum_{i=1}^{S} \lambda_i \Sigma^i_0.
\]
Since by C3 each \( \Sigma^i_0 \) is positive definite, so is \( \Sigma_0 \). Combining these results with the manner in which each \( m_i \) increases as \( m \) increases, we have
\[
\lim_{m \to \infty} \frac{1}{m} \sum_{i=1}^{s} \sum_{k_1=1}^{m_i} \sum_{k_i=1}^{m_i} \frac{i k_i}{g_{i1}} = 0 \quad (4.3.14a)
\]
and
\[
\lim_{m \to \infty} \frac{s}{m} \sum_{i=1}^{s} \sum_{k_1=1}^{m_i} \sum_{k_i=1}^{m_i} g_{uv} \Rightarrow \sigma_{uv}(\theta^0), \quad (4.3.14b)
\]
\( u, v = 1, 2, ..., q \). These limits will be used in the proofs of results given in the next section.


For each of the \( s \) independent sequences \( x^{im_i} \), \( i = 1, 2, ..., s \), the likelihood function is the joint p.d.f. of \( x^{im_i} \) as given by equation (4.3.3). Since these sequences are independent the joint likelihood function is the product of the individual likelihood functions
\[
\prod_{i=1}^{s} \prod_{k_1=1}^{m_i} f_{i1}^i. \quad \text{In order to calculate the maximum likelihood estimate}
\]
\( \hat{\theta} \) of \( \theta \) it will be convenient to work with the log likelihood function given by
\[
L_m(\theta) = \sum_{i=1}^{s} \sum_{k_1=1}^{m_i} g_{i1}^i \log f_{i1}^i, \quad m = \sum_{i=1}^{s} m_i, \quad (4.4.1)
\]
which exists under C1.

Theorem 4.4.1: There is a consistent maximum likelihood estimator \( \hat{\theta} \) of \( \theta^0 \). Moreover, \( \hat{\theta} \) is a local maximum of \( L_m(\theta) \) with probability
tending to one, and it is essentially unique.

**Proof:** By C1 and the mean value theorem

\[
m^{-1} \frac{\partial}{\partial \theta} \log L_m(\theta) = m^{-1} \sum_{i=1}^{s} \sum_{k_i=1}^{m_i} g_{u_i}^{i}(\theta^0) \\
+ \sum_{v=1}^{q} \frac{(\theta_v - \theta^0_v) m^{-1}}{\sum_{i=1}^{s} \sum_{k_i=1}^{m_i} g_{u_v}^{i}(\theta^0)} \\
+ m^{-1} |\theta - \theta^0|^2 \sum_{i=1}^{s} \sum_{k_i=1}^{m_i} \alpha_i G^{i}(\theta^0),
\]

(4.4.2)

u=1,2,...,q, for any \( \theta \) contained in the smallest of the neighborhoods of \( \theta^0 \), say \( N \), defined by C2, \( \theta' \) some point in \( N \), \( |\theta - \theta^0| \) denotes length in \( E^q \), each \( |\alpha^i| \leq q^2/2 \), and

\[
G^{i}_{k_i} = \sup_{\theta \in N} |g_{uv}^{i}(\theta)|.
\]

Since \( \Sigma_0 \) is positive definite it follows that there is a \( \beta > 0 \) such that

\[
\sum_{u=1}^{q} \sum_{v=1}^{q} \sigma_{uv}(\theta^0) t_u t_v \geq \beta \text{ for any vector } t \in E^q \text{ such that } |t| = 1.
\]

Suppose \( \epsilon > 0 \) is given. Choose \( \delta = \delta(\epsilon) > 0 \) in such a way that

\[
\{ \theta : |\theta - \theta^0| \leq \delta \} \subseteq N, \quad \delta^2 < \beta/3q^2(M+1), \quad \delta < \epsilon.
\]

(4.4.3)

Then choose \( m_0(\epsilon) \) large enough so that if \( m \geq m_0(\epsilon) \), then with probability exceeding 1-\( \epsilon \),

\[
\left| m^{-1} \sum_{i=1}^{s} \sum_{k_i=1}^{m_i} g_{u_i}^{i}(\theta^0) \right| < \delta \quad u=1,2,...,q
\]

(4.4.4a)

\[
0 \leq m^{-1} \sum_{i=1}^{s} \sum_{k_i=1}^{m_i} G^{i}_{k_i} < M+1 \quad u=1,2,...,q
\]

(4.4.4b)
\[
\left| m^{-1} \sum_{i=1}^{s} \sum_{k_i=1}^{m_i} \frac{ik_i}{g_{uv}^i(\theta^0) + \sigma_{uv}(\theta^0)} \right| < \delta \quad u,1,2,\ldots, q, \quad (4.4.4c)
\]

where \( M \) is some finite constant. These relations follow from C2, (4.3.14a), and (4.3.14b). If these relations are satisfied and
\[
|\theta - \theta^0| \leq \delta, \quad \text{then by (4.4.2)}
\]

\[
\left| m^{-1} \sum_{i=1}^{s} \sum_{k_i=1}^{m_i} \frac{ik_i}{g_{uv}^i(\theta^0) + \sigma_{uv}(\theta^0)} \right|
\leq \delta^2 + \sum_{v=1}^{q} (\theta_v - \theta^0_v) \left[ m^{-1} \sum_{i=1}^{s} \sum_{k_i=1}^{m_i} \frac{ik_i}{g_{uv}^i(\theta^0) + \sigma_{uv}(\theta^0)} \right]
\]

\[
+ |\theta - \theta^0|^2 m^{-1} \sum_{i=1}^{s} \sum_{k_i=1}^{m_i} G_{ij} \alpha_i
\]

\[
\leq \delta^2 + \delta q |\theta - \theta^0| + q^2 |\theta - \theta^0|^2 (M+1)/2
\]

\[
\leq 3q^2 \delta^2 (M+1).
\]

Hence, if \( |\theta - \theta^0| = \delta \) and using (4.4.3),
\[
\sum_{u=1}^{q} \left[ m^{-1} \sum_{i=1}^{s} \sum_{k_i=1}^{m_i} \frac{ik_i}{g_{uv}^i(\theta^0)} \right] (\theta_u - \theta^0_u)
\]

\[
\leq - \sum_{v=1}^{q} \sum_{u=1}^{q} (\theta_u - \theta^0_u) \sigma_{uv}(\theta^0) (\theta_v - \theta^0_v) + 3\delta^3 q^3 (M+1)
\]

\[
\leq - \beta |\theta - \theta^0| + 3\delta^3 q^3 (M+1) < 0.
\]

By Lemma 2 of Aitchison and Silvey [1] there is a value \( \hat{\theta} \) of \( \theta \) such that \( m^{-1} \sum_{i=1}^{s} \sum_{k_i=1}^{m_i} \frac{ik_i}{g_{uv}^i(\hat{\theta})} = 0, u,1,2,\ldots, q \), and such that
\[
|\theta - \theta^0| \leq \delta < \epsilon. \quad \text{Since } \epsilon \text{ was arbitrary } \hat{\theta} \text{ is a consistent solution of (4.4.2).}
We now show that $\hat{\theta}$ is a local maximum and essentially unique.

By the mean value theorem, if $\theta \in \mathbb{N}$,

$$
\frac{\partial^2 L_m(\theta)}{\partial \theta_u \partial \theta_v} = m^{-1} \sum_{i=1}^{\alpha_i} \sum_{k_i=1}^{\mathbf{i}} k_i \left( g_{uv}^{i}(\theta^0) + m^{-1} |\theta - \theta^0| \sum_{i=1}^{\alpha_i} \sum_{k_i=1}^{\mathbf{i}} G^{i} \right)
$$

(4.4.5)

for all $u, v$, $|\alpha_i| \leq q$. Choose $\delta$ so small that $\{\theta : |\theta - \theta^0| \leq \delta\} \subseteq \mathbb{N}$ and that $T_{uv}$ is a $q \times q$ symmetric matrix with $|T_{uv} - \sigma_{uv}(\theta^0)| < \delta[1 + q(M+1)]$ for all $u, v$, then $T_{uv}$ is positive definite. Now equations (4.4.4) hold with probability going to one. Then by (4.4.5)

$$
\left| m^{-1} \frac{\partial^2 L_m(\theta)}{\partial \theta_u \partial \theta_v} + \sigma_{uv}(\theta^0) \right| \leq \delta + q\delta(M+1)
$$

for any $\theta$ such that $|\theta - \theta^0| < \delta$. Then by our choice of $\delta$, with probability going to one, $\frac{\partial^2 L_m(\theta)}{\partial \theta_u \partial \theta_v}$ is negative definite for every $\theta$ such that $|\theta - \theta^0| < \delta$. It follows that with probability tending to one there is a unique solution of the maximum likelihood equations and any such solution is a local maximum of $L_m(\theta)$.

The next theorem gives the asymptotic distribution of the likelihood ratio test for the hypothesis that $\theta = \theta^0$, where $\theta^0$ is a set of points in $\mathbb{R}^d$. It also shows that the maximum likelihood estimates have asymptotically a normal distribution. First we define

$$
\gamma_u(m) = m^{-\frac{1}{2}} \sum_{i=1}^{\mathbf{m}} \sum_{k_i=1}^{\mathbf{i}} k_i g_u^{i}(\theta^0)
$$

(4.4.6)

and

$$
e_u(m) = m^{\frac{1}{2}} (\hat{\theta}_u - \theta_u^0).
$$

(4.4.7)

Let $\gamma(m)$ and $e(m)$ be the vectors with components $\gamma_u(m)$ and $e_u(m)$ respectively. Also let $\max_{\theta} L_m$ be the maximum of $L_m(\theta)$ over $\Theta$ and $L_m^0$ be the maximum on $\theta^0$. 
Theorem 4.4.2: Let \( \hat{\theta}^0 \) be the true set of parameter values and \( \hat{\theta} \) be the maximum likelihood estimates which exist by Theorem 4.4.1. Then

\[
P \lim_{m \to \infty} \frac{1}{m} \sum_{i=1}^{m} i \cdot k_i \cdot G_i(\theta^0) = \frac{1}{m} \sum_{i=1}^{m} \alpha_i \cdot \gamma_i(\theta^0) = 0\tag{4.4.8}
\]

and

\[
2 \left( \max_{\theta \in \Theta} L_m - L_m^0 \right) \leq \chi^2_q. \tag{4.4.9}
\]

Proof: Using C1, C3, C4, Theorem 4.2.2, and equations (4.3.14), we have that \( \gamma(m) \sim N_q(0, \Sigma_0) \). On setting \( \frac{\partial L_m(\theta)}{\partial \theta_u} = 0 \) and using (4.4.2), we have for \( \theta \in \Theta \)

\[
-\gamma_u(m) = \frac{1}{q} \sum_{v=1}^{q} e_v(m) \left[ \sum_{i=1}^{n} \frac{1}{s} \sum_{k=1}^{m} g_i(\theta^0) \right] + \sum_{i=1}^{s} \alpha_i \cdot |\theta - \theta^0|^2 \cdot \varepsilon(m) \cdot \left[ \sum_{k=1}^{m} G_i(\theta^0) \right]
\]

where \( |\alpha_i| \leq \frac{q}{2} \) and \( \theta \in \Theta \). Now \( p \lim_{m \to \infty} |\theta - \theta^0| = 0 \), so by equations (4.4.4b-c) for \( m \geq m_0(\varepsilon) \),

\[
|\gamma_u(m) - \sum_{v=1}^{q} e_v(m) \sigma_{uv}(\theta^0)| \leq \varepsilon_u(m) \varepsilon(m)
\]

where \( p \lim_{m \to \infty} \varepsilon_u(m) = 0 \). Let \( \varepsilon_m = \max_{u} \varepsilon_u(m) \). Then \( |\gamma(m) - \Sigma_0 \varepsilon(m)| \leq q \varepsilon_m \varepsilon(m) \), where \( p \lim_{m \to \infty} q \varepsilon_m = 0 \). Also \( |\Sigma_0 \varepsilon(m)| = c_m |\varepsilon_m| \), where \( c_m \) is a finite constant for each \( m \geq m_0(\varepsilon) \). Then

\[
|\gamma(m) - \Sigma_0 \varepsilon(m)| \leq q \frac{c_m}{c_m} \varepsilon_m \varepsilon(m)
\]

where \( p \lim_{m \to \infty} \frac{c_m}{c_m} = 0 \). By Theorem 4.2.3 \( p \lim_{m \to \infty} \Sigma_0 \varepsilon(m) = \chi(m) \) and thus \( p \lim_{m \to \infty} \varepsilon(m) = \Sigma_0^{-1} \chi(m) \) and \( \varepsilon(m) \sim N_q(0, \Sigma_0^{-1}) \). Again by the same theorem
we now have \( p \lim_{m \to \infty} y(m) = \Sigma \theta(m) \).

The asymptotic distribution of \( 2[\max_\Theta L_m^0 - L_m^0] \) remains to be found. 

For \( \theta \in \mathbb{N} \) we again use the mean value theorem to get

\[
\begin{align*}
\hat{g}^i(\theta) &= g^i(\theta^0) + \sum_{u=1}^{q} (\theta_u - \theta_u^0) g^{iu}(\theta^0) \\
&\quad + \sum_{u=1}^{q} \sum_{v=1}^{q} (\theta_u - \theta_u^0)(\theta_v - \theta_v^0) g_{uv}^{i}(\theta^0) + \alpha_i |\theta - \theta^0|^3 G^i
\end{align*}
\]

where \( |\alpha_i| \leq q^3/6 \). If \( \hat{\theta} \in \mathbb{N} \),

\[
2[\hat{L}_m(\hat{\theta}) - L_m^0] = 2 \sum_{u=1}^{q} (\hat{\theta}_u - \theta_u^0) \sum_{i=1}^{s} \sum_{k_i=1}^{m_i} ik_i(\theta^0) \\
+ \sum_{u=1}^{q} \sum_{v=1}^{q} (\hat{\theta}_u - \theta_u^0)(\hat{\theta}_v - \theta_v^0) \sum_{i=1}^{s} \sum_{k_i=1}^{m_i} g_{uv}^{i}(\theta^0) \\
+ |\theta - \theta^0|^3 \sum_{i=1}^{s} \sum_{k_i=1}^{m_i} \alpha_i G^i.
\]

By (4.4.4b)

\[
|\theta - \theta^0|^3 \sum_{i=1}^{s} \sum_{k_i=1}^{m_i} \alpha_i G^i = |\epsilon(m)|^3 \sum_{i=1}^{s} \sum_{k_i=1}^{m_i} \alpha_i G^i \\
\leq \frac{q^3}{6} |\epsilon(m)|^3 \sum_{i=1}^{s} \sum_{k_i=1}^{m_i} G^i
\]

\( \to 0 \) in probability.

Then

\[
p \lim_{m \to \infty} 2[\hat{L}_m(\hat{\theta}) - L_m^0] = p \lim_{m \to \infty} \left\{ 2 \sum_{u=1}^{q} e_u(m) y_u(m) + \sum_{u=1}^{q} \sum_{v=1}^{q} e_u(m) \sigma_{uv}(\theta^0) \right\}
\]
using (4.4.4c). Applying the first two limits of (4.4.8) to this result we have

$$\lim_{m \to \infty} 2[L_m^{\hat{\theta}} - L_m^0] = \gamma'(m) \Lambda^{-1} \gamma(m).$$

Finally, using the distributional convergence of $\gamma(m)$ and Theorem 4.2.5, (4.4.9) follows.

Now we come to the main problem of interest in this chapter, the testing of equality of parameters between different sequences of dependent random variables, i.e., that the true parameter values lie in some subset $\phi$ of $\Theta$. Suppose that $\{x^n\}$, $i=1,2,\ldots,s$, satisfy C1-C5.

Let $d: \phi \to \Theta$ be a mapping of an open subset $\phi \subset E^c$, csq, into $\Theta$.

Denote any point $\phi$ in $\phi$ by $(\phi_1, \phi_2, \ldots, \phi_c)$ and by $d_j(\phi)$, $j=1,2,\ldots,q$, the components of $d(\phi)$. Also suppose that $d$ satisfies the following condition.

**Condition 6:** The $q \times c$ matrix $D(\phi)$ with entries

$$D(\phi)_{ju} = \frac{\partial d_j(\phi)}{\partial \phi_u} \quad j=1,2,\ldots,q, \ u=1,2,\ldots,c \quad (4.4.11)$$

has rank $c$ throughout $\phi$ and the $d_j$ have continuous third partial derivatives.

**Example 4.1:** Let $r=3$, $s=3$, $q=rs=9$ and $\theta^1 = (\theta^1_1, \theta^1_2, \theta^1_3)$ be the parameter sets for the 3 sequences. Let $\Theta = (\theta^1, \theta^2, \theta^3) \in E^q$ and suppose that we want to test the hypothesis $H_1: \theta^1 = \theta^3, \ i=1,2,3$.

Then $\phi \in E^7$ and we may write

$$\phi = (\theta^1_1, \theta^1_2, \theta^1_3, \theta^2_1, \theta^2_2, \theta^2_3)$$

and

$$d(\phi) = (\theta^1_1, \theta^1_2, \theta^1_3, \theta^2_1, \theta^2_2, \theta^2_3, \theta^3_1, \theta^3_2, \theta^3_3).$$
The matrix $D(\phi)$ is given by

$$
D(\phi) = \\
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 
\end{bmatrix}
$$

which is of rank 7. \qed

Example 4.2: If we have a subset $\frac{\theta}{p}^i = (\theta^i_{p_1}, \theta^i_{p_2}, \ldots, \theta^i_{p_p})$ of size $p$ from each $\theta^i$ and we want to test the hypothesis

$$H_2: \frac{\theta}{p}^i = (\theta_{p_1}, \theta_{p_2}, \ldots, \theta_{p_p})$$

then the representation given above may be extended directly to this case, where $q = rs$ and $c = rs - (s-1)p$. We write $\phi = (\alpha^1, \alpha^2, \ldots, \alpha^s, \alpha^{s+1})$, where

$$\alpha^{s+1} = (\theta_{p_1}, \theta_{p_2}, \ldots, \theta_{p_p})$$

and $\alpha^i = (\theta^i_{p_1}, \theta^i_{p_2}, \ldots, \theta^i_{p_p})$, $i=1,2,\ldots,s$.

Then the components of the mapping are given by

$$d_j(\phi) = \begin{cases}
\theta_{p_k} & j = p_k + (i-1)r \quad i=1,2,\ldots,s, \quad k=1,2,\ldots,p \\
\theta^i_{p_k} & j = p_k + (i-1)r \quad i=1,2,\ldots,s, \quad k=p+1,p+2,\ldots,r.
\end{cases}$$

If we let $\phi^i$ be the set of components of $\phi$ which relate to the i-th sequence, $i=1,2,\ldots,s$, then we can let $d_i(\phi^i)$ represent the mapping $d_i: \phi^i \rightarrow \theta^i$, which is a restriction of $d$. If $\phi^i \in C_i$, then the matrix $D_i(\phi^i)$ defined similarly to $D(\phi)$ has rank $c_i$ throughout $\phi^i$. Let $f^i(\phi^i)$ and $g^i(\phi^i)$ represent $f^i$ and $g^i$ respectively under the assumption that $\phi$ is the true parameter space.
Theorem 4.4.3: Suppose that \( d: \phi \to \Theta \) satisfies C6. Then \( f^{ik}_i(\phi^i) \) satisfies C1-C5, \( i=1,2,\ldots,s \). Moreover, if \( \delta^0 = d(\phi^0) \) is the true value of the parameter set, then

\[
2[\max_{\phi} L_m - L_m^0] \overset{\text{L}}{\sim} \chi^2_c \tag{4.4.12a}
\]

and

\[
2[\max_{\Theta} L_m - \max_{\phi} L_m] \overset{\text{L}}{\sim} \chi^2_{q-c} \tag{4.4.12b}
\]

and these statistics are asymptotically independent.

Proof: By C6 the matrices

\[
E_{d_1}(\phi^i) \left\{ \frac{\partial}{\partial \phi_u^i} g_i^\phi \frac{\partial}{\partial \phi_v^i} g_i^\phi \right\} = D_i^i(\phi^i) \Sigma_d(\phi^i) D_i(\phi^i)
\]

are nonsingular, \( i=1,2,\ldots,s \), where \( d_1(\phi^i) \) indicates that \( \phi^i \) is the set of parameters we are considering and the \( k_i \) are defined by (4.3.6).

By the chain rule for differentiation \( m_i^{-1} \sum_{k=1}^{m_i} \frac{\partial}{\partial \phi_u^i} g_i^\phi \) is a linear function of the \( m_i^{-1} \sum_{k=1}^{m_i} g_u^i(\theta^i) , u=1,2,\ldots,r \). Since each \( f^{ik}_i \) satisfies C1-C5, so does each \( f^{ik}_i(\phi^i) \) and Theorem 4.4.1 holds. Then there is a consistent solution of the maximum likelihood equations

\[
\frac{\partial}{\partial \phi_u} L_m(\phi) = 0 \quad u=1,2,\ldots,c.
\]

Suppose that \( \phi \) is the true set of parameters and \( \phi^0 \) is the true set of values, i.e., \( f^{ik}_i(\phi^0) = f^{ik}_i \). Then

\[
L_m(\phi^0) = L_m(\phi^0) = L_m^0 \quad \text{and} \quad L_m(\phi) = \max_\phi L_m.
\]

Let \( w(m) \) be the random vector with components

\[
w_u(m) = \frac{1}{L} \sum_{i=1}^{m} \sum_{k=1}^{m_i} \frac{\partial}{\partial \phi_u^i} g_i^\phi.
\]

(4.4.13)
By the chain rule \( w(m) = D'(\phi^0)Y(m). \)

\[
\Sigma^m_0 = \left( \frac{1}{m} \sum_{i=1}^{m} \sum_{k_i=1}^{m_i} \sigma_{uv}^{(\phi^0)} \right)
\]

is the variance matrix of \( Y(m) \) and the variance matrix of \( w(m) \) is \( D'(\phi^0)\Sigma^m_0D(\phi^0). \) From the proof of the previous theorem we know that

\[
p\lim_{m \to \infty} 2[\max_{\Theta} L_m - L^0_m] = p\lim_{m \to \infty} w'(m)[D'(\phi^0)\Sigma^m_0D(\phi^0)]^{-1}w(m)
\]

\[
= p\lim_{m \to \infty} \gamma'(m)D(\phi^0)[D'(\phi^0)\Sigma^m_0D(\phi^0)]^{-1}D'(\phi^0)\gamma(m)
\]

\[
= p\lim_{m \to \infty} \gamma'(m)D(\phi^0)[D'(\phi^0)\Sigma^m_0D(\phi^0)]^{-1}D'(\phi^0)\gamma(m).
\]

Now let \( z_2(m) = 2[\max_{\Theta} L_m - L^0_m] \) and \( z_1(m) = 2[\max_{\phi} L_m - L^0_m] \), so that

\( z_2(m) - z_1(m) = 2[\max_{\Theta} L_m - \max_{\phi} L_m]. \) But \( Y(m) \overset{L}{\sim} N(0, \Sigma^m_0) \) and by Theorem 4.2.5 and its corollary the results of this theorem follow.

**Example 4.3:** For examples 4.1 and 4.2, respectively, we have that

\[
2[\max_{\Theta} L_m - \max_{H_1} L_m] \overset{L}{\sim} \chi^2_1 \quad \text{and} \quad 2[\max_{\Theta} L_m - \max_{H_2} L_m] \overset{L}{\sim} \chi^2_{(s-1)p}.
\]

4.5 **Summary**

Our objective in this chapter has been to find a method for estimating and comparing treatment effects when our observations consist of a number of observations of independent sequences of dependent random variables. If the conditional probability density function of each observation in a series exists and satisfies conditions 1 through 5, then by using the maximum likelihood method we have the following results:
(1) by Theorem 4.4.1 the estimates of the parameters of the conditional densities are consistent;

(2) by Theorem 4.4.2 the estimates have asymptotically a joint multinormal distribution; and

(3) by Theorem 4.4.3 the likelihood ratio tests of hypotheses of the form "H: \( \Theta \subset \Theta \) is the true set of parameters" have asymptotically a \( \chi^2 \) distribution if the mapping \( d: \Theta \rightarrow \Theta \), defined by \( H \), satisfies condition 6. These results are exactly analogous to those for the usual case of sequences of independent random variables.

It should be noted that some of the conditions 1-5 are weaker than other assumptions used to prove asymptotic normal distribution in the dependent case. See, e.g., Diananda [9], Hoeffding and Robbins [15], Silvey [27], and Billingsley [5],[6]. If the conditional densities under consideration satisfy one of these other assumptions, then some of C1-C5 will be unnecessary for the results to hold.
CHAPTER V
CONSTRUCTION OF SOME 2- AND 3- REPEATING DESIGNS

5.1 n-Repeating Designs

Here we consider some designs which may be useful in collecting observations for which the analysis of the previous chapter is applicable. Consider the situation in which each experimental unit receives a specified series of treatments, and we wish to estimate direct, residual, and interaction effects for dependent sets of observations. If we suppose that we know the form of the conditional densities and that they satisfy the conditions of Chapter 4, then we can use the method of maximum likelihood to estimate the treatment effects if they are present in the conditional densities. An obvious requirement of any design to be used for this purpose is that as the total number of observations on a unit increases so does the number of appearances of each treatment parameter of interest occurring on that unit. If some parameter, say the u-th, appears only a finite number of times with unit i, then $L_0^i$ will have a zero in position $(i,i)$ and will not be positive definite. This is a violation of C3.

Another desirable property of such designs is for each direct and residual effect to occur equally often, each first interaction to occur equally often, each second interaction to occur equally often, etc. Also we would like any member of a class of effects to occur equally often with each member of another class. In the particular case of asymptotic estimation, the estimates of elements in the variance matrix for the estimated parameters are given by the negative of the summations in the
LHS of (4.3.14b). If the above properties are satisfied by the designs, then an equal number of terms will contribute to each estimated variance for a particular class of effects, and similar results hold for the estimated covariances. These conditions are satisfied by letting each n-tuple or n-plet, whichever is being considered, occur equally often in the design. If we allow a treatment to occur more than once in any n consecutive applications, the serial arrays of Chapter 3 can be used if carefully chosen. If we allow a treatment to occur at most once in any n consecutive applications then we could use tied-double-change-over designs, in selected cases, for n=2 and triple designs for n=3. For these latter two cases we will give generally applicable methods of constructing designs based on the work of Williams [28],[29]. In fact, the designs for n=3 are constructed similarly to the method used for triple designs. However, triple designs required at least (v-1)(v-2) units, and here we will often be able to use fewer units. E.g., when v-2 is divisible by 3 only (v-1)(v-2)/3 units are required.

Let n-1 be the maximum number of periods for which residual or interaction effects are assumed to be present, and let v be the number of treatments being compared.

**Definition 5.1:** Let n,r,s, and v be integers. An n-repeating design (n-R design) is defined by a set of s basic treatment series, each of length rv. If we consider the last n-1 treatments of each basic series to be applied initially as conditioning preceding the basic series, rv n-plets are included in each series of applications. For an n-R design these series have the following three properties:
(i) Each successive set of \( n \) treatments in a basic series includes each treatment at most once.

(ii) The \( rv \) \( n \)-plets occurring in each basic series are all distinct.

(iii) In the \( srv \) \( n \)-plets occurring in the basic series, all the different \( n \)-plets occur equally often.

In a given experiment \( p \) units are assigned each basic series and the total number of units is \( sp \). The basic series are continually repeated until the experiment ends. If the basic series is repeated the same number of times for each unit then 

\[
\lambda_i = \lim_{m \to \infty} \frac{m_i}{m} = \frac{1}{sp}, \quad i=1,2,\ldots,sp.
\]

Thus if the likelihood method of the preceding chapter is used to estimate the parameters of a conditional density, the information from each unit contributes equally.

Because of the asymptotic nature of the results of Chapter 4, we would like each \( n \)-plet to occur as often as possible. For a fixed amount of time available for an experiment this can be accomplished by making \( r \) small. Condition 3 of Chapter 4 requires the nonsingularity of the limiting variance matrix for each experimental unit, so that as the number of parameters increase, \( r \) may have to also. For these reasons we will consider construction of \( n-R \) designs for \( n=2 \), \( r=2 \), and for \( n=3 \), \( r \geq 3 \).

Before proceeding with the constructions we give a simple example of a 2-R design in order to clarify the definition.
Example 5.1: If \( v=3 \) we have a special case because only one basic series is required for a 2-R design. If \( V = \{0,1,2\} \) is the set of treatments, then a single basic series defining a 2-R design is \( 0 \ 1 \ 2 \ 0 \ 2 \ 1 \). Using 1 for conditioning, the series which gives our 2-pllets is \( \begin{array}{c|cccc} 1 & 0 & 1 & 2 & 0 & 2 & 1 \end{array} \), which includes all 6 different 2-pllets. The values of the parameters for this design are \( n=2, r=2, s=1, \) and \( v=3 \).

5.2 Construction of 2-Repeating Designs

Assume that each treatment has an effect for one period after its application so that we are interested in estimating at most one residual and one interaction effect for each application. Then \( n=2 \) and we will give methods of constructing 2-R designs with \( r=2 \). Different constructions will be given for \( v \) odd and \( v \) even, but they both use the designs balanced for a single preceding effect with an odd number of treatments, as given by Williams [28].

First consider the case where \( v \) is odd.

Theorem 5.2.1: For any odd integer \( v = 2p+1, p = 1,2,... \), there exists a 2-repeating design.

Proof: We first give a method for constructing such designs and then show that the properties of 2-R designs are satisfied. From Williams [28] there exist a pair of Latin squares of size \( v \) such that every 2-plet occurs twice. The difference set used to generate square 1 includes each odd member of the set \( V' = \{1,2,...,v-1\} \) exactly twice, and the difference set used to generate square 2 includes each even member of \( V' \) exactly twice. For square 1 the first row is

\[
0 \ 1 \ v-1 \ 2 \ v-2 \ 3 \ldots \frac{v-1}{2} \ \frac{v+1}{2},
\]

(5.2.1a)
and for square 2 the first row is

\[ 0 \quad 1 \quad 2 \quad \cdots \quad \frac{v+1}{2} \quad rac{v-1}{2} \quad (5.2.1b) \]

Letting row \((i,h)\) denote row \(h\) of square \(i\), \(i=1,2\), and \(h=1,2,\ldots,v\), we may write the \(j\)-th element of row \((1,1)\) as

\[ \frac{1-j}{2} \quad \text{if } j \text{ is odd} , \quad (5.2.2a) \]

\[ \frac{j}{2} \quad \text{if } j \text{ is even} , \quad (5.2.2b) \]

and the \(j\)-th element of row \((2,1)\) as

\[ \frac{v+j}{2} + p \quad \text{if } j \text{ is odd} , \quad (5.2.3a) \]

\[ -\frac{j}{2} \quad \text{if } j \text{ is even} . \quad (5.2.3b) \]

(All arithmetic operations are to be performed modulo \(v\).) By the method of generating the Latin squares we may also write the \(j\)-th element of row \((1,v)\) as

\[ \frac{1-j}{2} + v-1 \quad \text{if } j \text{ is odd} , \quad (5.2.4a) \]

\[ \frac{j}{2} + v-1 \quad \text{if } j \text{ is even} , \quad (5.2.4b) \]

and the \(j\)-th element of row \((2,v)\) as

\[ \frac{v+j}{2} + p + v-1 \quad \text{if } j \text{ is odd} , \quad (5.2.5a) \]

\[ -\frac{j}{2} + v-1 \quad \text{if } j \text{ is even} . \quad (5.2.5b) \]

From these two squares we construct a 2-R design with \(v-1 = 2p\) basic series. The first \(p\) basic series consists of row \((1,i)\) followed by row \((2,p+1 - i)\), \(i=1,2,\ldots,p\), and the second \(p\) consists of row \((1,i)\) followed by row \((2,p-i)\), \(i=p+1,p+2,\ldots,2p\). We now show that every 2-plet occurs exactly twice in the basic series.

By the method of selecting the 2 Latin squares, each 2-plet occurs in them twice. In forming our basic series each of the 2-plets in row
v of the two squares is omitted. To each of the basic series we have added two 2-plets, one where the row from the second square follows the row from the first and one where the first element of the row from square 1 is repeated. These added 2-plets must be the same as the 2-plets omitted from the final rows if we are to have a 2-R design.

For row (1,v) the v-th element is p, the elements in all other odd-numbered columns are \( >p \), and the elements in the even-numbered columns are \( <p \). This may be seen by substituting the numbers 0,1,...,v-1 for \( j \) in (5.2.4a-b). Thus row (1,v) has the following 2-plets:

(i) \( p \) is not followed by any other number;

(ii) \( i < p \) is followed by \( v-i-2 \); and

(iii) \( i > p \) is followed by \( v-i-1 \).

We show, for example, that (ii) holds. From (5.2.4b), \( i \) occurs in an even-numbered column, say \( j \), so \( i = \frac{j}{2} + v-1 \). Then it is followed by

\[
\frac{1-\left(j+1\right)}{2} + v-1 = v-1 - \frac{i}{2}.
\]

But \( \frac{j}{2} = i+1 - v \), so \( v-1 - \frac{j}{2} = 2v - i-2 = v-i-2 \) (mod \( v \), and (ii) holds.

For row (2,v) the odd and even positioning of the elements is not so apparent. It is easy to see that \( v-1 = 2p \) occurs in position 1. Beginning with the pair in positions 2 and 3, every 2-plet with initial element in an even-numbered position has been reversed from the corresponding pair in row (1,v). To prove this we must show that

(a) for \( j \) even, \( -\frac{j}{2} + v-1 = \frac{1-(j+1)}{2} + v-1 \) and

(b) for \( j \) odd, \( \frac{v+j}{2} + p+v-1 = \frac{i-1}{2} + v-1 \).

(a) is obvious. For (b) we must show that \( \frac{v+j}{2} + p = \frac{i-1}{2} \). But

\[
p = \frac{2p}{2} = \frac{v-1}{2},
\]

so that the LHS of this last equation is \( \frac{2v+j-1}{2} = \frac{j-1}{2} \).
By making use of these results it is easily shown that the following 2-plets occur in row \((2,v)\):

(iv) \(p-1\) is not followed by any other number ;

(v) \(i, 0 \leq i \leq p - 2\) and \(i = 2p\), is followed by \(v - i - 3\); and

(vi) \(i, p \leq i \leq 2p - 1\), is followed by \(v - i - 2\).

We show, for example, that (v) holds. From (5.2.5a) \(i, 0 \leq i \leq p - 2\) and \(i = 2p\), occurs in an odd-numbered column, say \(j (j<v)\), so
\[
i = \frac{v+i}{2} + p + v - 1 = \frac{i-1}{2} + v - 1,\]
using (b). It is followed by \(-\frac{(j+1)}{2} + v - 1\). But \(j = 2i - 2v+3\), so \(-\frac{(j+1)}{2} + v-1 = 2v - i - 3 = v - i - 3\), and (v) holds.

Now we show that the new 2-plets formed from the \(v\)-th element of the square 1 row and the 1st element of the square 2 row in each basic sequence are exactly those 2-plets that occur in row \((1,v)\). For series \(i, 1 \leq i \leq p\), row \((1,i)\) is followed by row \((2,p-i+1)\). Element \(v\) of row \((1,i)\) is \(\frac{1-v}{2} + i - 1 = p+i\) which is in the range \(p+1\) to \(2p\). Element 1 of row \((2,p-i+1)\) is \(\frac{v+1}{2} + p + (p-i+1) - 1 = p-i\). In row \((1,v)\), \(p+i, 1 \leq i \leq p\), is followed by \(v - (p+i) - 1 = (2p+1) - (p+i) - 1 = p-i\), from (iii). But this is the same as the 1st element of row \((2,p-i+1)\).

For series \(i, p+1 \leq i \leq 2p\), row \((1,i)\) is followed by row \((2,p-i)\). Element \(v\) of row \((1,i)\) is \(p+i\) which is in the range 0 to \(p-1\). Element 1 of row \((2,p-i)\) is \(\frac{v+1}{2} + p + (p-i) - 1 = p-i-1\). In row \((1,v)\), \(p+i, 0 \leq i \leq p-1\), is followed by \(v - (p+i) - 2 = p-i-1\), from (ii). But this is the same as the first element of row \((2,p-i)\). Thus the result has been shown.

In an analogous manner it can be shown that the new 2-plets formed from the \(v\)-th element of the square 2 row and the 1st element of the square 1 row in each basic series are exactly those \(p\) 2-plets occurring in row
(2, v). As a result, in our 2p basic series each 2-plet occurs exactly 2
times and we have in fact constructed a 2-repeating design. □

Example 5.2: Let \( p=4 \) and \( v=9 \). The two Latin squares used to construct
the 2-R design have the following initial rows. (See (5.2.1a-b)).

\[
\begin{align*}
0 & 1 & 8 & 2 & 7 & 3 & 6 & 4 & 5 \\
0 & 8 & 1 & 7 & 2 & 6 & 3 & 5 & 4
\end{align*}
\]

(5.2.6a) (5.2.6b)
The series are shown in Figure 5.1. □

Theorem 5.2.2: For any even integer \( v \) there exists a 2-repeating
design.

Proof: Again we consider the pair of Latin squares with initial rows
(5.2.1a-b), this time of size \( v-1 \). The elements of these initial rows
are calculated from (5.2.2a-b) and (5.2.3a-b). In order to construct a
2-R design for \( v \) even, we use the following procedure.

(i) Write down the pair of squares of size \( v-1 \).

(ii) Place a \( v-1 \) at the end of each row in the two squares, giving
two \((v-1) \times v\) rectangles.

(iii) The \( i \)-th basic series is defined by row \( i \) of the first rectangle
followed by row \( i-1 \) of the second, \( i=2,3,\ldots,v-1 \). For the
other series row 1 is followed by row \( v-1 \).

Recall that by the method of constructing the two Latin squares each
2-plet of the numbers 0,1,\ldots,v-2 occurs in them exactly twice. Since
each of these numbers also occurs as the initial element of one row in
each square and as the final element in some other row, it is apparent
that each 2-plet including \( v-1 \) occurs exactly twice in the design. It
remains to be shown that no 2-plet occurs twice in any basic series.

We consider two separate cases.
**Figure 5.1**

First p=4 rows

\[
\begin{align*}
0 &\quad 1 &\quad 8 &\quad 2 &\quad 7 &\quad 3 &\quad 6 &\quad 4 &\quad 5 \\
1 &\quad 2 &\quad 0 &\quad 3 &\quad 8 &\quad 4 &\quad 7 &\quad 5 &\quad 6 \\
2 &\quad 3 &\quad 1 &\quad 4 &\quad 0 &\quad 5 &\quad 8 &\quad 6 &\quad 7 \\
3 &\quad 4 &\quad 2 &\quad 5 &\quad 1 &\quad 6 &\quad 0 &\quad 7 &\quad 8
\end{align*}
\]

Second p=4 rows

\[
\begin{align*}
4 &\quad 5 &\quad 3 &\quad 6 &\quad 2 &\quad 7 &\quad 1 &\quad 8 &\quad 0 \\
5 &\quad 6 &\quad 4 &\quad 7 &\quad 3 &\quad 8 &\quad 2 &\quad 0 &\quad 1 \\
6 &\quad 7 &\quad 5 &\quad 8 &\quad 4 &\quad 0 &\quad 3 &\quad 1 &\quad 2 \\
7 &\quad 8 &\quad 6 &\quad 0 &\quad 5 &\quad 1 &\quad 4 &\quad 2 &\quad 3
\end{align*}
\]

Last row omitted in design

\[
\begin{align*}
8 &\quad 0 &\quad 7 &\quad 1 &\quad 5 &\quad 2 &\quad 5 &\quad 3 &\quad 4 \\
8 &\quad 7 &\quad 0 &\quad 6 &\quad 1 &\quad 5 &\quad 2 &\quad 4 &\quad 3
\end{align*}
\]

SQUARE 1

SQUARE 2

Arrows indicate rows of the second square which follow rows of the first in forming the basic sequences.
(a) The differences between adjacent elements in square 1 are odd and in square 2 are even. Thus no 2-plet which does not include v-1 is duplicated in a basic series.

(b) For the 2-plets including v-1 to occur singly we must have the initial elements of the two rows used for a basic sequence different, and the same for the final elements. The initial elements are different since row i of each square has initial element i-1. Now suppose that the (v-1)th elements of rows (1,i) and (2,i-1) are identical. Then we must have (mod v-1 = 2p+1)

\[ \frac{1-(v-1)}{2} + (i-1) = \frac{(v-1)+(v-1)}{2} + \frac{v-2}{2} + (i-2) \quad \text{(5.2.7a)} \]

using (5.2.2a) and (5.2.3a). This reduces to

\[ p + i = p + i - 2 \quad \text{(5.2.7b)} \]

which is impossible unless v-1 = 2. But v-1 is odd, and the 2-plets including v-1 in each basic series are unique. □

By the methods of Theorems 5.2.1 and 5.2.2 we have constructed 2-R designs for all v, where the values of the remaining parameters are r=2 and s = v-1.

Example 5.3: We demonstrate the construction of a 2-R design for v even. Let v=4, r=2, and s = v-1 = 3.

\[
\begin{array}{ccc}
3 & 0 & 1 & 2 & 3 \\
3 & 1 & 2 & 0 & 3 \\
3 & 2 & 0 & 1 & 3 \\
\end{array}
\]

(a) Square 1  (b) Square 2  (b)

\[ \text{Rectangle 1} \quad \text{Rectangle 2} \]
(a) Conditioning application

(b) Final elements $v-1 = 3$ added to end of each row.

The arrows indicate which row of rectangle 2 is to follow a given row of rectangle 1 to form the basic series.

5.3 Construction of 3-Repeating Designs

Now assume that each treatment has a residual and/or interaction effect that is present two periods after application, in addition to its direct effect. For a design to be balanced for every pair of residual effects and interaction effects it must have every 3-plet occurring equally often. In order to construct a class of 3-repeating designs we will again make use of the designs balanced for pairs of residual effects as given by Williams [28],[29]. If $v$ treatments are being considered these designs consist of $v-1$ Latin squares of size $v$. The $i$-th row, $i=1,2,...,v$, of each square has initial element $i-1$, and we will use only those designs for which the final element of the first row of each square is different. Then the final element of the $i$-th row of each square is different. Recall from Chapter 1 that every 3-plet occurs in these designs exactly once and that every 2-plet occurs at the beginning of some row and the end of some other row exactly once. Whether the designs used are orthogonal sets of Latin squares or constructed from difference sets does not matter.

First consider the case where $v-2$ is divisible by 3. We shall give a method for constructing 3-R designs for this case which can be extended to cover the cases when $v-2$ is not divisible by 3.
Theorem 5.3.1: A 3-repeating design with \( r=3 \) exists for \( v-2 \) divisible by 3, where \( v \) is the number of treatments being considered.

Proof: To define the basic series, first take a design for \( v-1 \) treatments balanced for pairs of residual effects, as described above. Add a final column with all entries \( v-1 \) to each square, and write down side-by-side the resulting \( v-2 \) rectangles of size \((v-1)\times v\). In each row of this \((v-1)\times v(v-2)\) array let the first \( 3v \) entries, second \( 3v \) entries, etc., constitute the basic series. The last 2 elements of each basic series are placed at the beginning of that series as conditioning in order to determine the \( rv \) 3-plets which occur. It is obvious from the discussion preceding this theorem that every 3-plet not containing \( v-1 \) occurs exactly once in some basic series. We show that every 3-plet containing \( v-1 \) occurs exactly once.

(a) Since every 2-plet occurs at the beginning of some row in the original squares, every 3-plet with \( v-1 \) as the initial element occurs exactly once. Similarly, every 3-plet with \( v-1 \) as the third element occurs exactly once.

(b) The first row of each of the original squares begins with 0 and ends with one of the elements \( 1, 2, \ldots, v-2 \), where the last element for each square is different. Because of the addition of the final column of \((v-1)\)’s to each square, each 3-plet of the form \( \{j \ (v-1) \ 0\} \), \( j=1, 2, \ldots, v-2 \), occurs once in the basic series constructed from the first rows of the rectangles. Similarly, for row \( i \), \( i=1, 2, \ldots, v-1 \), the 3-plets \( \{j \ (v-1) \ i\} \), \( i \neq j \), each occur once in the basic series constructed from that row. Thus every 3-plet with \( v-1 \) as the second element occurs exactly once in the design. Since each 3-plet occurs exactly once in some basic series, no 3-plet is repeated in a given series.
Example 5.4: Let $v=5$ and $v-2 = 3$ be divisible by 3. Each row of
the following rectangle represents a basic series and its conditioning
treatments.

\[
\begin{array}{ccccccccc}
2 & 4 & 0 & 1 & 2 & 3 & 4 & 0 & 2 & 3 & 1 & 4 & 0 & 3 & 1 & 2 & 4 \\
3 & 4 & 1 & 0 & 3 & 2 & 4 & 1 & 3 & 2 & 0 & 4 & 1 & 2 & 0 & 3 & 4 \\
0 & 4 & 2 & 3 & 0 & 1 & 4 & 2 & 0 & 1 & 3 & 4 & 2 & 1 & 3 & 0 & 4 \\
1 & 4 & 3 & 2 & 1 & 0 & 4 & 3 & 1 & 0 & 2 & 4 & 3 & 0 & 2 & 1 & 4 \\
\end{array}
\]

(a) Conditioning treatments
(b) Columns of $v-1 = 4$.

Note that the corresponding rows of each square have the same initial
element. Because of this and the fact that the 3 Latin squares are
mutually orthogonal, no two last elements of the same row in different
squares can be identical.

If $v-2$ is not divisible by 3 we can use the same type of procedure
to construct 3-R designs, but we must increase either $r$ or $s$. This is
because the number of elements in any row of the derived $(v-1)xv(v-2)$
rectangle is not divisible by 3$v$. We cannot just go to the next row
when we come to the end of a given row in constructing the basic series.
In doing so we might repeat a 3-plet with $v-1$ in the second position and
leave out some other 3-plet. There are 2 different procedures which over-
come this problem.

(i) If $v-2$ is divisible by the integer $q\neq 3$, use the first $qv$
entries, the second $qv$ entries, etc., of each row of the
derived $(v-1)xv(v-2)$ rectangle as the basic series. Then
\[ r = q \text{ and } s = \frac{(v-2)(v-1)}{q}. \]

(ii) If \( s = (v-1)(v-2) \) experimental units are available and \( r = 3 \) is to be used, write down the derived \((v-1)\times(v-2)\) rectangle 3 times to form an extended \((v-1)\times3v(v-3)\) rectangle. Now use each successive set of \(3v\) elements in a row as a basic series.

That (i) and (ii) in fact give 3-repeating designs is clear from the proof of Theorem 5.3.1.
CHAPTER VI

SERIAL FACTORIAL DESIGNS

6.1 The Serial Factorial Model and Notation

Instead of \( v \) distinct treatments, we now consider the case of having \( w \) different factors, each of which may occur at several different quantitative levels. A treatment combination then consists of the simultaneous application of one level of each factor. When one unit is assigned to each possible combination we have a factorial experiment, and when only a certain subset of the combinations is used we have a fractional factorial. These designs are used to estimate the main effects of each factor level and the interaction effects between levels of different factors.

In some situations an experimenter must decide on a single treatment combination to be used repeatedly in the future. Then the experimenter might be more interested in the direct and residual effects of each factor level in a series of treatments rather than in interactions between factors in a single application. For this case we would like to be able to arrange series of applications for several experimental units that would allow us to estimate all the direct and residual effects with equal variances within, and covariances between, classes of estimates (e.g., direct effects, first residuals), and at the same time use as few combinations and units as is practicable.

A natural approach to this problem is to use fractional factorials and apply the combinations of each fraction to a single unit in a speci-
fied order. Here we will consider designs of this type when each of the \( \omega \) factors has 2 levels, the \( 2^\omega \) series, for \( \omega = 2,3,4,5,6 \). These designs are closely related to those described by Patterson [21] and called serial factorial designs. Here we shall use this same title for our designs, and we now describe this serial factorial model.

Denote the factors by capital letters and the entire set by \( \Lambda = \{A,B,C,\ldots,\Omega\} \), \( |\Lambda| = \omega \), and each factor's two levels by the subscripts 0 and 1. Also, letting the subscripts \( d \) and \( r \) represent direct and residual effects respectively, each factor \( \Gamma \in \Lambda \) has four effects denoted by \( \Gamma_{d0}, \Gamma_{d1}, \Gamma_{r0}, \text{ and } \Gamma_{r1} \). We will use subscripted lower case letters to represent actual factor levels being applied. E.g., \( a_{d0} \) denotes factor \( A \) occurring at direct effect level 0 and \( b_{r1} \) that factor \( B \) occurred at level 1 in the previous application, and thus for the present application residual effect 1 of \( B \) occurs.

Rather than estimating both \( \Gamma_{d0} \) and \( \Gamma_{d1} \) for each \( \Gamma \in \Lambda \), we are interested only in any difference between them. Thus to reduce the number of parameters to be estimated we use the restrictions

\[
\Gamma_{d0} + \Gamma_{d1} = 0, \quad \forall \Gamma \in \Lambda. \tag{6.1.1a}
\]

Similarly for residual effects we have

\[
\Gamma_{r0} + \Gamma_{r1} = 0, \quad \forall \Gamma \in \Lambda. \tag{6.1.1b}
\]

This leaves us with \( 2\omega \) independent factor effects to be estimated.

We will be interested in testing hypotheses of the form \( H_0: \Gamma_{r1} = 0, \forall \Gamma \in \Lambda \), and \( H_2: \Gamma_{r1} + \Gamma_{d1} = 0, \forall \Gamma \in \Lambda \).

For the \( k \)-th observation, \( k=1,2,\ldots,n \), on unit \( j, j=1,2,\ldots,m \), we will represent the levels present by a vector \( \alpha_{jk} \) consisting of
0's and 1's and having the form
\[ \alpha_{jk} = (a^d(jk), a^r(jk), b^d(jk), \ldots, o^d(jk), o^r(jk)) \]
where
\[ \gamma^d(jk) = \begin{cases} 1 & \text{if } \gamma_{d1} \text{ is present} \\ 0 & \text{if } \gamma_{d0} \text{ is present} \end{cases} \]
and a similar definition holds for \( \gamma^r(jk) \). We write the response model as
\[ y_{jk} = \mu + \delta_j + \sum_{\Gamma \epsilon \Lambda} N_{d1} \Gamma_{dj1} + \sum_{\Gamma \epsilon \Lambda} N_{r1} \Gamma_{rj1} + \epsilon_{jk} \]
where \( \mu \) is the overall mean,
\( \delta_j \) is the effect of unit \( j, j=1,2,\ldots,m \),
\( \Gamma_{dj} = \begin{cases} -1 & \text{if } \gamma^d(jk) = 0 , \quad \gamma, \Gamma \epsilon \Lambda \\ 1 & \text{if } \gamma^d(jk) = 1 \end{cases} \)
\( \Gamma_{d1} \) is the direct effect of factor \( \Gamma \) at level 1, \( \Gamma \epsilon \Lambda \),
\( \Gamma_{rj} = \begin{cases} -1 & \text{if } \gamma^r(jk) = 0 , \quad \gamma, \Gamma \epsilon \Lambda \\ 1 & \text{if } \gamma^r(jk) = 1 \end{cases} \)
\( \Gamma_{r1} \) is the residual effect of factor \( \Gamma \) at level 1, \( \Gamma \epsilon \Lambda \), and the
\( \epsilon_{jk} \) are all independent with common distribution \( \mathcal{N}(0,\sigma^2) \).

There are \( 2\omega + m+1 \) effects to be estimated in addition to the common variance \( \sigma^2 \).

**Example 6.1:** As an example of a serial factorial design consider the simple case where \( \omega = 2 \). Each unit will receive all four treatment combinations, and there must be at least two units as the number of effects to be estimated is \( 5+m \). For \( m = 2 \) units we may use the following design.
Each row for each unit represents an application of a treatment combination and its corresponding $\alpha_{jk}$. The first application on each unit is to be considered as a conditioning treatment and no observation is made until after the second application. In this design there are 8 observations and 7 effects. To allow sufficient degrees of freedom for the estimation of $\sigma^2$ we could repeat these series on two additional units.

Note that in the example $a^{d(1k)} + b^{r(1k)} = 1, \forall k$. I.e., $a_{d1}$ always occurs with $b_{r0}$ and $a_{d0}$ with $b_{r1}$ on unit 1.

**Definition 6.1:** $\gamma_{ig}$ and $\phi_{i'g'}$ are **unit confounded** on unit $j$ if $\gamma_{i(jk)} + \phi_{i'(jk)} = g + g'$ for all $k$, where $i, i' = r$ or $d$ and $g, g' = 0$ or 1.

**Definition 6.2:** $\gamma_{ig}$ and $\phi_{i'g'}$ are **design confounded** $q$ times if $\gamma_{ig}$ is unit confounded with $\phi_{i'g'}$ $q$ more times than it is unit confounded with $\phi_{i'g'}$, where $i, i' = r$ or $d$, $g, g' = 0$ or 1, and

$$
\bar{g}' = \begin{cases} 
0 & \text{if } g' = 1 \\
1 & \text{if } g' = 0
\end{cases}
$$

If the unit confounding occurs equally often with each level we say there is no design confounding. (If two effects are unit confounded, they are
not separately estimable using observations on that unit only. Two
design confounded effects are separately estimable using all observa-
tions unless they are unit confounded on every unit. Design confound-
ing indicates non-zero covariance between the estimated effects.)
Thus on unit 1 \( a_{d1} \) is unit confounded with \( b_{r0} \) (and at the same time
\( a_{d0} \) is unit confounded with \( b_{r1} \), but since this confounding always
occurs in pairs we need only mention one of the pair), and on unit 2
\( a_{d1} \) is unit confounded with \( b_{r1} \), so there is no design confounding
between the direct level of \( A \) and the residual level of \( B \). When con-
structing serial factorial designs we want the amount of design con-
 founding to be as small as possible in order to keep the covariance be-
tween estimated effects small.

Example 6.1 (continued): Using the notation of (6.1.3) we write the
observational equations of the given design as

\[
\begin{align*}
y_{11} &= \mu + \delta_1 + A_{d1} + B_{d1} + A_{r1} - B_{r1} + \varepsilon_{11} \\
y_{12} &= \mu + \delta_1 - A_{d1} + B_{d1} + A_{r1} + B_{r1} + \varepsilon_{12} \\
&\vdots \\
y_{24} &= \mu + \delta_2 - A_{d1} + B_{d1} - A_{r1} - B_{r1} + \varepsilon_{24}.
\end{align*}
\]  

(6.1.4)

Serial factorial designs for \( 2^\omega, \omega = 3,4,5,6 \), will be given in the
following sections. In the final sections we give a general method for
solving the normal equations, as well as some special results for the
given designs with \( \omega \) even. All the designs will have four observations
per experimental unit.
6.2 $2^3$ Designs

There are $7+m$ effects to be estimated, so we need at least three series of treatments. Since by the method of selecting factorial fractions there will be unit confounding for each series, some multiple of three units is required for the design confounding to be in some sense "balanced." Thus the simplest possible design with balance requires only three units. In addition to this overall balance, we want to select the fractions so that for each unit both levels of each factor will occur as direct and residual effects twice each. Then treatment effects will be orthogonal to unit effects. We will demonstrate construction of these designs with two different types of balance, but first we give another definition.

Definition 6.3: If on a specific unit a factor is applied at levels 0 and 1 alternately, then on that unit it is an alternating factor.

For the first type of $2^3$ serial factorial design only design confounding between direct levels of one factor and residual levels of another occur. The procedure followed is:

1) each unit receives a 1/2 fraction defined by
\[ a^d(jk) + b^d(jk) + c^d(jk) = \rho, \]  
(6.1.5)

where $\rho = 0$ or 1, $j=$unit number, and $k=1,2,3,4$; and

2) each unit has a different alternating factor.

(For simplicity, in the sequel we will drop the superscripts in equations of the form (6.1.5) and write $a+b+c = \rho$.)

Example 6.2: For each unit we give the defining sum, treatment series, and unit confounding.
<table>
<thead>
<tr>
<th>Unit 1</th>
<th>Unit 2</th>
<th>Unit 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>sum</td>
<td>a + b + c = 0</td>
<td>a + b + c = 1</td>
</tr>
<tr>
<td>factor</td>
<td>A B C</td>
<td>A B C</td>
</tr>
<tr>
<td>series</td>
<td>0 0 0</td>
<td>1 0 0</td>
</tr>
<tr>
<td></td>
<td>1 1 0</td>
<td>0 1 0</td>
</tr>
<tr>
<td></td>
<td>1 0 1</td>
<td>1 1 1</td>
</tr>
<tr>
<td></td>
<td>0 1 1</td>
<td>0 0 1</td>
</tr>
<tr>
<td></td>
<td>0 0 0</td>
<td>1 0 0</td>
</tr>
<tr>
<td>confounding</td>
<td>( a_d l - c_r 0 )</td>
<td>( a_d l - a_r 0 )</td>
</tr>
<tr>
<td></td>
<td>( b_d l - b_r 0 )</td>
<td>( b_d l - c_r 0 )</td>
</tr>
<tr>
<td></td>
<td>( c_d l - a_r 0 )</td>
<td>( c_d l - b_r l )</td>
</tr>
</tbody>
</table>

The alternating factors for the three units are B for unit 1, A for unit 2, and C for unit 3. Note that each direct level of each factor is design confounded once with one residual level of each factor, and in this sense the design is balanced.

For the second type of \( 2^3 \) design different factors have design confounding between the same type of effects, i.e., direct with direct and residual with residual. Each factor is also confounded between its direct and residual effects. Again the method of construction involves two steps.

1) Each unit receives a different \( 1/2 \) fraction, the defining sums being \( a+b = \rho_1 \), \( b+c = \rho_2 \), and \( a+c = \rho_3 \) respectively, where \( \rho_1, \rho_2, \rho_3 = 0 \) or 1.

2) The factor not occurring in the defining sum is taken as the alternating factor.

Example 6.3: In this design the alternating factors are C for unit 1, A for unit 2, and B for unit 3, as given in (2).
<table>
<thead>
<tr>
<th>Unit 1</th>
<th>Unit 2</th>
<th>Unit 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>sum</td>
<td>a + b = 0</td>
<td>b + c = 1</td>
</tr>
<tr>
<td>factor</td>
<td>A B C</td>
<td>A B C</td>
</tr>
<tr>
<td>series</td>
<td>0 0 1</td>
<td>1 0 1</td>
</tr>
<tr>
<td></td>
<td>0 0 0</td>
<td>0 0 1</td>
</tr>
<tr>
<td></td>
<td>1 1 1</td>
<td>1 1 0</td>
</tr>
<tr>
<td></td>
<td>1 1 0</td>
<td>0 1 0</td>
</tr>
<tr>
<td></td>
<td>0 0 1</td>
<td>1 0 1</td>
</tr>
</tbody>
</table>

### Confounding

<table>
<thead>
<tr>
<th></th>
<th>d1 - b1</th>
<th>d1 - r0</th>
<th>d1 - c0</th>
</tr>
</thead>
<tbody>
<tr>
<td>a1</td>
<td>a1</td>
<td>a1</td>
<td></td>
</tr>
<tr>
<td>b1</td>
<td>b1</td>
<td>b1</td>
<td></td>
</tr>
<tr>
<td>c1</td>
<td>c1</td>
<td>c1</td>
<td></td>
</tr>
<tr>
<td>d1</td>
<td>d1</td>
<td>d1</td>
<td></td>
</tr>
</tbody>
</table>

Whether an experimenter uses a design of the first or second type depends on the covariance pattern between effect estimates that he is willing to accept. If residual effects are of little interest the first type of design would be better, but if it is of primary interest to test for the absence of residual effects the second type would be preferred.

### 6.3 Design

There are 9+4 parameters to be estimated, so we require at least 4 fractions of four treatments each. As in the 2^3 designs, at least one factor must alternate for each unit. Since each time a factor alternates its direct and residual effects are confounded, we want each factor to alternate equally often in the design. These conditions can be satisfied by a design with a different alternating factor on each of four units. We give a six step procedure for a class of 2^4 designs.
1) Separate the factors into two pairs \( \Pi_1, \Pi_2 \) and \( \psi_1, \psi_2 \), say. Each of these pairs will be used for defining sums on two units.

2) For unit 1 use the defining sum \( \pi_1 + \pi_2 = 0 \) and for unit two use \( \pi_1 + \pi_2 = 1 \).

3) The sums in (2) leave us with a 1/2 fraction for each of the first two units, so further restrictions are needed to get 1/4 fractions. For both units use the same three-factor sum \( \psi_1 + \psi_2 + \pi_1 = \rho \), where \( i = 1 \) or 2 and \( \rho = 0 \) or 1.

4) Order the treatments of the 1/4 fractions defined by (2) and (3) such that \( \psi_1 \) and \( \psi_2 \) each alternate in one of the two units.

5) For units 3 and 4 reverse the roles of \( \Pi_1, \Pi_2 \) and \( \psi_1, \psi_2 \). The three defining sums for these units will be \( \psi_1 + \psi_2 = 0, \psi_1 + \psi_2 = 1, \) and \( \pi_1 + \pi_2 + \psi_1 = \rho, \ i = 1 \) or 2 and \( \rho = 0 \) or 1.

6) In addition to \( \Pi_1 \) and \( \Pi_2 \) each being an alternating factor on one of the last two units, we must select the order of application to reduce the amount of design confounding. After ordering the treatments for the first two units there will be unit confounding between direct and residual levels of different factors. Arrange the treatment order on units 3 and 4 so that the opposite unit confounding between direct and residual levels of different factors occurs, and in this way all design confounding between different factors is eliminated.

As a result of this procedure the only design confounding is once between the direct and residual levels of each factor.

**Example 6.4:** Let the factor pairs be \( A, C \) and \( B, D \).
The design confounding is given by the last line as a consequence of (6). □

6.4 $2^5$ Designs

At least 4 fractions of size 4 are required in order to estimate the 11+m parameters of the $2^5$ factorial. As in the previous designs each fraction will have at least one alternating factor, so a multiple of five fractions is needed for each factor to alternate equally often and the design to remain balanced. We give a procedure for constructing $2^5$ designs with two alternating factors on each of five units, thus having each factor alternate twice. Each 1/8 fraction will be defined by 3 independent two-factor sums.
1) From the letters \{A,B,C,D,E\} form five sets of triplets such that every letter occurs in three triplets and each pair of letters occurs in the same triplet either once or twice.

2) There are 3 pairs of letters in each triplet. Select two of these pairs to be used in defining the fraction associated with the triplet. Assign to the sum of the levels of the two factors in each pair the value 0 or 1. Once these values are given the value assigned to the sum of the levels of the remaining pair is determined. If the first two values are the same the last is 0; if they are different it is 1. Of the 10 distinct pairs available from 5 letters, 5 will occur twice and 5 once in these two-factor sums.

Example 6.5: We will demonstrate the method described in (2). Let the triplet be ABC, and the three pairs are AB, BC, AC. Select the pairs AB and BC to be independent and use the sums \(a+b = 1\) and \(b+c = 0\) in defining the fraction. Then the restriction \(a+c = 1\) will also hold for each treatment combination. (See also Example 6.6.)

Definition 6.4: For each triplet in (1) there are two letters from \{A,B,C,D,E\} which are not present. We will call this pair the associate pair of the triplet.

3) Assign to each triplet its associate pair. These pairs will be the same as those in (1) that occur together twice in the same triplet. (To see this, let AB be such a pair. There are 3 remaining triplets, one of which contains A but not B, and one which contains B but not A, by (1). This leaves one triplet containing neither A nor B, and its associate pair must be AB.) Now assign values as in (2) to the sum of the levels of the associate pair. Do this in a way
such that the three values assigned to this pair in (2) and (3) are not all the same.

4) We now have 5 sets of 4 two-factor sums, three of the four being independent. Each of these sets defines a 1/8 fraction, and the treatments of each fraction are ordered so that the associate pair will be alternating factors.

As a result of (2) and (3), each factor's direct level is design confounded once with each other factor's direct level and similarly for residual levels. The only design confounding between direct and residual levels occurs with the alternating pairs. From (4) each factor's direct level is design confounded once with the residual level of two other factors and twice with its own residual. In this sense the design is balanced.

Once all the two-factor sums of the design have been written down the design confounding is determined in the following manner.

i) For the two factor sums that occur once.

If the assigned value is 0, direct level 1 of the two factors are confounded, and similarly for residual levels. If the assigned value is 1, direct level 1 of a factor is confounded with direct level 0 of the other, and similarly for residual levels.

ii) For the two-factor sums that occur thrice.

Confounding between direct levels and between residual levels is determined as in (i), except that the value assigned twice is used rather than the single assigned value.

Confounding between direct and residual levels is determined using the value assigned to the sum of two factors when they occur as associate pairs and the reverse of the method for finding direct with direct
confounding in (i). (That is, if the sum is 1, direct level 1 of a factor is confounded with residual level 1 of the other factor, and if the sum is 0, direct level 1 of a factor is confounded with residual level 0 of the other.)

Example 6.6: We give the step by step construction of a \(2^5\) serial factorial design.

1) Let the five triplets be ABC, ABD, ACE, BDE, and CDE.

2) For each triplet the three two-factor sums are given, the first two of which are taken to be independent.

<table>
<thead>
<tr>
<th>triplet</th>
<th>ABC</th>
<th>ABD</th>
<th>ACE</th>
<th>BDE</th>
<th>CDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>sums</td>
<td>a+b=0</td>
<td>a+b=1</td>
<td>a+c=0</td>
<td>b+d=0</td>
<td>c+d=1</td>
</tr>
<tr>
<td></td>
<td>b+c=1</td>
<td>b+d=1</td>
<td>c+e=1</td>
<td>d+e=0</td>
<td>d+e=1</td>
</tr>
<tr>
<td></td>
<td>a+c=1</td>
<td>a+d=0</td>
<td>a+e=1</td>
<td>b+e=0</td>
<td>c+e=0</td>
</tr>
</tbody>
</table>

Note that the third sum for each triplet can be gotten by adding the first two (mod 2).

3) Associate pairs and sums.

<table>
<thead>
<tr>
<th>pair</th>
<th>DE</th>
<th>CE</th>
<th>BD</th>
<th>AC</th>
<th>AB</th>
</tr>
</thead>
<tbody>
<tr>
<td>sum</td>
<td>d+e=0</td>
<td>c+e=1</td>
<td>b+d=0</td>
<td>a+c=1</td>
<td>a+b=0</td>
</tr>
</tbody>
</table>

4) Treatment series with associate pairs alternating.

<table>
<thead>
<tr>
<th>factor</th>
<th>Unit 1</th>
<th>Unit 2</th>
<th>Unit 3</th>
<th>Unit 4</th>
<th>Unit 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>series</td>
<td>A B C D E</td>
<td>A B C D E</td>
<td>A B C D E</td>
<td>A B C D E</td>
<td>A B C D E</td>
</tr>
<tr>
<td></td>
<td>1 1 0 0 0</td>
<td>1 0 0 1 1</td>
<td>0 1 0 1 1</td>
<td>1 0 0 0 0</td>
<td>0 0 0 1 0</td>
</tr>
<tr>
<td></td>
<td>1 1 0 1 1</td>
<td>1 0 1 1 0</td>
<td>0 0 0 0 1</td>
<td>0 0 1 0 0</td>
<td>1 1 1 0 1</td>
</tr>
<tr>
<td></td>
<td>0 0 1 0 0</td>
<td>0 1 0 0 1</td>
<td>1 1 1 1 0</td>
<td>1 0 1 1 0</td>
<td>1 1 0 1 0</td>
</tr>
<tr>
<td></td>
<td>0 0 1 1 1</td>
<td>0 1 1 0 0</td>
<td>1 0 1 0 0</td>
<td>0 1 1 0 1</td>
<td>1 0 1 0 1</td>
</tr>
<tr>
<td></td>
<td>1 1 0 0 0</td>
<td>1 0 0 1 1</td>
<td>0 1 0 1 1</td>
<td>1 0 0 0 0</td>
<td>0 0 0 1 0</td>
</tr>
</tbody>
</table>
The five pairs that occur once in (2) and (3) and their assigned values are:

\[ \text{AD-0, AE-1, BC-1, BE-0, CD-1}. \]

The five pairs that occur thrice in (2) and (3) and the values assigned to them twice are:

\[ \text{AB-0, AC-1, BD-0, CE-1, DE-0}. \]

Using (i) and (ii) on these results we get the design confounding between direct levels. (Confounding between residual levels is the same.)

\[
\begin{align*}
\text{ad}_1-b_1, \quad \text{ad}_1-c_0, \quad \text{ad}_1-d_1, \quad \text{ad}_1-e_0 \\
\text{bd}_1-c_0, \quad \text{bd}_1-d_1, \quad \text{bd}_1-e_1 \\
\text{cd}_1-d_0, \quad \text{cd}_1-e_0 \\
\text{dd}_1-e_1.
\end{align*}
\]

The design confounding between direct and residual levels is derived from (3) and (ii).

\[
\begin{align*}
\text{ad}_1-b_0, \quad \text{bd}_1-a_0, \quad \text{ad}_1-c_1, \quad \text{cd}_1-a_1, \quad \text{bd}_1-d_0, \quad \text{dd}_1-b_0 \\
\text{dd}_1-e_0, \quad \text{ed}_1-d_0, \quad \text{cd}_1-e_1, \quad \text{ed}_1-c_1.
\end{align*}
\]

If the experimenter has 5 additional experimental units available and wants to eliminate design confounding between different factors, the same two-factor sums as in (2) and (3) may be used to define the new fractions, but with all the assigned values switched. (I.e., 1 replacing 0 and 0 replacing 1.) In this new design each factor's direct level is design confounded 4 times with its own residual, but all other design confounding is eliminated.
6.5 $2^6$ Designs

There are $13 + m$ parameters to be estimated, so at least 5 fractions of size 4 are required. If each unit has two alternating factors, as in the $2^5$ designs, the number of units must be a multiple of 3. Thus 6 is the fewest number of units we can use for a $2^6$ serial factorial. We give a procedure for constructing these designs on 6 units, two alternating factors per unit, such that the only design confounding is twice between each factor's direct and residual levels.

1) From the letters \{A, B, C, D, E, F\} select three disjoint pairs to be used in two-factor sums for defining fractions.

2) Select one letter from each pair in (1). These three letters will be used in three-factor sums.

3) Divide the units into 3 groups of 2 units each. Each group uses a different pair from (1) as alternating factors on its two units. The two-factor sum for this pair is assigned the value 0 on one of the units of the group and 1 on the other.

4) Each pair from (1) occurs in two groups in which its factors are not alternating. In one of these groups its two-factor sum is assigned the value 0 and in the other 1.

5) The three-factor sum is used on every unit for defining the fraction and is assigned a different value for each unit of a group.

6) On each unit the direct levels of each non-alternating pair are unit confounded with the residual levels of the other non-alternating pair. The order of treatment application to the two units of a group is arranged so that this confounding is for opposite levels on the different units. (I.e., if $\pi_{d1}$ is confounded with $\psi_{r0}$ on one unit, then it is confounded with $\psi_{r1}$ on the other.)
7) As a result of (3-6) the only design confounding is twice between the direct and residual levels of each factor.

**Example 6.7:** Choose as the three disjoint pairs AB, CE, and DF, and choose as the triplet of (2) BCF. For groups 1,2, and 3, let the alternating pair of factors be CE, DF, and AB, respectively. We give the groups and the sums that define the fractions for each unit.

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit 1</td>
<td>Unit 2</td>
</tr>
<tr>
<td>a+b=1</td>
<td>a+b=1</td>
</tr>
<tr>
<td>c+e=0</td>
<td>c+e=1</td>
</tr>
<tr>
<td>d+f=1</td>
<td>d+f=1</td>
</tr>
<tr>
<td>b+c+f=1</td>
<td>b+c+f=0</td>
</tr>
</tbody>
</table>

**Group 3**

<table>
<thead>
<tr>
<th>Unit 5</th>
<th>Unit 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>a+b=0</td>
<td>a+b=1</td>
</tr>
<tr>
<td>c+e=0</td>
<td>c+e=0</td>
</tr>
<tr>
<td>d+f=0</td>
<td>d+f=0</td>
</tr>
<tr>
<td>b+c+f=0</td>
<td>b+c+f=1</td>
</tr>
</tbody>
</table>

Note how the values assigned to the sums agree with (3-5). In Figure 6.1 are the treatment series derived from the above sums. The only design confounding is twice between direct level 1 of each factor and its residual level 0.
### Figure 6.1

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Unit 1</strong></td>
<td><strong>Unit 3</strong></td>
<td><strong>Unit 5</strong></td>
</tr>
<tr>
<td>A B C D E F</td>
<td>A B C D E F</td>
<td>A B C D E F</td>
</tr>
<tr>
<td>0 1 0 1 0 0</td>
<td>1 1 0 1 1 1</td>
<td>1 1 1 0 1 0</td>
</tr>
<tr>
<td>0 1 1 0 1 1</td>
<td>0 0 0 0 1 0</td>
<td>0 0 1 1 1 1</td>
</tr>
<tr>
<td>1 0 0 0 0 1</td>
<td>0 0 1 1 0 1</td>
<td>1 1 0 1 0 1</td>
</tr>
<tr>
<td>1 0 1 1 1 0</td>
<td>1 1 1 0 0 0</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>0 1 0 1 0 0</td>
<td>1 1 0 1 1 1</td>
<td>1 1 1 0 1 0</td>
</tr>
</tbody>
</table>

| **Unit 2** | **Unit 4** | **Unit 6** |
| A B C D E F | A B C D E F | A B C D E F |
| 0 1 0 0 1 1 | 1 1 1 0 0 1 | 1 0 1 1 1 1 |
| 0 1 1 1 0 0 | 0 0 1 1 0 0 | 0 1 1 0 1 0 |
| 1 0 0 1 1 0 | 0 0 0 0 1 1 | 1 0 0 0 0 0 |
| 1 0 1 0 0 1 | 1 1 0 1 1 0 | 0 1 0 1 0 1 |
| 0 1 0 0 1 1 | 1 1 1 0 0 1 | 1 0 1 1 1 1 |

### 6.6 General Analysis of $2^k$ Designs

For the designs given in the previous sections each experimental unit receives a series of five treatment combinations. The first application is considered only as conditioning, and observations are made after each of the succeeding applications so that there are 4 observations per unit. Because each unit receives a defined fraction of the entire factorial, each level of direct and residual effects for each factor affects 2 observations per unit. We used $\omega$ units in each of the given $2^k$ designs, but they are not all practical because of the small number of degrees of freedom available for estimating $\sigma^2$. This drawback may be
overcome by replicating a design a certain number of times, otherwise
extending the design as suggested for the \(2^5\) designs, or by constructing
more complicated designs. Because of these possible extensions we give
the analysis for the general case of \(n\) observations on each of \(m\) units.

From model (6.1.3) it is seen that there are \(2\omega+m+1\) independent
effects to be estimated: \(\mu, \delta_j, j=1,2,\ldots,m, \Gamma_{d1}, \) and \(\Gamma_{r1}, \Gamma \in \Lambda\). In
order to rewrite (6.1.3) in matrix form we relabel the observations.
Those made on unit \(j\) will be denoted by \(Y_{n(j-1)+1}, Y_{n(j-1)+2},\ldots,\)
\(Y_{nj}\), and there are a total of \(N = mn\) observations represented by the
vector
\[
\mathbf{y} = (y_1, y_2, \ldots, y_N)'
\]  
(6.6.1)

Now rewrite (6.1.3) as
\[
\begin{align*}
\mathbf{y} &= \mathbf{I_N} \mu + G' \delta + D' \Gamma + R' \Phi + \epsilon \\
\end{align*}
\]  
(6.6.2)

where \(G'\) is an \(N \times m\) matrix with elements \(g'_{ij}\) given by
\[
g'_{ij} = \begin{cases} 
1 & \text{if observation } i \text{ is made on unit } j \\
0 & \text{otherwise} 
\end{cases}
\]

\(\delta = (\delta_1, \delta_2, \ldots, \delta_m)'\) is the \(m\)-vector of unit effects,

\(D'\) is an \(N \times \omega\) matrix with elements \(d'_{ik}\) given by
\[
d'_{ik} = \begin{cases} 
1 & \text{if the } k\text{-th factor occurs at direct level } 1 \text{ for observation } i \\
-1 & \text{if it occurs at direct level } 0 
\end{cases}
\]

\(\Gamma = (A_{d1}, B_{d1}, \ldots, \Omega_{d1})'\) is the \(\omega\)-vector of direct level 1 effects,

\(R'\) is an \(N \times \omega\) matrix with elements \(r'_{ik}\) given by
\[
\begin{aligned}
\mathbf{r}_{ik} &= \begin{cases} 
1 & \text{if the } k\text{-th factor occurs at residual level 1 for observation } i \\
-1 & \text{if it occurs at residual level 0 ,}
\end{cases} \\
\mathbf{\theta} &= (A_{11}', B_{11}, \ldots, R_{11}) \text{ is the } \omega\text{-vector of residual level 1 effects, and} \\
\mathbf{\varepsilon} &= \text{ an } N\text{-vector with } N(0, \sigma^2 I_N) \text{ distribution.}
\end{aligned}
\]

The expected value of \( \mathbf{y} \) may be written as

\[ 
\mathbf{E} \mathbf{y} = \begin{pmatrix} \mathbf{1}_N' \mathbf{G}' \mathbf{D}' \mathbf{R}' \end{pmatrix} 
\begin{pmatrix} \mu \\
\mathbf{\delta} \\
\mathbf{\Gamma} \\
\mathbf{\phi} 
\end{pmatrix} = \mathbf{A} \mathbf{\theta},
\]  

which is in the form of the general linear model. To solve for the parameters \( \mathbf{\theta} \) we have the usual normal equations

\[ 
\mathbf{A} \mathbf{A}' \mathbf{\theta} = \mathbf{A} \mathbf{y},
\]  

We will examine the two sides of this equation separately.

The matrix \( \mathbf{A} \mathbf{A}' \) may be written

\[ 
\begin{pmatrix} \mathbf{1}_N' \\
\mathbf{G}' \\
\mathbf{D}' \\
\mathbf{R} 
\end{pmatrix} \begin{pmatrix} \mathbf{1}_N \\
\mathbf{G}' \\
\mathbf{D}' \\
\mathbf{R}' 
\end{pmatrix} = \begin{pmatrix} \mathbf{1}_N' \mathbf{1}_N \\
\mathbf{G}' \mathbf{G}' \\
\mathbf{D}' \mathbf{D}' \\
\mathbf{R}' \mathbf{R}' 
\end{pmatrix}.
\]  

This matrix is symmetric, so we need only examine the products on and above the main diagonal.

\[ 
\mathbf{1}_N' \mathbf{1}_N = N.
\]  

\( \mathbf{1}_N' \mathbf{G}' \) is an \( m \)-dimensional row vector of column sums of \( \mathbf{G}' \). The sum of the \( j \)-th column of \( \mathbf{G}' \) is the number of observations made on unit \( j \), which is \( n \).
\(1' \cdot G' = n1' \cdot \frac{1}{m}\) \hspace{1cm} (6.6.6b)

\(1' \cdot D'\) is an \(\omega\)-dimensional row vector of column sums of \(D'\). The sum of the \(k\)-th column is the number of times the direct effect of factor \(k\) occurs at level 1 minus the number of times it occurs at level 0, which is 0.

\(1' \cdot D' = 0' \cdot \frac{1}{\omega}\) \hspace{1cm} (6.6.6c)

where \(0' \cdot \frac{1}{\omega}\) is an \(\omega\)-vector of zeros.

\(1' \cdot R' = 0' \cdot \frac{1}{\omega}\) \hspace{1cm} (6.6.6d)

for the same reason as (6.6.6c) holds, with residual levels taking the place of direct levels.

The \((i,p)\) element of \(G'\) is given by

\[
\sum_{k=1}^{N} g'_{ki}g'_{kp} = \text{number of times an observation occurs on both unit } i \text{ and unit } p
\]

\[
= \begin{cases} 
0 & \text{if } i \neq p \\
n & \text{if } i = p 
\end{cases}
\]

\(G' = nI_m\) \hspace{1cm} (6.6.6e)

The \((i,p)\) element of \(GD'\) is given by

\[
\sum_{k=1}^{N} g'_{ki}d'_{kp} = \text{number of times the direct effect of factor } p \text{ occurs at level 1 on unit } i \text{ minus the number of times it occurs at level 0}
\]

\[
= 0
\]

\(GD' = 0_{m \times \omega}\) \hspace{1cm} (6.6.6f)

Similarly,

\(GR' = 0_{m \times \omega}\) \hspace{1cm} (6.6.6g)
The \( (i,p) \) element of \( DD' \) is given by

\[
\sum_{k=1}^{N} d'_k d'_{kp} = \text{number of times the direct effect of factor } i \text{ occurs at the same level as the direct effect of factor } p \text{ minus the number of times they occur at different levels.}
\]

\[
= \begin{cases} 
0 & \text{if } i \neq p \text{ and no design confounding between these effects} \\
\frac{n}{2q} & \text{if } i \neq p \text{ and these effects are design confounded } q \text{ times at the same level} \\
-\frac{n}{2q} & \text{if } i \neq p \text{ and these effects are design confounded } q \text{ times at different levels} \\
N & \text{if } i = p
\end{cases}
\]

\[
DD' = NI_{\omega} + (DD')', \quad (6.6.6h)
\]

where \((DD')'\) is \(DD'\) with the main diagonal replaced by zeros.

The \((i,p)\) element of \( DR' \) is given by

\[
\sum_{k=1}^{N} d'_k r'_{kp} = \text{number of times the direct effect of factor } i \text{ occurs at the same level as the residual effect of factor } P \text{ minus the number of times they occur at different levels}
\]

\[
= \begin{cases} 
0, \frac{n}{2q}, \text{ or } -\frac{n}{2q} & \text{if } i \neq p, \text{ as above} \\
-nt & \text{if } i = p \text{ and } t \text{ is the number of times each factor alternates}
\end{cases}
\]

\[
DR' = -ntI_{\omega} + (DR')', \quad (6.6.6i)
\]

where \((DR')'\) is \(DR'\) with the main diagonal replaced by zeros.

The \((i,p)\) element of \( RR' \) is given by
\[
\sum_{k=1}^{N} r_i^* r_p^* = \text{number of times the residual effect of factor i occurs at the same level as the residual effect of factor p minus the number of times they occur at different levels}
\]
\[= 0, \frac{n}{2q}, -\frac{n}{2q}, \text{ or } N \text{ as for DD'}
\]
\[RR' = NI_{\omega} + (RR'), \quad (6.6.6j)
\]

where \((RR')\) is \(RR'\) with the main diagonal replaced by zeros.

Using these results we write
\[
AA' = \begin{bmatrix}
N & n_{1m}' & 0' & 0' \\
n_{1m} & n_{m} & 0_{m\times\omega} & 0_{m\times\omega} \\
0 & 0 & NI_{\omega} + (DD') & -ntI_{\omega} + (DR') \\
0 & 0 & -ntI_{\omega} + (DR') & NI_{\omega} + (RR')
\end{bmatrix}.
\quad (6.6.7)
\]

The symmetric matrices \((DD')\), \((DR')\), and \((RR')\), represent the design confounding. From the pattern of this matrix it is seen that the estimates for factor effects are orthogonal to those for the mean and unit effects.

The right hand side of \((6.6.4)\) can be written as
\[
\begin{bmatrix}
\frac{1}{nN} \\
G \\
D \\
R
\end{bmatrix} \chi = \begin{bmatrix}
T \\
G \\
D \\
R
\end{bmatrix},
\quad (6.6.8)
\]

where \(T\) is the sum of all observations,
\[
G = (G_1, G_2, \ldots, G_m)', \text{ and } G_i \text{ is the sum of observations on unit i},
\]
\[
D = (D_1, D_2, \ldots, D_\omega)', \text{ and } D_i \text{ is the sum of observations for direct level 1 of factor i minus the sum of observations for direct level 0 of factor i}, \text{ and}
\]
$R = (R_1, R_2, \ldots, R_\omega)$ and $R_1$ is as for $D_1$ but with residual levels replacing direct.

Combining (6.6.7) and (6.6.8) the normal equations (6.6.4) become

$$
\begin{bmatrix}
N & 1' & 0' & 0' \\
1_{-m} & nI_m & 0_{m \times \omega} & 0_{m \times \omega} \\
0_{\omega} & 0_{\omega \times m} & NI_\omega + (DD') & -ntI_\omega + (DR') \\
0_{\omega} & 0_{\omega \times m} & -ntI_\omega + (DR') & NI_\omega + (RR')
\end{bmatrix}
\begin{bmatrix}
\mu \\
\delta \\
\Gamma \\
\Phi
\end{bmatrix}
= 
\begin{bmatrix}
T \\
G \\
D \\
R
\end{bmatrix} \quad (6.6.9)
$$

Using the constraint $\sum_{j=1}^{m} \delta_j = 0$, the first $m+1$ equations of (6.6.9) give

the estimates

$$
\hat{\mu} = T/N \quad (6.6.10)
$$

$$
\hat{\delta}_j = G_j/n - T/N \quad (6.6.11)
$$

The estimates of factor effects are given by

$$
\begin{bmatrix}
\hat{\Gamma} \\
\hat{\Phi}
\end{bmatrix}
= 
\begin{bmatrix}
NI_\omega + (DD') & -ntI_\omega + (DR') \\
-ntI_\omega + (DR') & NI_\omega + (DD')
\end{bmatrix}
^{-1}
\begin{bmatrix}
D \\
R
\end{bmatrix} \quad (6.6.12)
$$

(Note that (RR') has been replaced by (DD') since they are identical matrices.)

If we let the generalized inverse in this equation system be $B$, then the estimated covariance matrix of the factor effects is

$$
B\hat{\sigma}^2, \quad \hat{\sigma}^2 = \frac{\chi'[I_N - A' (AA')^{-1} A] \chi}{N - 2\omega - m} .
$$

To test whether or not there is a difference between using level 0 and level 1 of a particular factor, we test

$$
H_2: \Gamma_{dl} + \Gamma_{rl} = 0 \quad \text{some } \Gamma \in \Lambda . \quad (6.6.13)
$$
If there is a difference between levels it will be indicated by a 
rejection of \( H_2 \), as a result of restrictions (6.1.1a-b). The usual \( t \)-
test is given by rejecting \( H_2 \) with significance level \( \alpha \) \((0 < \alpha < 1)\) if

\[
\left| \frac{\hat{\Gamma}_\gamma + \hat{\Phi}_\gamma}{\hat{\sigma}^2 (b_{y\gamma} + b_{w+y,\omega+y^+} + 2b_{y,\omega+y})} \right| \geq t_{(1-\alpha)/2,u}
\]

where \( t_{(1-\alpha)/2,u} \) is the 100\((1-\alpha)/2\)% point of the \( t \) distribution with \( u \)
degrees of freedom, \( u = N - 2\omega - m \), and \( B = (b_{r_e}) \).

Tests of the form (6.6.14) for different \( \gamma \) are not necessarily
independent. For the designs given in this chapter the tests are inde-
dependent for \( \omega \) even but not for \( \omega \) odd.

6.7 Simplified Analysis for \( \omega \) Even

Because the designs given for \( \omega \) even have no design confounding
between levels of different factors, we can derive simplified formulas
for (i) estimating direct and residual effects, (ii) estimating direct
effects under the assumption that residual effects are absent, and (iii)
testing for the absence of all residual effects. In fact, these results
hold for any design with no design confounding between factors.

Suppose a particular design for \( \omega \) even is repeated \( s \) times so
that \( m = 2\omega, n = 4 \), and \( N = 4s\omega \). In (6.6.12) all the elements of (DD')
and (DR') are 0, so we write

\[
\begin{pmatrix}
\hat{\Gamma} \\
\hat{\Phi}
\end{pmatrix} = 
\begin{bmatrix}
4s\omega I_{\omega} & 4t I_{\omega} \\
-4t I_{\omega} & 4s\omega I_{\omega}
\end{bmatrix}^{-1}
\begin{bmatrix}
D \\
R
\end{bmatrix}.
\]

(6.7.1)

Note that the generalized inverse will be a true inverse here. In terms
of Kronecker matrix products this inverse is
\[
\begin{bmatrix}
4I_\omega \otimes (sw & -t) \\
\quad & \quad \\
-t & sw \\
\end{bmatrix}^{-1} = \frac{1}{4} I_\omega \otimes \begin{bmatrix}
sw & -t \\
-t & sw \\
\end{bmatrix}^{-1}
\]

\[
= \frac{1}{4(s^2 + t^2)} I_\omega \otimes \begin{bmatrix}
sw & t \\
t & sw \\
\end{bmatrix}.
\] (6.7.2)

Placing this result in (6.7.1) we get for factor $\gamma \in \Lambda$

\[
\begin{bmatrix}
\hat{I}_\gamma \\
\hat{\phi}_\gamma \\
\end{bmatrix} = \frac{1}{4(s^2 + t^2)} \begin{bmatrix}
sw & t \\
t & sw \\
\end{bmatrix} \begin{bmatrix}
D_\gamma \\
R_\gamma \\
\end{bmatrix}
\] (6.7.3)

Tests of hypotheses of the form (6.6.13) are independent for different $\Gamma$ and (6.6.14) becomes

\[
\left| \frac{(sw+t)(D_\gamma + R_\gamma)}{4(s^2 + t^2)} \right| \geq \frac{2(sw+R)}{4(s^2 + r^2)} \hat{\sigma}^2 \geq t(1-\alpha)/2, u
\]

or

\[
\frac{|D_\gamma + R_\gamma|}{2\hat{\sigma}^2} \geq t(1-\alpha)/2, u = \omega(3s-2).
\] (6.7.4)

If we make the assumption that no residual effects are present in our model, (6.6.3) becomes

\[
E_Y = (1_N; G'; D') \begin{bmatrix}
\delta \\
\Gamma \\
-0 \\
\end{bmatrix}
\]

\[
= A_1^\theta \Delta_1,
\] (6.7.5)

and

\[
(A_1 A_1^\dagger)^{-} = \begin{bmatrix}
4s_w & 41_m^\dagger \\
41_m & 4I_m \\
0_{\omega \times (m+1)} & 1_{4s_w I_\omega} \\
\end{bmatrix}.
\] (6.7.6)
Then the estimate of the direct effect of treatment $\gamma$ is

$$\hat{\Gamma}_{\gamma 0} = D_{\gamma} / 4s_{\omega}.$$  (6.7.7)

If we repeat the analysis assuming now that residual effects are present, we can get the new direct effect estimates from the previous ones as

$$\hat{\Gamma}_{\gamma} = \frac{s_{\omega}^{2}}{s_{\omega}^{2} - t^{2}} \hat{\Gamma}_{\gamma 0} + \frac{t}{s_{\omega}^{2} - t^{2}} R_{\gamma}.$$  (6.7.8)

For the case of $\omega$ even and testing the hypothesis that no residual effects are present, i.e., $H_0: \Phi = 0$, we use the usual $F$ statistic

$$\frac{\chi' [A'(AA')^{-1} A - A_{1}'(A_{1}A_{1}')^{-1}A_{1}] \chi}{\chi[I - A'(AA')^{-1}A] \chi} = F$$  (6.7.9)

which under $H_0$ has the $F$ distribution with $\omega$ and $\omega(3s-2)$ degrees of freedom. Letting

$$M = \begin{bmatrix} 4s_{\omega} & 4I_m^{1'} \\ 4I_m & 4I_m \end{bmatrix} = \begin{bmatrix} 0 & 0' \\ 0 & I_m \end{bmatrix}$$  (6.7.10)

and $P$ be the matrix in (6.7.2), we have the following simplification of matrices from (6.7.9):

$$A'(AA')^{-1} A = (1_N:G':D':R') \begin{bmatrix} M_{(m+1)\times(m+1)} & 0_{(m+1)\times2\omega} \\ 0_{\omega\times(m+1)} & P_{2\omega\times2\omega} \end{bmatrix} \begin{bmatrix} 1'_N \\ \cdot \cdot \cdot \\ G \\ \cdot \cdot \cdot \\ D \\ \cdot \cdot \cdot \\ R \end{bmatrix}$$

$$= (1_N:G')M \begin{bmatrix} 1'_N \\ \cdot \cdot \cdot \\ G \\ \cdot \cdot \cdot \\ D \\ \cdot \cdot \cdot \\ R \end{bmatrix} + (D':R')P \begin{bmatrix} 1'_N \\ \cdot \cdot \cdot \\ G \\ \cdot \cdot \cdot \\ D \\ \cdot \cdot \cdot \\ R \end{bmatrix}$$  (6.7.11)
\[ A_1'(A_1A_1')^{-1}A_1 = (1_{\mathbf{N}}:G'':D'') \begin{pmatrix} M_{(m+1)\times(m+1)} & 0_{(m+1)\times\omega} \\ 0_{\omega \times (m+1)} & \frac{1}{4s\omega} I_{\omega} \end{pmatrix} \begin{pmatrix} 1'_{\mathbf{N}} \\ \frac{1}{\omega} \\ \frac{G}{\omega} \\ D \end{pmatrix} \]

\[ = (1_{\mathbf{N}}:G')M \begin{pmatrix} 1'_{\mathbf{N}} \\ -\frac{1}{\omega} \end{pmatrix} + \frac{1}{4s\omega} D'D' \quad (6.7.12) \]

Using \( (1_{\mathbf{N}}:G')M \begin{pmatrix} 1'_{\mathbf{N}} \\ \frac{G}{\omega} \end{pmatrix} = \frac{1}{4} G'G \) (6.7.9) becomes

\[
\frac{\mathcal{X}' \left[ \begin{pmatrix} (D':R')P \begin{pmatrix} D' \\ R \end{pmatrix} - \frac{1}{4s\omega} D'D' \end{pmatrix} \right] \mathcal{X}}{\mathcal{X}' \left[ \begin{pmatrix} I_{\mathbf{N}} - \frac{1}{4} G'G - (D':R')P \begin{pmatrix} D' \\ R \end{pmatrix} \end{pmatrix} \right] \mathcal{X}} = F \quad (6.7.13)
\]

We reject \( H_0 \) if \( F > F_{1-\alpha,\omega,\omega(3s-2)} \), with significance level \( \alpha \) \( (0 < \alpha < 1) \), where \( F_{1-\alpha,\omega,\omega(3s-2)} \) is the 100(1-\(\alpha\)% point of the F distribution with \( \omega \) and \( \omega(3s-2) \) degrees of freedom.
BIBLIOGRAPHY


21. _________ (1968), Serial factorial designs, Biometrika, 55, 67-81.


25. _________ (1965), The theory of least squares when the parameters are stochastic and its application to growth curves, Biometrika, 52, 447-458.


