MINIMAL MULTIDIMENSIONAL DESIGNS

by

John T. Sennetti

Thesis submitted to the Graduate Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Statistics

APPROVED:

Dr. Klaus Hinkelmann, Chairman

Dr. Boyd Harshbarger
Department of Statistics

Dr. Clyde Kramer
Department of Statistics

Dr. Ernest Jennelle
Department of Civil Engineering

Dr. Raymond Myers
Department of Statistics

July, 1972

Blacksburg, Virginia

ACKNOWLEDGEMENTS

I would like to express my appreciation to Dr. Klaus Hinkelmann for his patience, timely advice, and scholarly guidance in this research. Appreciation is also expressed to Dr. Boyd Harshbarger whose dedication to the success of his graduate students will long be remembered, to Dr. Clyde Kramer whose suggestions helped clarify the analysis and organize its presentation, to Dr. Raymond Myers for his contributions and interest and to Dr. Ernest Jennelle for his moral support and encouragement. Sincere gratitude is given to the National Institute of Environmental Health Sciences which supported this investigation under NIH Training Grant, numbers 7T01 ES000 33-08 and 7T01 ES000 33-09.

Finally, I wish to thank my wife, Cynthia, for her patience, endurence, and encouragement during the preparation of this thesis. Heartfelt appreciation is due to Cynthia, who undertook the arduous task of typing the rough draft in its various stages and the manuscript in its final form. It may be truly said that this work could not have been accomplished without her untiring support. Therefore, this thesis is dedicated to her.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF TABLES

CHAPTER I

INTRODUCTION

An experimental design may be considered an arrangement of the degrees (levels) of identified forces (factors) in a plan of investigation. The purpose of this arrangement is to control or measure factorial effects on experimental units to which factor level combinations are applied. Multidimensional designs (MD's) are multifactor fixed effect experimental designs. Those MD's where not all factor level combinations occur (which we will call incomplete designs) shall be the subject of our investigation. An example of such a design is a Latin square where rows, columns, and treatments are considered fixed effect factors, each at k levels. Other examples include Graeco-latin squares, lattice designs, balanced and partially balanced incomplete block designs, F-squares, and multidimensional partially balanced designs.

The above designs are used when no interaction is assumed between any factors such as treatments, blocks, rows, columns, etc. This is not always a valid assumption. In those cases where this assumption may not hold, it is natural to investigate designs which allow for the presence of certain two-factor interactions. Some designs which allow for this possibility are given by Potthoff (1958, 1962), Causey (1967, 1968), and Anderson (1968). These latter designs are special cases of fractional factorials where some interactions are assumed negligible. Therefore MD's are fairly common. Researchers in industry, education, and almost all scientific fields have used these plans of investigation. (Potthoff, 1963; Bose and Srivastava, 1964a)

1.1 Minimal Designs

One problem that arises in such plans of investigation is that the number of factors, and thus the number of factor level combinations, may become too large to carry out. In an attempt to economize on experimental units, experimenters use experience or prior knowledge concerning the factors involved and hence choose designs to estimate only certain effects considered important. An experimenter may assume two-factor and higher order interactions to be negligible. Orthogonal main effect plans, for example, permit the estimation of only the main effects. Assuming no interactions present, one may obtain these estimates with a desired property, orthogonality. Also, some economy of experimental units is achieved in these designs. However, with the same number of experimental units or even less, it may be possible to allow for two-factor interactions, but give up orthogonality. Having information on two-factor interactions may be considered more important than having orthogonal main effect estimates. Perhaps in using the experimenter's knowledge about the assumed effects it may be possible to emphasize economy entirely. Presently there is no literature on experimental designs that (1) permit the estimation of all assumed main effects and two-factor interactions and (2) use as few assemblies (factor level combinations) as possible. The MD's available require more assemblies than absolutely necessary. These "extra" assemblies provide such properties as orthogonality and balance. By ignoring these properties, we may require fewer assemblies.

Consider, as a point of illustration, an example given by Potthoff (1963). The example concerns a pricing experiment on certain types (A,B,C) of frozen fruit juice. This experiment will be a preliminary test which will provide information for future tests. The experimenters wish to determine the effect of three different prices (1,2,3) on the sales of each of three different types of fruit juices. The experiment is to take place on the days, Thursday, Friday, and Saturday (I,II,III), when store sales are the greatest. Thus the three factors in the experiment are price, type of fruit juice, and day.

A complete factorial experiment would require that each type of juice be sold every day at all three prices, simultaneously. Potthoff (1963) says,

> This would be out of the question: the same type of juice could not be offered for sale in the store at more than one price at any one time. Hence it is necessary to consider using an incomplete design which would specify the selling price of each type of juice for each day of the experiment.

Potthoff (1963) in this example later considers one type of possible two-factor interaction, a price-juice interaction. We shall, for the purpose of illustration, assume there is no such interaction. A price-juice interaction would mean that lower prices would increase the sales of some juice more than of others. Also, the experimenters assume no price-day interaction:

> The socio-economic characteristics of the shoppers are not different depending on the day of the week and/or the weather, and a price elasticity of demand is not assumed to be different for the socio-economic groups of consumers.

As for a juice-day interaction, the experimenters are willing to assume that the choice of juice would not be influenced by the day of the week, weather, or socio-economic factors. From this example and these assumptions, we will now begin to illustrate a certain type of design.

Assuming no interaction occurs between or among types of fruit juices, prices or days, two degrees of freedom (d.f.) are required for each factor in order to estimate the differences in juice types, pricing levels, and days. Therefore a total of seven particular periods involving only the days Thursday, Friday, and Saturday, and different price-juice level combinations, are needed to estimate the desired effects. An orthogonal main effect plan may be considered here: a 1/3 replicate of a 3^3 factorial using the factor level combinations from a 3x3 Latin square. However, this design requires nine level combinations (assemblies) and only seven are absolutely necessary. These seven could be the level combinations (1AI), (2BII), (3CII), (1BI), (2CIII), (3CIII), and (1AII). More shall be said on this type of design in Chapter V.

Since careful selection and monitoring of particular test periods will incur considerable expense, it is desired that as few periods as possible be used, especially since the study will be a preliminary test. Further, it may be that only very few periods are available for such a test.

For the moment let us ignore experimental error. A design utilizing only seven periods would have to be non-orthogonal and could not have the usual analysis of a Latin square, but would at least be preferred to not experimenting at all, if this were the choice. An estimate of error then, could perhaps come from prior information (see Cochran and Cox, 1957, p.262) or some other source. Note that with only eight test periods available a design using seven periods would still apply. If 9 periods were still available, it still is feasible then to perform a particular set of seven necessary to estimate desired effects. Given the necessary seven assemblies, any more assemblies would then provide d.f. for experimental error.

This fruit juice study illustrates the type of problem one may encounter in constructing and analyzing a design with a minimum of factor level combinations: a minimal design. The real usefulness of minimal designs is made more apparent when we consider many factors each with many levels. For these cases even "economic" designs (Bose and Srivastava, 1964a) may require more assemblies than necessary. Our interest will be directed toward minimal designs of m factors F_i, with n_i levels, $i = 1,\ldots m$, that permit not only the estimation of main effects, but certain assumed interactions as well. Chapter V will present a class of such designs with examples that assume two factor interaction models.

1.2 Augmenting Designs

Parallel to the need for a minimal design is the need for a minimum set of new factor level combinations added to an established design for the purpose of comparing effects not estimable or not assumed present in the original design. For example, in many factorial main effect plans, the question is occasionally raised on the possible presence of two-factor interactions. The orthogonal main effect plans of Addelman and Kempthorne (1961) enable an experimenter to estimate the differences among levels of each factor using some degree of economy in level combinations. Therefore a procedure for adding factor level combinations to such a main effect plan to allow for interaction would be highly desirable.

This is particularly true, if due to the limitations of experimental units, the usual factorial fractions (incomplete factorials) do not apply. The advantage here is to make use of the existing main effect design rather than to construct new factorial fractions which may require more assemblies than really necessary. Also, since incomplete designs are chosen for their characteristic economy of experimental units, it may seem especially desirable to the experimenter if the factor level combinations to be added are minimum in number; i.e., minimal. Such an augmented design would provide information on suspected interaction and future experiments may then be performed in light of this information.

Let us consider the study of fruit juices mentioned earlier. Assume that the nine assemblies necessary for the orthogonal 1/3 replicate are available. The questions may be raised on what assemblies (and thus periods) are to be added to the original nine in order to allow for the possible effects of a price-juice interaction. Since test periods may be at a premium, the fewest periods necessary may be preferred. If a preliminary test is made and no significant interaction noted, future studies may require fewer periods than used by the augmented design.

For example, let the nine assemblies of a 1/3 replicate be represented by (1AI), (2BI), (LCII), (2AII), (3BII), (LBIII), (2CIII), and (3AIII). The pair of assemblies (1BII) and (1BI) could be used to augment the original nine for the presence of a price-juice interaction. More shall be said of this in Chapter IV.

1.3 Optimal Designs

The very ability to construct a minimal design or a procedure to augment designs with the minimum number of assemblies may be in itself an optimality criterion. Certainly these designs may not have the orthogonal properties of common MD's. One must pay a price for using less information or fewer experimental units. However, given this constrain of limited assemblies, we may choose from a set of available minimal designs a design that is optimal according to common optimality criteria. For example, from a set of augmented designs we may determine which augmentation procedure enables one to estimate certain desired effects with the least average variance. Other common optimality criteria shall be discussed in Chapter V. Methods of constructing "optimum" sets from which one chooses an optimal design shall be presented also.

1.4 Literature Review

Before pursuing our investigation any further, let us provide some foundation for our inquiry. As mentioned, no previous work is given in the literature where development is directed specifically toward minimal designs. However, we may present some related works in the analysis and construction of some MD's where main effects and certain interactions are of interest. We shall limit our review to this area of interest rather than present a review of extensive development in experimental designs that includes investigation by Fisher (1935,) Yates (1936), Youden (1937), Cornish (1938) and many more. Kempthorne (1952), Scheffe (1959), and Hinkelmann (1968a) cover most of this work to date. After discussing useful research in the analysis and construction of MD's, we will present previous work in augmenting MD's.

Potthoff (1958) first used the term "multidimensional" to describe multifactor fixed effect designs. His investigations are in the construction and analysis of designs that estimate main effects and some interactions, and involve some economy of experimental units.

Potthoff (1958, 1962a, 1962b, 1963) constructed his design classes from Kronecker products of balanced and partially balanced incomplete block designs. This enabled Potthoff to develop from the normal equations a reduced form for one factor similar to the reduced from one receives for treatments in the analysis of incomplete block designs. However, Potthoff's designs are not minimal and are actually special cases of a more general class: multidimensional partially balanced designs (MPBD's) first described by Bose and Srivastava (1964b) and later by Anderson (1968). Bose and Srivastava present an algorithm for solving the full normal equations which result from such designs. This analysis, although more general than Potthoff's is considered tedious and unappealing (Hoke, 1971). Further, MPBD's are not minimal designs. However, in the construction of some MPBD's, Anderson (1968) indicated some helpful theorems on the concept of connected designs which shall later prove useful in constructing and augmenting MD's.

Interest toward minimal designs is currently increasing. Daniel (1971) presents a minimal main effect 2^3 factorial plan. Minimum designs were the subjects of discussion by Federer, Hedayat, and Raktoe (1972), and more work in the area of minimal designs is expected to appear.

The concept of augmenting designs is not new. Methods of adding factorial fractions to established designs in order to estimate effects not estimable or not assumed present have been given by Daniel (1962), Addelman (1963), and John (1966). These methods involve constructing assemblies in order to break the alias chains of two factor interactions, so that certain interaction effects may be estimated. However, the direction of this work is toward developing balanced fractions, rather than fractions involving the minimum number of assemblies.

Augmenting block designs with more treatments (and thus factor level combinations of treatments and blocks) is discussed by Federer (1956, 1961). Federer develops a procedure for adding new treatments (more levels of one factor) to a design and thus the nature of augmenting is slightly different. However, the addition of these new treatments is achieved with the minimum of experimental units. Also if one consideres the set of treatments (old and new) as levels of one factor, Federer's work for fixed effects then may be considered a special case of what we are about to present in Chapters IV and V.

1.5 Comments On The Available Literature

In reviewing the current literature on experimental designs, one observations can not be overlooked. It is noted by J. T. Tocher in the critical review following Kiefer (1959) that the construction of many designs involves such mathematical expertise, that experimenters often view such construction more as a mathematical game, rather than as a practical plan of investigation. Consequently, it is not uncommon for a researcher, trained in agriculture for example, not to be familiar with a Galois field, hypersphere, or pratially balanced array. (For example, see Vadja, 1967).

Our research has been motivated, therefore, by strong practical considerations. Due to unique experimental constraints on factor level combinations or on the number of experimental units, common MD's often do not apply. Hopefully, since the designs about to be discussed require fewer experimental constraints, an experimenter will be able to fit a design to the problem at hand, rather than change the problem to fit available designs.

In accordance with the experimenter's viewpoint, we shall try to establish designs which are also easy to construct and analyze.

Attempts to simplify analysis are given in Chapters II and III. In these Chapters we shall present a general form for the reduced normal equations. This method generalizes Potthoff's approach and is applicable to all non-orthogonal, incomplete, fixed effect designs, where two-factor interactions may or may not be assumed present. By reducing the equations to be solved, the analysis is made simpler. In place of the analysis given by Bose and Srivastava (1964b) our analysis is recommended. Also, our approach to the analysis serves as a guide to the construction of minimal designs given in Chapter V.

By extending the work of Anderson (1968) to include interaction models, and by applying these results, we present a new approach to the construction of experimental designs in Chapter IV. With this new approach, an experimenter can construct the needed factor level combinations directly, rather than first investigate such mathematical concepts as vector spaces or investigate designs which may require changing the experiment. In summary then, we hope to present statistical designs and methods that will be easily grasped by any experimenter.

A good factorial fraction is said to have (1) economy, (2) little correlation between desired estimates, (3) fairly uniform variance on these estimates, and (4) an easy method of application and analysis (Bose and Srivastava 1964a). Fractions with all four properties are not at all common. In order to achieve any one property, other properties may have to be sacrificed.

For minimal designs, (2) and (3) may be difficult to achieve. Therefore, in our work we shall concentrate on properties (1) and (4), but shall not ignore (2) and (3) completely. It is with this philosophy that we construct minimal designs and augment designs with a minimum of experimental units.

CHAPTER II

THE ANALYSIS OF DESIGNS WHERE NO INTERACTION IS ASSUMED

Designs which estimate only main effects and use the minimum number of experimental units, may not be orthogonal. If this is true, then an analysis more general than that of orthogonal designs is necessary. In this chapter we shall develop a special approach to the analysis of non-orthogonal m factor designs where no interaction is assumed present.

The purpose of this special approach is to simplify the analysis of the design. This method also provides a device for the construction of the minimal designs discussed in Chapter V. However, this approach shall be similar in part to that of Potthoff (1958).

Potthoff's approach follows the usual procedure for finding treatment estimates adjusted for blocks in incomplete block designs (for example, see Hinkelmann, 1968a). Using this approach, Potthoff finds the sum of squares due to one factor adjusted for m-1 factors for an m factor design. This sum of squares results from a reduced form of the full normal equations (i.e. solving the reduced normal equations and multiplying these solutions by the right hand side of the reduced normal equations). Potthoff developed his analysis for only partially balanced designs. We shall not restrict ourselves to partially balanced designs, but provide an analysis for all incomplete designs and thereby generalize Potthoff's results.

This chapter shall discuss the following: Section 2.1 presents the two-factor case in order to better understand the analysis of m factors; the usefulness of generalizing this analysis is given in section 2.2; and section 2.3 presents a general form for the reduced normal equations for m factor fixed effect models; and methods for computing this form are given in section 2.4.

2.1 An Introduction To The Analysis Of Two-Factor Designs

We shall first establish notation which shall serve us later for m factors. Assume we have a two-factor design such that a typical observation is represented by $y_{ij\omega}$, where $i = 1,\ldots,n_1$, $j = 1,\ldots,n_2$, and $\omega = 0, 1,\ldots,h_{ij}$. Here h_{ij} is an element of the incidence matrix, $H_{12} = (h_{ij})$, and represents the number of times a combination (i,j) is replicated (or the number of times factor one at level i appears with factor two at level j). Note that if $h_{ij} = 0$, $y_{ij\omega}$ does not appear in the set of observations. Let us assume error terms are normally, independently, homoscedastically distributed such that $y_{ij\omega} \sim N[E(y_{ij\omega}),\sigma^2]$; for purposes of estimation the normality assumption can be dropped.

For the two factor no interaction model, $E(y_{ij\omega})$ is written as

$$E(y_{ij\omega}) = \mu + f_1^i + f_2^j \,, \tag{2.1.1}$$

where μ is the general mean and f_1^i and f_2^j are the i-th and j-th level effects of the first and second factor respectively. We may rewrite (2.1.1) as

(2.1.2) $$E(y_{ij\omega}) = p_1^i + p_2^j$$

where, for example, $p_1^i = f_1^i$ and $p_2^j = f_2^j + \mu$. Also, for convenience, we shall use standard matrix and vector notation. Therefore, let $\underline{p}_1' = (p_1^1, \ldots, p_1^{n_1})$ and $\underline{p}_2' = (p_2^1, \ldots, p_2^{n_2})$ be the factor parameters in the in the model, and $X = [X_1 \vdots X_2]$ be the design matrix with respective design matrices, X_1 of order $h x n_1$ and X_2 of order $h x n_2$, to be associated with $\underline{p}_1$ and $\underline{p}_2$. In defining X_1 and X_2 in this manner, H_{12} may then be expressed as $H_{12} = X_1' X_2 = (h_{ij})$. Let $h = \sum_i \sum_j h_{ij}$ and $\underline{y}$ be the hx1 observation vector. Then

(2.1.3) $$E(\underline{y}) = X_1 \underline{p}_1 + X_2 \underline{p}_2 \qquad \text{where}$$

(2.1.4)

$$\underline{y}' = (y_{111}, y_{112}, \ldots, y_{11h_{11}}; y_{121}, \ldots, y_{12h_{12}}; \ldots, y_{n_1 n_2 h_{n_1 n_2}})$$

Further, let us define factor totals which will be used later:

(2.1.5) $$X_1' \underline{y} = \underline{Y}_1 \text{ and } X_2' \underline{y} = \underline{Y}_2$$

with

$$\underline{Y}_1' = (Y_{11}, Y_{12}, \ldots, Y_{1n_1}),$$

$$\underline{Y}_2' = (Y_{21}, Y_{22}, \ldots, Y_{2n_2})$$

and

$$Y_{11} = \sum_{j}^{n_2} \sum_{\omega}^{h_{1j}} y_{1j\omega}; \ldots ; Y_{1n_1} = \sum_{j}^{n_2} \sum_{\omega}^{h_{n_1 j}} y_{n_1 j\omega},$$

$$Y_{21} = \sum_{i}^{n_1} \sum_{\omega}^{h_{i1}} y_{i1\omega}; \ldots ; Y_{2n_2} = \sum_{i}^{n_1} \sum_{\omega}^{h_{in_2}} y_{in_2\omega}.$$

Let Δ be a linear unbiased estimator of some function of the first factor effects only. $E[\Delta]$, then, must not contain any second factor effects. Since Δ is a linear estimate, it must represent a linear combination of the observations, as for example

(2.1.6) $$\Delta = \underline{q}_1' \underline{Y}_1 + \underline{q}_2' \underline{Y}_2 \qquad \text{or}$$

(2.1.7) $$\Delta = \sum_i q_{1i} Y_{1i} + \sum_j q_{2j} Y_{2j} ,$$

where the $\underline{q}_i'$ are vectors of real numbers and

(2.1.8) $$\underline{q}_1' = (q_{11}, q_{12}, \ldots, q_{1n_1})$$

$$\underline{q}_2' = (q_{21}, q_{22}, \ldots, q_{2n_2})$$

Using first (2.1.6) and then (2.1.5) and (2.1.3), we can express $E(\Delta)$ as:

(2.1.9) $$E(\Delta) = E\,(\underline{q}_1' \underline{Y}_1 + \underline{q}_2' \underline{Y}_2)$$

$$= E([\underline{q}_1' X_1' + \underline{q}_2' X_2'] \underline{y})$$

$$E(\Delta) = (\underline{q}_1' X_1' + \underline{q}_2' X_2')(X_1 \underline{p}_1 + X_2 \underline{p}_2)$$

$$= (\underline{q}_1' X_1' + \underline{q}_2' X_2') X_1 \underline{p}_1 + (\underline{q}_1' X_1' + \underline{q}_2' X_2') X_2 \underline{p}_2$$

We observe here that $E(\Delta)$ will be free of factor two if and only if

(2.1.10) $$\underline{q}_1' X_1' X_2 + \underline{q}_2' X_2' X_2 = 0'$$

If we let $H_{21} = X_2' X_1$, $H_{11} = X_1' X_1$, and $H_{22} = X_2' X_2$ we may write

(2.1.10) as follows

(2.1.11) $$\underline{q}_1' H_{12} + \underline{q}_2' H_{22} = 0$$

We should note that the above notation is found in the normal equation

as

(2.1.12) $$\begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} \underline{p}_1 \\ \underline{p}_2 \end{bmatrix} = \begin{bmatrix} \underline{Y}_1 \\ \underline{Y}_2 \end{bmatrix},$$

Note that $H_{11} = \text{diag}\,(h_{i.})$ and $H_{22} = \text{diag}\,(h_{.j})$, with

$$h_{i.} = \sum_{j=1}^{n_2} h_{ij} \quad \text{and } h_{.j} = \sum_{i=1}^{n_1} h_{ij} \quad .$$

For the purpose of analysis, Potthoff (1958) requires $h_{i.}$, the occurrence of the i-th level of factor one, to be equal to $\frac{h}{n_1}$, and similarly $h_{.j} = \frac{h}{n_2}$; whereas Anderson (1968) requires $h_{i.} = \alpha_1$, be a constant for all levels of factor one, and similarly $h_{.j} = \alpha_2$, a constant for factor two. Note that we require no such restriction here.

From (2.1.11) we get

(2.1.13) $$\underline{q}'_2 = -\underline{q}'_1 H_{12} H^{-1}_{22}$$

Using (2.1.13) in (2.1.9) we then may express E(Δ) as a function of $\underline{p}_1$ only:

(2.1.14) $$E(\Delta) = (\underline{q}'_1 X'_1 - \underline{q}'_1 H_{12} H^{-1}_{22} X'_2) X_1 \underline{p}_1$$

$$= \underline{q}'_1 (H_{11} - H_{12} H^{-1}_{22} H_{21}) \underline{p}_1$$

$$= \underline{q}'_1 C_1 \underline{p}_1$$

where the term involving $\underline{p}_2$ is zero, and C_1 is defined:

$$C_1 = H_{11} - H_{12} H^{-1}_{22} H_{21}$$

The (i,I) element of C_1 is given by

$$c_{iI} = h_{i.} \delta_{iI} - \sum_j \frac{h_{ij} h_{Ij}}{h_{.j}} \qquad (i,I = 1,2,\ldots,n_1)$$

where δ_{iI} is the Kronecker delta. If we substitute (2.1.13) into (2.1.6) we get

(2.1.15) $$\Delta = \underline{q}'_1 (\underline{Y}_1 - H_{12} H^{-1}_{22} \underline{Y}_2)$$

$$= \underline{q}'_1 \underline{Q}_1$$

where

$$\underline{Q}' = (Q_{11}, Q_{12}, \ldots, Q_{1n_1}) = (\underline{Y}_1 - H_{12} H_{22}^{-1} \underline{Y}_2)'$$

Taking (2.1.14) and (2.1.15) we find

$$\underline{q}_1' E(\underline{Q}_1) = \underline{q}_1' C_1 \underline{p}_1$$

or since the above is true for any $\underline{q}_1'$,

$$E(\underline{Q}_1) = C_1 \underline{p}_1$$

so that for a solution $\hat{\underline{p}}_1$ to the above we have

$$\underline{Q}_1 = C_1 \hat{\underline{p}}_1 \tag{2.1.16}$$

A solution to (2.1.16) is a familiar result in least squares theory, (see Hinkelmann, 1968a) $\hat{\underline{p}}_1 = C_1^- \underline{Q}_1$, where C_1^- is a matrix such that

$$C_1 C_1^- C_1 = C_1 \quad .$$

C_1^- is called a generalized inverse (g-inverse) of C_1. More shall be said of g-inverses in section 2.4. Let $r(C_1) = n_1 - r_1$ be the rank of C_1. Note that $n_1 - r_1 \leq n_1 - 1$ since

$$\sum_I c_{iI} = \sum_i c_{iI} = 0 \quad .$$

If $r_1 = 1$, the design is said to be connected and every contrast in $\underline{p}_1$ is then estimable. If $\underline{L}'\,\underline{p}_1$ is a contrast, its best linear unbiased estimate is $\underline{L}'\,\hat{\underline{p}}_1$, and the variance of $\underline{L}'\hat{\underline{p}}_1$ is $\underline{L}'\,C_1^-\,\underline{L}\,\sigma^2$. Finding a matrix C_1^- is therefore quite important for analysis, as we see later.

In a manner similar to Potthoff (1958), we develop the analysis of variance of our two-factor model, and then extend this development to m factor models. The development is based on the conditional error principle which is well known and can be found in Scheffe (1959, Chapter 2).

Under the hypothesis that all first factor effects are equal to g, say, (2.1.3) becomes

$$E(\underline{y}) = X_2\underline{p}_2^* \quad \text{where} \quad \underline{p}_2^* = \underline{p}_2 + g\underline{J}_{n_2 1}$$

and $\underline{J}_{n_2 1}$ is a n_2x1 unity vector.

The conditional error sum of squares (s.s.) under this hypothesis shall be denoted by ss_{ce} and is

$$ss_{ce} = \underline{y}'\underline{y} - \underline{Y}_2'\,H_{22}^{-1}\,\underline{Y}_2$$

From the reduced normal equations (r.n.e.) in (2.1.16), $\hat{\underline{p}}_1$ is found and the s.s. for factor one is

$$ss_{p_1} = \hat{\underline{p}}_1' \underline{Q}_1 .$$

The error s.s. under model (2.1.3) is then

$$ss_e = ss_{ce} - ss_{p_1}$$

$$= \underline{y}'\underline{y} - \underline{Y}_2' H_{22}^{-1} \underline{Y}_2 - \hat{\underline{p}}_1' \underline{Q}_1$$

Therefore in summary,

Source	d.f.	s.s	E(MS)
Factor One	$n_1 - r_1$	$\hat{\underline{p}}_1' \underline{Q}_1$	$\dfrac{\underline{p}_1' C_1 \underline{p}_1}{n_1 - r_1} + \sigma^2$
Factor Two	n_2	$\underline{Y}_2' H_{22}^{-1} \underline{Y}_2$	
Error	$h-n_1 - n_2 + r_1$	$\underline{y}'\underline{y} - \underline{Y}_2' H_{22}^{-1} \underline{Y}_2 - \hat{\underline{p}}_1' \underline{Q}_1$	σ^2
Total	h	$\underline{y}'\underline{y}$	

2.2 The Analysis of m-Factor Designs

Before we extend the results of section 2.1, let us review what has been presented. The results of section 2.2 can be found essentially in Hinkelmann (1968a). This analysis represents no contribution to what is already known, but this development is given to provide a basis for extension to m-factor models.

Let us now extend our results. Assume an m-factor design with a typical observation given by $y_{ij\cdots z\omega}$, where $i=1,\ldots n_1$, $j=1\ldots,n_2,\ldots,$ $z=1,\ldots,n_m$, and $\omega=0,1,\ldots,h_{ij\cdots z}$. Again $h_{ij\cdots z}$ represents the number of times the combination of factor levels $(i,j,\ldots,z)$ appears, and the total number of observations is $h = \sum_i \sum_j,\ldots,\sum_z h_{ij}\ldots z$. As before $y_{ij\cdots z\omega} \sim N[E(y_{ij\cdots z\omega}), \sigma^2]$. The design matrix is

$$X = [\, X_1 \vdots X_2 \vdots \vdots \vdots X_m]$$

and the parameter vectors are

$$\underline{p}_u' = (p_u^1, p_u^2, \ldots, p_u^{n_u})$$

for $u=1,\ldots,m$. It shall be understood that $\underline{p}_m$ includes the mean μ, as did $\underline{p}_2$ in section 2.1. Our model is then

$$E(\underline{y}) = X_1\underline{p}_1 + X_2\underline{p}_2 + X_3\underline{p}_3 + \ldots + X_m\underline{p}_m \tag{2.2.1}$$

Further use of this notation is made in this and later sections. A general form for C_α, the matrix for the reduced normal equations (r.n.e.) for $\underline{p}_\alpha$, will be presented in section 2.3. The s.s. for factor α, adjusted for the other m-1 factors, is then $ss_{p_\alpha} = \hat{\underline{p}}_\alpha' \underline{Q}_\alpha$, with a general form for $\underline{Q}_\alpha$ also given in section 2.3.

By calculating $\underline{p}_\alpha' \underline{Q}_\alpha$, the error s.s. under (2.2.1) follows in a manner similar to section 2.1. Without loss of generality, let $\alpha = 1$. The unadjusted estimates are obtained by assuming the conditional model,

$$E(\underline{y}) = X_2\underline{p}_2 + X_3\underline{p}_3 + X_4\underline{p}_4 + \ldots + X_m\underline{p}_m \,.$$

For this model, the normal equations (n.e.) are,

$$(2.2.2) \qquad \begin{bmatrix} X_2'X_2 & X_2'X_3 & \ldots & X_2'X_m \\ X_3'X_2 & X_3'X_3 & \ldots & X_3'X_m \\ \cdot & \cdot & \ldots & \cdot \\ X_m'X_2 & X_m'X_3 & \ldots & X_mX_m \end{bmatrix} \begin{bmatrix} \hat{\underline{p}}^*_2 \\ \hat{\underline{p}}^*_3 \\ \cdot \\ \hat{\underline{p}}^*_m \end{bmatrix} = \begin{bmatrix} X_2'\underline{y} \\ X_3'\underline{y} \\ \cdot \\ X_m'\underline{y} \end{bmatrix}$$

Let H be the coefficient matrix for the unadjusted estimates $\hat{\underline{p}}^{*\prime} = (\hat{\underline{p}}^*_2, \hat{\underline{p}}^*_3, \ldots, \hat{\underline{p}}^*_m)$ in (2.2.2) and H^- be a generalized inverse for H.

Then $\hat{\underline{p}}^* = H^-\underline{Y}^*_2$ where $\underline{Y}^{*\prime}_2 = (X_2'\underline{y}, X_3'\underline{y}, \ldots, X_m'\underline{y})$ and the error s.s. for (2.2.1) is $\quad ss_e = \underline{y}'\underline{y} - \underline{Y}^{*\prime} H^- \underline{Y}^* - \hat{\underline{p}}_1' \underline{Q}_1$.

In calculating a C_1 matrix for $\underline{p}_1$ in section 2.3, a g-inverse for H^-, is also found, and thus ss_e follows. Also, should a solution to the full n.e. be required, $(\hat{\underline{p}}_1, \hat{\underline{p}}_2, \ldots, \hat{\underline{p}}_m)$, these estimates may be found directly. A solution $\hat{\underline{p}}_1$ is already found from the r.n.e. for $\underline{p}_1$. Then for u=2,...,m, let $\underline{Z}_u = X_u'\underline{y} - X_u'X_1\hat{\underline{p}}_1$, and $\underline{Z}' = (\underline{Z}_2, \underline{Z}_3, \ldots, \underline{Z}_m)$. The following estimates are solutions to the full n.e., since they are adjusted for $\hat{\underline{p}}_1$,

$$(\hat{\underline{p}}_2, \hat{\underline{p}}_3, \ldots, \hat{\underline{p}}_m)' = H^- \underline{Z} \,.$$

Since an H^- matrix required for the analysis is calculated in finding a C_1 matrix in the manner about to be discribed, much may be said in favor of the general form for the C_1 matrix to be given in the next section.

2.3 A General Form For Calculating C_α The Coefficient Matrix In The Reduced Normal Equations For $\underline{p}_\alpha$

There are certain advantages in finding a coefficient matrix, C_α : (1) the variances and covariances of estimable functions are easier to find, since a generalized inverse is easier to find for the reduced normal equations; (2) the s.s. due to a certain factor can be obtained easily and directly, rather than fitting the full and conditional models; (3) for additive models, the rank of the reduced form as one of several methods (Chakrabarti, 1963) determines the number of estimable functions available; (4) for some interaction models, the main effect and interaction effect estimates may be easier to find.

Without loss of generality, we shall develop a general form for C_1, the coefficient matrix for $\underline{p}_1$. Let us generalize on developments (2.1.6) to (2.1.11). Let Δ be a linear unbiased estimate for some linear function in $\underline{p}_1$ only, i.e. $E(\Delta)$ does not contain any other factor effects. We represent Δ then as

$$\Delta = \underline{q}_1' X_1'\underline{y} + \underline{q}_2'X_2'\underline{y} + \underline{q}_3'X_3'\underline{y} + \ldots + \underline{q}_m'X_m\underline{y}$$

where as before, $\underline{q}_i$ are vectors of real numbers.

$$\underline{q}_i' = (q_{i1}, q_{i2}, \ldots, q_{in_i}) \qquad i=1,\ldots,m.$$

We have already found $E(\Delta)$ for $m = 2$, and calculated C_1. It follows then for general m,

$$(2.3.1) \qquad E(\Delta) = (\underline{q}_1'X_1' + \underline{q}_2'X_2' + \ldots + \underline{q}_m'X_m')\, X_1\underline{p}_1$$
$$+ (\underline{q}_1'X_1' + \underline{q}_2'X_2' + \ldots + \underline{q}_m'X_m')\, X_2\underline{p}_2$$
$$+ (\underline{q}_1'X_1' + \underline{q}_2'X_2' + \ldots + \underline{q}_m'X_m')\, X_m\underline{p}_m \; .$$

If $E(\Delta)$ is to be free of $\underline{p}_i$, $2\leq i\leq m$, the following equations must be satisfied regardless of the true value of $\underline{p}_i$, $2 \leq i \leq m$,

$$(2.3.2) \qquad \underline{q}_1'H_{1i} + \underline{q}_2'H_{2i} + \ldots + \underline{q}_m'H_{mi} = 0$$

where

$$H_{\ell t} = X_\ell'X_t = (h_{\lambda\tau}^{\ell t}) \qquad 2\leq\ell,\ t\leq m$$

with

$$h_{\lambda\tau}^{\ell t} = \sum_i \ldots \sum_k \sum_n \ldots \sum_z h_{i\ldots k\lambda\ldots n\tau\ldots z}$$

and λ,τ being the subscript in the ℓ and t position, respectively ($\lambda = 1,2,\ldots,n_\ell$; $\tau = 1,2,\ldots,n_t$). Note that $H_{\lambda\lambda}$ are diagonal matrices.

For the sake of notation, let $H = (H_{\ell t})$, $2\leq\ell$, $t\leq m$.

From (2.3.2) we then have

$$(2.3.3) \qquad (\underline{q}_2', \ldots, \underline{q}_m') \begin{bmatrix} H_{22} & \ldots H_{2m} \\ \vdots & \vdots \\ H_{m2} & \ldots H_{mm} \end{bmatrix} = -\underline{q}_1' \, (H_{12} \ldots H_{1m})$$

or

$$(2.3.4) \qquad (\underline{q}_2', \ldots, \underline{q}_m') = -\underline{q}_1' \, (H_{12} \ldots H_{1m}) \begin{bmatrix} H_{22} & \ldots & H_{2m} \\ \vdots & & \vdots \\ H_{m2} & \ldots & H_{mm} \end{bmatrix}^{-}$$

where a generalized inverse for H is written as

$$(2.3.5) \qquad \begin{bmatrix} H_{22} & \ldots & H_{2m} \\ \vdots & & \vdots \\ H_{m2} & \ldots & H_{mm} \end{bmatrix}^{-} = \begin{bmatrix} M_{22} & \ldots & M_{2m} \\ \vdots & & \vdots \\ M_{m2} & \ldots & M_{mm} \end{bmatrix}$$

with $M_{\ell t}$ of order $n_\ell \times n_t$. The constant vectors $\underline{q}_i'$, $i>1$, are then

$$(2.3.6) \qquad \underline{q}_i' = -\underline{q}_1' \, (H_{12} M_{2i} + H_{13} M_{3i} + \ldots + H_{1m} M_{mi}).$$

From equation (2.3.1) in connection with (2.3.6) we obtain

$$(2.3.7) \qquad E(\Delta) = (\underline{q}_1' X_1 + \underline{q}_2' X_2 + \ldots + \underline{q}_m' X_m) \, X \, \underline{p}_{1-1}$$

$$= \underline{q}_1' \, (H_{11} - H_{12} M_{22} H_{21} - H_{13} M_{32} H_{21} - \ldots - H_{1m} M_{mm} H_{m1}) \underline{p}_{-1}$$

so that, similar to (2.1.15) $E(\Delta)$ can be written as

(2.3.8) $$E(\Delta) = \underset{-1}{q}'\, C_1 \underset{-1}{p}$$

with

(2.3.9) $$\begin{aligned} C_1 = H_{11} &- (H_{12}M_{22} + H_{13}M_{32} + \ldots + H_{1m}M_{m2})\, H_{21} \\ &- (H_{12}M_{12} + H_{13}M_{33} + \ldots + H_{1m}M_{m3})\, H_{31} \\ &- \vdots \qquad\qquad\qquad\qquad\qquad\qquad \vdots \\ &- (H_{12}M_{2m} + H_{13}M_{3m} + \ldots + H_{1m}M_{mm})\, H_{m1} \end{aligned}$$

or

(2.1.10) $$C_1 = H_{11} - (H_{12}\, H_{13} \ldots H_{1m})\ H^{-}\ (H_{12}\, H_{13} \ldots H_{1m})'$$

As in section 2.1, the reduced normal equations for $\underset{-1}{p}$ are given by $C_1 \underset{-1}{\hat{p}} = \underset{-1}{Q}$, and we now need to calculate the corresponding $\underset{-1}{Q}$ vector.

We may develop $\underset{-1}{Q}$ in a manner similar to section 2.1 and (2.3.6). We have

(2.3.11) $$\Delta = \underset{-1}{q}'\,\underset{-1}{Y} + \underset{-2}{q}'\,\underset{-2}{Y} + \ldots + \underset{-m}{q}'\,\underset{-m}{Y}$$

or

(2.3.12) $$\begin{aligned} \Delta = \underset{-1}{q}' \{\underset{-1}{Y} &- (H_{12}M_{22} + H_{13}M_{32} + \ldots + H_{1m}M_{m2})\, \underset{-2}{Y} \\ &\cdots\cdots\cdots\cdots\cdots \\ &- (H_{12}M_{2m} + H_{13}M_{3m} + \ldots + H_{1m}M_{mm})\, \underset{-m}{Y}\} \end{aligned}$$

and

(2.2.13) $$\Delta = \underset{-1}{q}'\, \underset{-1}{Q}$$

with

$$(2.3.14) \qquad \underset{\sim}{Q}_1 = \underset{\sim}{Y}_1 - (H_{12}\ H_{13} \ldots H_{1m})\ H^- \ (\underset{\sim}{Y}_2\ \underset{\sim}{Y}_3 \ldots, \underset{\sim}{Y}_m).$$

In order to calculate C_1 and $\underset{\sim}{Q}_1$ one therefore needs to find a generalized inverse for H. Also, to complete the analysis, a generalized inverse for C_1 is needed. Therefore, methods for computing generalized inverses are given in the following section.

2.4 Calculating Generalized Inverses

Among the many classes of generalized (g-inverse) or pseudo-inverses best summarized by Rhodes (1964) and later by Rao (1967) we shall consider only the "weak" form.

Definition 2.1.: A g-inverse of a matrix G of order k x k is a matrix of order k x k denoted by G^-, such that for any $\underline{y}$ for which $G\underline{x} = \underline{y}$ is consistent, $\underline{x} = G^- \underline{y}$ is a solution.

For our purposes if G^- is a g-inverse for G then $G\, G^-\, G = G$ will hold and conversely (see Rao, 1967).

The next two theorems are quite programmable and are compatible with section 2.3.

Theorem 2.1 Let B be a singular square symmetric matrix, k x k, that may be partitioned into

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

where B_{11} is $k_1 \times k_1$, B_{22} is $k_2 \times k_2$, $B_{12} = B'_{21}$ is $k_1 \times k_2$ and $k_1 + k_2 = k$.

If B_{11}^- is a g-inverse for B_{11}, then a g-inverse for B is

$$B = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}$$

where

$$M_{22} = (B_{22} - B_{21} B_{11}^- B_{12})^-$$

$$M_{12} = -B_{11}^- B_{12} M_{22} = M'_{21}$$

$$M_{11} = B_{11}^- - B_{11}^- B_{12} M_{21} .$$

This theorem is taken directly from Rhodes (1964) and has great value in inverting H in (2.3.10) and (2.3.5), since $H_{\ell\ell}$, $\ell=2,\ldots,m$, are non-singular matrices and hence $H_{\ell\ell}^- = H_{\ell\ell}^{-1}$. For calculating M_{11} in the above theorem we may repeat the use of this theorem until a unique inverse is attained or use the following theorems.

<u>Theorem 2.2</u>: If B is a singular (n x n) matrix of rank n - 1, then a generalized inverse for B is

$$(B + aJ_{nn})^{-1}$$

where a is any non-zero real number and J_{nn} (n x n) is a unity matrix.

A proof for this theorem is given essentially by Shah (1959) (see also Hinkelmann, 1968a). This latter theorem is most appealing, since practical considerations often require that all contrasts in desired factor levels be estimable, which means that C_1 and often other matrices used in calculating C_1, be of rank one less than their order. Also, this theorem permits simple computer programming. Another theorem that lends itself to programming is the following given by Bose (1958).

Theorem 2.3: Let X be a design matrix for a multidimensional design of factors F_1, F_2..., F_m, with factor F_i having n_i levels. If the degrees of freedom associated with the i-th factor is n_i-1, then a generalized inverse for X'X is $(X'X + \Gamma)^{-1}$ where

$$\Gamma = \text{DIAG}(\Theta_1 J_{n_1 n_1}; \Theta_2 J_{n_2 n_2}; \ldots; \Theta_{m-1} J_{n_{m-1} n_{m-1}}, O_{n_m n_m})$$

with $O_{n_m n_m}$ a n_m x n_m matrix of zeros, $\Theta_1, \Theta_2, \ldots, \Theta_{m-1}$ real numbers, and $J_{n_i n_i}$ are n_i x n_i unity matrices.

Theorems 2.1, 2.2, and 2.3 are particularly useful for inverting the coefficient matrix for the r.n.e. This matrix will be used in analysis for m factors, with and without interaction present among them. The latter case (without interaction) has essentially been presented in this chapter. The former case will now be discussed.

CHAPTER III

THE ANALYSIS OF DESIGNS WHERE SOME TWO-FACTOR INTERACTIONS ARE ASSUMED

An analysis of incomplete designs where no interaction is assumed was presented in Chapter II. This analysis can be applied to all such incomplete designs, and therefore applies to those cases where the works of Potthoff (1958), Bose and Srivastava (1964b) and Anderson (1968) do not apply.

Using the approach to solving the normal equations presented in section 2.3, we shall now present an analysis of experimental plans where certain types of two-factor interactions are assumed. An example was given in Chapter I where a price-juice interaction was suspected. The model in this case would contain a vector of parameters for the effects of the price-juice interaction.

As in Chapter I, our attention will be focused on those plans where not all factor level combinations are present. These plans are then incomplete, i.e. are fractional factorials. Further, these fractions can be both asymmetric and irregular, having factors of different levels and having these levels occur with non-proportional frequencies.

Bose and Srivastava (1964a) discuss an approach to the analysis of irregular fractions of $2^{k_1} \times 3^{k_2}$ factorials. However, this work is difficult to apply, and the analysis of irregular fractions with levels greater than three has not been discussed. Potthoff (1958, 1963), Causey (1967), and Anderson (1968) considered the analysis of asymmetric fractions.

However, their work concentrated on designs having partial balance on the factor levels involved. We will investigate the analysis of designs where balance on factor levels is not necessarily present.

Further, the analysis presented in this chapter provides for the possibility of investigating the main effects in the presence of certain two factor interactions. By presenting this analysis, we do not intend to imply that such analysis shall always be important, but just note here that for some types of interaction, it is indeed useful to investigate main effects. In those cases where it is not useful to do so one investigates "simple" main effects (Cochran and Cox, 1957).

This chapter shall discuss the following: section 3.1 shall present an analysis of a two-factor model with an assumed two-factor interaction and shall introduce the analysis of m factor models with two-factor interactions; section 3.2 presents an analysis of an m-factor model with one interaction; section 3.3 discusses models with two interactions; and section 3.4 discusses the analysis for models with three or more two-factor interactions. Also discussed in section 3.4 is a "weakness" of the approach to the analysis presented in this chapter.

3.1 The Two-Factor Design With Interaction

The analysis of a two-factor model with interaction present for an incomplete design is presented by Scheffé (1959). However, we shall just mention these results and introduce notation that shall serve us later. Our model is now

(3.1.1) $\quad E(y_{ij\omega}) = p_{12}^{ij} \qquad i = 1, \ldots, n_1;\ j = 1, \ldots, n_2;\ \omega = o, \ldots, h_{ij};$

where $p_{12}^{ij} = \mu + p_1^i + p_2^j + t_{12}^{ij}$ and t_{12}^{ij} represents the effect of interaction between factors one (F_1) and two (F_2) at levels i and j respectively. In matrix notation our model can be written as

(3.1.2) $\quad E(\underset{\sim}{y}) = X_{(12)}\, \underset{\sim}{p}_{12}$

Where $\underset{\sim}{y}$ is h x 1 and $h = \sum_i \sum_j h_{ij}$. If w_{12} is the number of level combinations of F_1 and F_2 that appear in $\underset{\sim}{y}$, then p_{12} is w_{12} x 1. We shall present p_{12} as if all $n_1\ n_2$ level combinations do appear, with the understanding that those level combinations which do not appear shall be deleted from p_{12}. This vector is then

$$\underset{\sim}{p}'_{12} = (p_{12}^{11}, p_{12}^{12}, \ldots, p_{12}^{1n_2}, \ldots, p_{12}^{n_1n_2})$$

and $X_{(12)}$ is the design matrix, h x w_{12}, for $\underset{\sim}{p}_{12}$. The parameter vector $\underset{\sim}{t}_{12}$ then is defined as $\underset{\sim}{t}'_{12} = (t_{12}^{11}, t_{12}^{12}, \ldots, t_{12}^{1n_2}, \ldots, t_{12}^{n_1n_2})$, 1 x w_{12}.

For the corresponding sum of squares (s.s.) for error and for the parameter $\underset{\sim}{p}_1$, we will use the notation introduced in Chapter II, ss_e and ss_{p_1} respectively. The s.s. due to the parameter vector $\underset{\sim}{t}_{12}$ will be expressed as $ss_{t_{12}}$.

Let $H_{(12)(12)} = X'_{(12)}X_{(12)}$ and $X'_{(12)}\,\underset{\sim}{y} = \underset{\sim}{Y}_{(12)}$ Then we may ex-

the s.s. due to $\underset{\sim}{p}_{12}$ as

$$ss_{p_{12}} = Y'_{(12)}H^{-1}_{(12)(12)}\underline{Y}_{(12)}, \text{ with d.f.} = w_{12}.$$

Note that this s.s. is the s.s. due to the combined main effects and interaction effects of factors one and two. If the conditional model is (2.1.2) the s.s. due to $\underline{t}_{12}$ is

$$\begin{aligned} ss_{t_{12}} &= ss_{ce} - ss_{e} \\ &= (\underline{y}'\underline{y}-\hat{\underline{p}}_1'Q_1 - Y_2' H_{22}^{-1} \underline{Y}_2) - (\underline{y}'\underline{y} - ss_{p_{12}}) \\ &= Y'_{(12)}H^{-1}_{(12)(12)}\underline{Y}_{(12)} - \underline{Y}_2' H_{22}^{-1} \underline{Y}_2 - \hat{\underline{p}}_1' \underline{Q}_1 \end{aligned}$$

with $w_{12} - n_1 - n_2 + r_1$ d.f.

<u>The S. S. Due to Main effects and Interaction</u>

If the d.f. for $ss_{t_{12}}$ are $(n_1 - 1)(n_2 - 1)$ or if $w_{12} = n_1 n_2$ one may use the approach of Scheffe (1959, p. 118), or proceed as Potthoff (1958, 1963) and Anderson (1968). The s.s. due to main effects is then found from the s.s. due to contrasts in the following parameters, $f_{i.}$ and $f_{.j}$:

$$f_{i.} = \bar{p}^{i.}_{12} - \bar{p}^{..}_{12}$$

(3.1.3)

$$f_{.j} = \bar{p}^{.j}_{12} - \bar{p}^{..}_{12}$$

where
$$\bar{p}^{i\cdot}_{12} = \frac{1}{n_2} \sum_j p^{ij}_{12}$$
$$\bar{p}^{\cdot j}_{12} = \frac{1}{n_1} \sum_i p^{ij}_{12}$$
$$\bar{p}^{\cdot\cdot}_{12} = \frac{1}{n_1 n_2} \sum_i \sum_j p^{ij}_{12}$$

Note that if one applied the restrictions

$$(3.1.4) \qquad \sum_i p^i_1 = \sum_j p^j_2 = \sum_i t^{ij}_{12} = \sum_j t^{ij}_{12} = 0,$$

then $f_{i\cdot} = p^i_1$ and $f_{\cdot j} = p^j_2$. Potthoff (1958, p.60) proved that the expression for the s.s. due to contrasts in $f_{i\cdot}$ and in $f_{\cdot j}$ is the same as the expression given by Scheffé (1959, p. 118) for the s.s. due to main effects. Also, the s.s. due to contrasts in the following parameters,

$$(3.1.5) \qquad f_{ij} = p^{ij}_{12} - \bar{p}^{i\cdot}_{12} - \bar{p}^{\cdot j}_{12} + \bar{p}^{\cdot\cdot}_{12}$$

when $w_{12} = n_1 n_2$, is the s.s. due to t^{ij}_{12} since by applying (3.1.4), $f_{ij} = t^{ij}_{12}$.

Now if the d.f. for $ss_{t_{12}}$ are not $(n_1 - 1)(n_2 - 1)$, or if $w_{12} \neq n_1 n_2$ we can proceed to find a test on main effects as suggested by Elston and Bush (1961). Elston and Bush (1961) indicated how some comparisons on the true cell means, i.e. $E(y_{ij\omega})$, can be used such that tests on first order interaction and on main effects in the presence of assumed interaction can be made. This procedure amounts to testing main effects when there are no observations in some assemblies. We shall refer to Elston and Bush (1961) later in discussing m-factor models.

The Analysis of m-Factor Models

An analysis summary of m-factor models with interaction present will first be given in order to better acquaint the reader with later developments. The following models shall be referred to:

$$(1)\quad E(\underline{y}) = X_{(12)}\underline{p}_{12} + X_3\underline{p}_3 + X_4\underline{p}_4 + \dots + x_m\underline{p}_m$$

$$(2)\quad E(\underline{y}) = X_{(12)}\underline{p}_{12} + X_{(34)}\underline{p}_{34} + X_5\underline{p}_5 + \dots + X_m\underline{p}_m$$

$$(3)\quad E(\underline{y}) = X_{(12)}\underline{p}_{12} + X_{(13)}\underline{p}_{13} + X_4\underline{p}_4 + \dots + X_m\underline{p}_m$$

$$(4)\quad E(\underline{y}) = X_{(12)}\underline{p}_{12} + X_{(13)}\underline{p}_{13} + X_{(23)}\underline{p}_{23} + X_4\underline{p}_4 + \dots + X_m\underline{p}_m$$

In each case we solve the r.n.e. for $\underline{p}_{12}$ in order to obtain C_{12}, Q_{12}; and an estimate $\hat{\underline{p}}_{12}$. The s.s. due to $\underline{p}_{12}$ (not $\underline{t}_{12}$) is then $ss_{p_{12}} = \hat{\underline{p}}_{12}'\ \underline{Q}_{12}$.

The unadjusted estimates for p* are found as in section 2.2., and the unadjusted s.s., $\underline{Y}^{*\prime}H\ \underline{Y}^*$, then follows. (The nature of $\underline{p}^*$, $\underline{Y}^*$, and H will change with each model). The s.s. due to error in each case will be expressed as $\underline{y}'\underline{y} - \underline{Y}^{*\prime}H^-\underline{Y}^* - \hat{\underline{p}}_{12}'\ \underline{Q}_{12}$.

Some interaction s.s. and main effect s.s. can be found either by subtraction or from the s.s. due to contrasts in "f" parameters similar to (3.1.3) defined in terms of p_{12}^{ij}. Other interaction and main effect s.s. can be found from the s.s. due to contrasts involving parameters p_{23}^{ik}, or p_{23}^{jk}, etc. In each case the corresponding estimates are found from a set of reduced normal equations. Without loss of generality we shall present estimates for the effects involving F_1 and F_2. Other estimates may be found similarly. The corresponding s.s. will be presented for each model for the two cases: (a) when all the d.f. for $ss_{p_{12}}$ are available and (b) when not all the d.f. for $ss_{p_{12}}$ are not available. The d.f. available for $ss_{p_{12}}$ may be checked either by inspecting the rank of C_{12} or by methods of Chapters IV and V.

For the analysis of models (3) and (4), we will present a reduced form of the normal equations for $\underline{p}_{12}$ for each level, i, of F_1. The intent here is to simplify the solution to the normal equations.

3.2 The m-Factor Design With One Interaction

The Analysis of Model (1)

Assume that we have an m-factor incomplete design where, without loss of generality, interaction between factors one and two may be present. A typical observation is represented by $y_{ijk\ldots z\omega}$ where the notation of section 2.2 applies to our observations and the parameters p_u^{ℓ}, $\ell = 1,\ldots,n_u$, $u = 1,\ldots m$. The model is now

$$E(y_{ijk\ldots z\omega}) = p_{12}^{ij} + p_3^{k} + \ldots + p_m^{z} \tag{3.2.1}$$

where $p_{12}^{ij} = p_1^{i} + p_2^{j} + t_{12}^{ij}$. Note that p_{12}^{ij} does not include μ as it did in section 3.1. Using the notation of 2.2 and 3.1, (3.2.1) may be expressed as model (1):

$$(3.2.2) \qquad E(\underline{y}) = X_{(12)}\underline{p}_{12} + X_3\underline{p}_3 + \ldots + X_m\underline{p}_m .$$

The above model is similar to an (m-1)-factor model with no interaction. Therefore, the analysis follows that of Chapter II. Let $X'_{(12)}X_{(12)} = H_{(12)(12)}$ and $X'_{(12)}\, X_u = H_{(12)\,u} = H'_{u(12)}$, $w_{12} \times n_u$, for $u = 3,\ldots,m$.

The reduced normal equations (r.n.e.) for the parameter $\underline{p}_{12}$ are then

$$C_{12}\,\hat{\underline{p}}_{12} = \underline{Q}_{12}$$

where C_{12} is similar to (2.3.10) and $\underline{Q}_{12}$ is similar to (2.3.14). Note that the matrix H in (2.3.10) is now

$$(3.2.3) \qquad H = \begin{bmatrix} H_{33} & H_{34} & \cdots & H_{3m} \\ \cdot & \cdot & \cdots & \cdot \\ H_{m3} & H_{m4} & \cdots & H_{mm} \end{bmatrix}$$

$C_{(12)}$ is $w_{12} \times w_{12}$ of rank $w_{12} - r_{12}$, where r_{12} is an integer, $1 \leq r_{12} \leq w_{12} - 1$. The s.s. due to $\underline{p}_{12}$ is $ss_{p_{12}} = \hat{\underline{p}}'_{12}\,\underline{Q}_{12}$, with d.f. $= w_{12} - r_{12}$.

<u>The S.S. due to Main effects And Interaction</u>

If the d.f. available for $ss_{p_{12}}$ are $n_1 n_2 - 1$, i.e. $w_{12} - r_{12} = n_1 n_2 - 1$

and, or equivalently, if the d.f. for $ss_{t_{12}}$ is $(n_1 - 1)(n_2 - 1)$, we can define the parameters f_{ij}, $f_{i.}$, and $f_{.j}$ as in (3.1.3) and (3.1.5). The s.s. for interactions, $ss_{t_{12}}$, follows from the s.s. due to contrasts in f_{ij}. The s.s. for main effects under the presence of interaction follows from contrasts in $f_{i.}$ and $f_{.j}$. For example, an estimate of $t_{12}^{11} + t_{12}^{22} - t_{12}^{21} - t_{12}^{12} = \underline{L}'\underline{t}_{12}$ is then $\hat{f}_{11} + \hat{f}_{22} - \hat{f}_{21} - \hat{f}_{12}$ or alternatively $\hat{p}_{12}^{11} + \hat{p}_{12}^{22} - \hat{p}_{12}^{21} - \hat{p}_{12}^{12} = \underline{L}'\hat{\underline{p}}_{12}$. The s.s. due to $\underline{L}'\underline{t}_{12}$ is then

$$(3.2.4) \qquad (\underline{L}'\hat{\underline{p}}_{12})' \, [\underline{L}' C_{12}^{-} \underline{L}]^{-1} \, (\underline{L}'\hat{\underline{p}}_{12})$$

where C_{12}^{-} is a g-inverse for C_{12}. The main effect contrasts in $\underline{p}_1$ and $\underline{p}_2$, and their corresponding s.s. are also found from contrasts in $\underline{p}_{12}$. Appendix I presents a numerical example of such an analysis.

If the d.f. available for $ss_{p_{12}}$ is not $n_1 n_2 - 1$ but are $w_{12} - r_{12} \neq n_1 n_2 - 1$, or equivalently, the d.f. for $ss_{t_{12}}$ are not $(n_1 - 1)(n_2 - 1)$, there still may be comparisons on two-factor interactions and main effects that can be tested. First, it still may be possible to define meaningful effects in terms of the "f" parameters, f_{ij}, $f_{i.}$, and $f_{.j}$, as in section 3.1. If so, then the s.s. due to individual contrasts may be found as before. However, if it is not possible to define such meaningful contrasts, it still may be able to compare factor level combinations as do Elston and Bush (1961).

Further, if $w_{12} - r_{12} \neq n_1 n_2 - 1$, the s.s. due to $\underline{t}_{12}$ can be obtained by subtraction as in section 3.1. Let the error s.s. under the conditional model given by (2.2.1) be $ss_{e(1.2.3)}$. Let the error s.s. from (3.2.2) be denoted by $ss_{e(12.3)}$. From a solution to the r.n.e. for $\underline{p}_{12}$, $ss_{e(12.3)} = \underline{y}'\underline{y} - \underline{Y}^{*\prime}H^{-}\underline{Y}^{*} - \hat{\underline{p}}_{12}'\underline{Q}_{12}$ with d.f. $= h - w_{12} + r_{12} - r$, where H is of rank r, and is given in (3.2.3), and $\underline{Y}^{*\prime}H^{-}\underline{Y}^{*}$ is the unadjusted s.s. found as in section 2.2. Then $ss_{t_{12}} = ss_{e(1.2.3)} - ss_{e(12.3)}$.

Testing the absence of F_1 and Interaction

Also using the conditional model $E(\underline{y}) = X_2\underline{p}_2 + X_3\underline{p}_3 + \ldots + X_m\underline{p}_m$ and (3.2.2) one could test $\underline{p}_1 = \underline{0}$ and $\underline{t}_{12} = \underline{0}$, or simply the absence of factor one and of the first order interaction. This may be especially appealing if one would desire to test the main effects of factor one when not all level combinations of factors one and two are present. Let the error s.s. under conditional model be denoted by $ss_{e(2.3)}$ and the error s.s. under (3.2.2) be denoted by $ss_{e(12.3)}$. Then the s.s. due to $\underline{p}_1 = \underline{0}$ and $\underline{t}_{12} = \underline{0}$ would be $ss_{p_1, t_{12}} = ss_{e(2.3)} - ss_{e(12.3)}$. The d.f. for $ss_{p_1, t_{12}}$ are also received by subtraction. A similar procedure can be used for testing the absence of F_2 and interaction.

This approach may be needed later in discussing main effects which are defined by more parameters than p_{12}^{ij}, since using our form of the reduced normal equations for $\underline{p}_{12}$ only, and corresponding g-inverse C_{12}^{-}, we do not have the variance-covariance matrix for other parameters. More shall be said on this in section 3.4.

The Analysis of Factors Other Than F_1 and F_2

The analysis on factors one and two is now complete. For individual tests on factors three to m, one can obtain a set of reduced normal equations for each factor, solve these equations and obtain the respective s.s. for each factor. For example, to adjust $\underline{p}_3$ for the parameters $\underline{p}_{12}$, $\underline{p}_4$, ..., $\underline{p}_m$, a coefficient matrix for the reduced normal equations is found by interchanging the integer subscripts (12) and 3 in the expression for C_{12}.

Note that since μ is associated with $\underline{p}_m$, we redefine our model so that μ is associated with $\underline{p}_{m-1}$ before the matrix C_m is calculated for factor m. This terminates our discussion of m factor models with one interaction assumed.

3.3 The m-Factor Design With Two Interactions

The Analysis of Models (2) and (3)

For models of the form $E(y_{ijk\ell...z\alpha}) = p_{12}^{ij} + p_{34}^{k\ell} +...+p_m^z$, model (2), where two first order interactions are present and no factor interacts with another more than once, the analysis follows the methods of Chapter II, and sections 3.1 and 3.2. The analysis of model (3) where a factor interacts with two other factors given by

$$(3.3.1) \qquad E(y_{ijk\;...z\alpha}) = p_{12}^{ij} + p_{13}^{ik} + p_4^{\ell} + ... + p_m^z$$

shall be discussed in this section. Without loss of generality we dis- discuss this in terms of model (3.3.1). The parameters are similar to those defined in section 3.2 with the exception that now

$$p_{12}^{ij} = \frac{1}{2}\, p_1^i + p_2^j + t_{12}^{ij}$$

$$p_{13}^{ik} = \frac{1}{2}\, p_1^i + p_3^k + t_{13}^{ik}$$

Let w_{13} be the number of factor level combinations from F_1 and F_3 that actually occur. Let $\underline{p}'_{13} = (p_{13}^{11}, p_{13}^{12}, \ldots, p_{13}^{1n_3}, \ldots, p_{13}^{n_1 n_3})$ be of order $1 \times w_{13}$, and $X_{(13)}$ the $h \times w_{13}$ corresponding design matrix for $\underline{p}_{13}$. Then, with $\underline{p}_{12}$ and $X_{(12)}$ as defined before (3.3.1) may be expressed as

$$E(\underline{y}) = X_{(12)}\underline{p}_{12} + X_{(13)}\underline{p}_{13} + X_4\underline{p}_4 + \ldots + X_m\underline{p}_m \; . \tag{3.3.2}$$

This model is somewhat similar to an (m-2)-factor model with no interaction present and the analysis of this model, then, will follow procedures similar to previous development. However, there will be some different steps taken in our approach.

For a fixed level of F_1, i, we will find an expression for the co-coefficient matrix of the estimate $\hat{\underline{p}}_{12}$ in the r.n.e. for F_2. This coefficient matrix shall be expressed as $C_{12}^{i\cdot}$. The reduced normal equations for $\underline{p}_{12}$ can then be expressed as sets of equations $C_{12}^{i\cdot}\, \underline{p}_{12}^{i\cdot} = \underline{Q}_{12}^{i\cdot}$, $i = 1, \ldots, n_1$, where the vectors $\underline{p}_{12}^{i}$ and $\underline{Q}_{12}^{i\cdot}$ are similar to those defined in section 2.3. (We shall define these vectors explicitly in the following developments.) An expression for this coefficient matrix, $C_{12}^{i\cdot}$ shall be given first. We need to define the matrices involved. Let $X_{(12)}^{(i\cdot)}$ be a design matrix that is obtained from $X_{(12)}$ by fixing the i-th level of factor F_1. Note that $X_{(12)}$ is $h \times w_{12}$ and

$$X_{(12)} = (X_{(12)}^{(1.)}, X_{(12)}^{(2.)}, \ldots, X_{(12)}^{(n_1.)})$$

and $X_{(12)}^{(i.)}$ is $h \times n_{12}^{i.}$ where $n_{12}^{i.}$ is the number of levels of factor F_2 that appear with the i-th level of factor F_1, and $\sum_i n_{12}^{i.} = w_{12}$. Then $H_{(12)(12)}^{(i.)(i.)} = X_{12}^{(i.)'} X_{(12)}^{(i.)}$ is an $n_{12}^{i.} \times n_{12}^{i.}$ diagonal matrix with elements $h_{i1}, h_{i2}, \ldots, h_{in_{12}^{i.}}$ where $h_{ij} = \sum_k \sum_\ell, \ldots, \sum_z h_{ijk\ell\ldots z}$, and $X_{(12)}^{(i.)'} X_{(13)} = H_{(12)(13)}^{(i.)}$, $X_{(12)}^{(i.)'} X_u = H_{(12)u}^{(i.)}$ for $u = 4, \ldots, m$. Let $X_u' X_v = H_{uv}$, $u, v = 4, \ldots, m$, $X'_{(13)} X_u = H_{(13)u}$, and $X'_{(13)} X_{(13)} = H_{(13)(13)}$, then

$$H = \begin{bmatrix} H_{(13)(13)} & H_{(13)4} & \cdots & H_{(13)m} \\ \cdot & \cdot & \cdots & \cdot \\ H_{m(13)} & H_{m4} & \cdots & H_{mm} \end{bmatrix} \tag{3.3.3}$$

Given a g-inverse H^-, the r.n.e. for F_2 with i fixed, follow in a manner similar to section 2.3 and the coefficient matrix is calculated in a form similar to (2.3.10).

A formal presentation of the development for $C_{12}^{i.}$, is because of the notation involved, difficult to follow. However, this development is quite similar to that in section 2.3 and is therefore, not really necessary. From (2.3.10), fixing the i-th level of F_1, we have the expression for $C^{i.}_{12}$,

$$C^{i.}_{12} = H^{(i.)(i.)}_{(12)(12)} - [H^{(i.)}_{(12)(13)} H^{(i.)}_{(12)4} \cdots H^{(i.)}_{(12)m}]\; H^{-} [H^{(i.)}_{(12)(13)} H^{(i.)}_{(12)4} \cdots H^{(i.)}_{(12)m}]'.$$

$C^{i.}_{12}$ is a $n^{i.}_{12} \times n^{i.}_{12}$ matrix of rank $n^{i.}_{12} - r^{i.}_{12}$, where $r^{i.}_{12}$ is an integer function of i. $C^{i.}_{12}$ is a coefficient matrix for $\underline{p}^{i.}_{12}$, where $\underline{p}^{i.}_{12}$ is given by

$$\underline{p}^{i.}_{12} = (p^{i1}_{12}, p^{i2}_{12}, p^{i3}_{12}, \ldots, p^{in^{i.}_{12}}_{12})' .$$

The right side of the r.n.e. is similar to (2.3.14)

$$\underline{Q}^{i.}_{12} = X^{(i.)}_{(12)}\underline{y} - [H^{(i.)}_{(12)(13)} H^{(i.)}_{(12)4} \cdots H^{(i.)}_{(12)m}] H^{-} (X_{(13)} X_4 \cdots X_m)'\underline{y}$$

where corresponding to $\underline{p}^{i.}_{12}$, $\underline{Q}^{i.}_{12}$ is defined by

$$\underline{Q}^{i.}_{12} = (Q^{i1}_{12}, Q^{i2}_{12}, Q^{i3}_{12} \ldots Q^{in^{i.}_{12}}_{12})'$$

A solution to the r.n.e. is found for each i, $i=1,\ldots,n_1$, and a parameter $\hat{\underline{p}}_{12} = (\hat{p}^{1.}_{12}, \hat{p}^{2.}_{12}, \ldots, \hat{p}^{n_1 .}_{12})$ is now available for calculating the corresponding s.s. for testing $\underline{p}_{12} = \underline{0}$,

$$ss_{p_{12}} = \sum_i \hat{\underline{p}}^{i.}_{12}{}' \underline{Q}^{i.}_{12}, \text{with } \sum_i (n^{i.}_{12} - r^{i.}_{12}) \text{ d.f.}$$

Note that a g-inverse for C_{12} is found easily, since $C_{12} = \text{diag}\,(C^{1.}_{12}, C^{2.}_{12}, \ldots, C^{n_1 .}_{12})$. Let $ss_{e(12\cdot 13)}$ be the error s.s. under the model (3.3.2) and $ss_{e(13.4)}$ the error s.s. under the model

(3.3.4) $\quad E(\underline{y}) = X_{(13)}\underline{p}_{13} + X_4\underline{p}_4 + \ldots + X_m\underline{p}_m$.

This model is similar to (3.2.2) and may be considered a "conditional" model. Then $ss_{e(12.13)} = ss_{e(13.4)} - ss_{p_{12}}$. Note $ss_{e(13.4)}$ is easy to calculate, since it uses the H matrix of (3.3.3); $ss_{e(13.4)}$ is calculated in a manner similar to approach given in section 2.2. Let the design matrix $X = [X_{(13)}, X_4, \ldots, X_m]$ of (3.3.4) have rank a, then

$h - a - \sum_{i=1}^{n_1} (n_{12}^{i\cdot} - r_{12}^{i\cdot})$ are the d.f. for $ss_{e(12.13)}$.

<u>The S.S. due to Main Effects of F_2 and Interaction</u>

If the d.f. available for $ss_{p_{12}}$ are $n_1(n_2 - 1)$, i.e., if the rank of C_{12} is $\sum_{i=1}^{n_1} (n_{12}^{i\cdot} - r_{12}^{i\cdot}) = n_1(n_2 - 1)$, the main effects of factor two and the interaction effects of factors one and two may be received from

$$f_{\cdot j} = \bar{p}_{12}^{\cdot j} - \bar{p}_{12}^{\cdot\cdot}$$

$$f_{ij} = p_{12}^{ij} - \bar{p}_{12}^{\cdot j} - \bar{p}_{12}^{i\cdot} + \bar{p}_{12}^{\cdot\cdot}$$

where $\bar{p}_{12}^{\cdot j}$, $\bar{p}_{12}^{i\cdot}$, and $\bar{p}_{12}^{\cdot\cdot}$ are defined as in section 3.1. Note that by applying (3.1.4), $f_{\cdot j} = p_2^j$ and $f_{ij} = t_{12}^{ij}$. Then the s.s. contrasts of factor two of or the two-factor interaction may be received by (3.2.4).

If the d.f. available for $ss_{p_{12}}$ are not $n_1(n_2 - 1)$, then we proceed, as we did for the similar case in section 3.2. Some meaningful effects may still be defined in terms of $f_{.j}$, and f_{ij}, and the s.s. for these may be found by (3.3.4). Another approach could be that of Elston and Bush (1961). By comparing factor level combinations, we still may be able to define some meaningful tests. Further, the s.s. due to $\underline{t}_{12}$ (interaction), $ss_{t_{12}}$, can be found as the difference $ss_{e(13.2)} - ss_{e(12.13)}$, where $ss_{e(13.2)}$ is the error s.s. under

$$(3.3.5) \qquad E(y) = X_{(13)}\underline{p}_{13} + X_2\underline{p}_2 + X_4\underline{p}_4 + \ldots + X_m\underline{p}_m \ .$$

The d.f. for $ss_{t_{12}}$ is then $\sum_{i=1}^{n_1} [n_{12}^{i.} - r_{12}^{i.}] - b$, where b is the rank of the coefficient matrix in the r.n.e. for p_2 under (3.3.3). Note that if $\sum_{i=1}^{n} [n_{12}^{i.} - r_{12}^{i.}] = n_1(n_2 - 1)$ and $b = n_2 - 1$, then the d.f. for $ss_{t_{12}}$ are $(n_1 - 1)(n_2 - 1)$.

The S.S. Due to The Main Effects of F_1

In order to estimate main effect contrasts of F_1, we could proceed as follows. From the r.n.e. for $\underline{p}_{12}$, an estimate, $\hat{\underline{p}}_{12}$, can be found. Then, using the H matrix defined in (3.3.3) and the procedures of section 2.2, an estimate $\hat{\underline{p}}_{13}$ can be found. Let the d.f. available for F_1 be $n_1 - 1$. (This may be checked by methods of Chapter IV). Then the estimates of F_1 effects may be define in terms of

$$(3.3.6) \qquad f_{i.} = \bar{p}^{i.}_{12} + \bar{p}^{i.}_{13} - \bar{p}^{..}_{12} - \bar{p}^{..}_{13} .$$

If we apply (3.1.4), the $f_{i.}$ are then $p^{i}_{1}.$ (The parameters $\bar{p}^{i.}_{13}$, $\bar{p}^{..}_{13}$ are defined similar to $\bar{p}^{i.}_{12}$ and $\bar{p}^{.j}_{12}$ and $\bar{p}^{..}_{12}$ in section 3.1.) A test in $f_{i.} = 0$ would require the s.s. necessary to test the absence of factor one. Let this s.s. defined by ss_{fi}. Then $ss_{fi} = ss_{e(2.3)} - ss_{e(12.13)}$ where $ss_{e(2.3)}$ is the s.s. due to error under the conditional model

$$E(\underset{\sim}{y}) = X_2\underset{-}{p}_2 + X_3\underset{-}{p}_3 + \ldots + X_m\underset{-}{p}_m$$

The d.f. for ss_{fi} would then be $(n_1 -1) + (n_1 -1)(n_2 -1)+(n_1 -1)(n_3-1) = (n_1-1)(n_2 + n_3-1)$, [if the d.f. for $ss_{p_{12}}$ are $n_1(n_2-1)$, and the d.f. for $ss_{p_{13}}$ (due to $\underset{-}{p}_{13}$) are $n_1(n_3 - 1)$].

However, a test on the main effects of F_1 under the presence of interaction would involve a little more work. Note that the variances of these effects cannot be defined in terms of one g-inverse, C^{-}_{12}, as in section 3.2. This situation then, illuminates a weakness in the general approach developed in Chapter III for models of the type (3.3.2). (Note that this weakness is only for certain types of main effects.) In order to find the variances of estimates of effects in factor one, we can not proceed as before. However, the s.s. due to these effects still may be expressed. Let $X = [X_{(12)}, X_{(13)}, X_4,\ldots,X_m]$ be the design matrix for model (3.3.2). Let $\underset{-}{p}' = (\underset{-}{p}_{12}, \underset{-}{p}_{13}, \underset{-}{p}_4,\ldots,\underset{-}{p}_m)$ and let $\underset{-}{L}'\underset{-}{p}$ be a main effect contrast (i.e. defined in terms of $\underset{-}{p}_{12}$ and $\underset{-}{p}_{13}$). Then the s.s. due to this contrast is

$$(3.3.7) \qquad (\underline{L}'\hat{\underline{p}})'(\underline{L}'(X'X)^{-}\underline{L})^{-1}(\underline{L}'\hat{\underline{p}})$$

where $\hat{\underline{p}}$ is a solution to the normal equations and $(X'X)^{-}$ is a g-inverse for $(X'X)$. Further, the variance of $\underline{L}'\hat{\underline{p}}$ is $\underline{L}'(X'X)^{-}\underline{L}\sigma^2$. [These last results can be found in Rao, 1967].

If the d.f. available for factor one is not $n_1 - 1$, we still may be able to make some tests on factor one under the presence of first order interaction. For these tests we may be able to compare factor level combinations as do Elston and Bush (1961), as mentioned earlier.

This essentially covers the analyses of factors one and two under the assumed presence of two first order interactions. A similar approach may be used in the analysis of factors one and three. For tests on $\underline{p}_u$, $u = 4,\ldots,m$, the corresponding r.n.e. for each $\underline{p}_u$ may be found and estimates $\hat{\underline{p}}_u$ calculated. The necessary s.s. and variances of estimates may then be found from the g-inverses of C_u, $u = 4, \ldots m$.

3.4 The M-Factor Design With Three Interactions

The Analysis of Model (4)

The analysis of a design model with three first order interaction parameter vectors, as for example, Model (4):

$$(3.4.1) \qquad E(\underline{y}) = X_{(12)}\underline{p}_{12} + X_{(13)}\underline{p}_{13} + X_{(23)}\underline{p}_{23} + X_4\underline{p}_4 + \ldots + X_m\underline{p}_m$$

is similar to what has been presented. The design matrices $X_{(12)}$, $X_{(13)}$ and X_u, $u = 4,\ldots, m$ are defined as before, the parameters $\underline{p}_u$, are as before, and the parameters, $\underline{p}_{12}$, $\underline{p}_{13}$, and $\underline{p}_{23}$, w_{23} x 1 with design matrix $X_{(23)}$, h x w_{23}, are now defined from the following:

$$p_{12}^{ij} = \frac{1}{2} p_1^i + \frac{1}{2} p_2^j + t_{12}^{ij}$$

$$p_{13}^{ik} = \frac{1}{2} p_1^i + \frac{1}{2} p_3^k + t_{13}^{ik}$$

$$p_{23}^{jk} = \frac{1}{2} p_2^j + \frac{1}{2} p_3^k + t_{23}^{jk}$$

where t_{23}^{ik} is the interaction effect of F_2 and F_3 at levels j and k, and w_{23} different level combinations of F_2 and F_3 occur. One then solves the r.n.e. for $\underline{p}_{12}$ at fixed level, i, of F_1. The $C_{12}^{i.}$, $\underline{p}_{12}^{i.}$ and $\underline{Q}_{12}^{i.}$ matrices and vectors are similar to those presented for model (3.3.2). To find these explicitly, one follows the analysis of sections, 2.3 and 3.3.

The s.s. for testing $p_{12}^{ij} = 0$ is then $ss_{p_{12}} = \sum_i \hat{p}_{12}^{i.'} Q_{12}^{i.}$. This s.s. now corresponds to the s.s. for testing interaction effects (for testing $t_{12}^{ij} = 0$). If the d.f. for $ss_{t_{12}}$ are $\sum_i (n_{12}^{i.} - r_{12}^{i.})$, ($n_{12}^{i.}$ and $r_{12}^{i.}$ are defined in section 3.3) and are equal to $(n_1 - 1)(n_2 - 1)$, then $ss_{t_{12}}$ is the s.s. due to contrasts on $f_{ij} = p_{12}^{ij} - \bar{p}_{12}^{i.} - \bar{p}_{12}^{.j} + \bar{p}_{12}^{..}$, since applying conditions (3.1.4), $f_{ij} = t_{12}^{ij}$. Then one uses procedures given in section 3.2 to find the variances and necessary s.s. due to the interaction contrasts. A similar approach would be used for interaction contrasts involving p_{13}^{ik} or for contrasts involving p_{23}^{jk}.

If $\sum_i (n_{12}^{i.} - r_{12}^{i.})$ are not $(n_1 - 1)(n_2 - 1)$ (and this may be checked using methods of Chapter IV) then $ss_{t_{12}}$ may be found by subtraction. Let $ss_{e(13.23)}$ be the error s.s. under

$$E(\underline{y}) = X_{(13)}\underline{p}_{13} + X_{(23)}\underline{p}_{23} + X_4\underline{p}_4 + \ldots + X_m\underline{p}_m$$

and let the error s.s. for (3.4.1) be $ss_{e(12.13.23)}$. Then $ss_{t_{12}}$ = $ss_{e(13.23)} - ss_{e(12.13.23)}$ with $\sum_i (n_{12}^{i\cdot} - r_{12}^{i\cdot})$ d.f. A similar approach follows for interaction between F_1 and F_3 or between F_2 and F_3.

<u>The S.S. Due to Main Effects</u>

For estimates of main effects, we have a situation similar to section 3.3. If the d.f. for $ss_{t_{12}}$, $ss_{t_{13}}$, and $ss_{t_{23}}$, are respectively $(n_1 - 1)(n_2 - 1)$, $(n_1 - 1)(n_3 - 1)$, and $(n_2 - 1)(n_3 - 1)$, we can define the main effects of factors one, two, and three in terms of the contrasts

$$f_{i..} = \bar{p}_{12}^{i\cdot} + \bar{p}_{13}^{i\cdot} - \bar{p}_{12}^{\cdot\cdot} - \bar{p}_{13}^{\cdot\cdot}$$

$$f_{..k} = \bar{p}_{13}^{\cdot k} + \bar{p}_{23}^{\cdot k} - \bar{p}_{13}^{\cdot\cdot} - \bar{p}_{23}^{\cdot\cdot}$$

$$f_{.j.} = \bar{p}_{12}^{\cdot j} + \bar{p}_{23}^{j\cdot} - \bar{p}_{12}^{\cdot\cdot} - \bar{p}_{13}^{\cdot\cdot}$$

where $\bar{p}_{12}^{i\cdot}$, $\bar{p}_{13}^{i\cdot}$, $\bar{p}_{12}^{\cdot j}$, $\bar{p}_{13}^{\cdot k}$, $\bar{p}_{12}^{\cdot\cdot}$, and $\bar{p}_{13}^{\cdot\cdot}$ follow from before and

$$\bar{p}_{23}^{j\cdot} = \frac{1}{n_3} \sum_k p_{23}^{jk}$$

$$\bar{p}_{23}^{\cdot k} = \frac{1}{n} \sum_j p_{23}^{jk}$$

$$\bar{p}_{23}^{\cdot\cdot} = \frac{1}{n_2}\frac{1}{n_3}\sum_j\sum_k p_{23}^{jk}$$

Note that applying the restrictions,

$$\sum_i p_1^i = \sum_j p_2^j = \sum_i t_{12}^{ij} = \sum_j t_{12}^{ij} = \sum_i t_{13}^{ik} = \sum_k t_{13}^{ik} = \sum_j t_{23}^{jk} = \sum_k t_{23}^{jk} = \sum_k p_3^k = 0,$$

one then defines $f_{i..} = p_1^i$, $f_{.j.} = p_2^j$, $f_{..k} = p_3^i$.

We may approach the analysis of these effects as we did in section 3.3. An estimate $\hat{\underline{p}}_{12}$ can be found from the r.n.e. for $\underline{p}_{12}$. Estimates $\hat{\underline{p}}_{13}$ and $\hat{\underline{p}}_{23}$ then may be found by the procedure indicated in section 2.2, where using $\hat{\underline{p}}_{12}$ and the full normal equations, we find adjusted estimates $\hat{\underline{p}}_{12}$ and $\hat{\underline{p}}_{13}$. Therefore estimates $\hat{f}_{i.}$, $\hat{f}_{.j.}$ and $\hat{f}_{..k}$ can be found. As in section 3.2. we may test $f_{i..} = 0$, a test for the absence of factor one. But this test would involve $p_1^i = 0$, $t_{12}^{ij} = 0$, and $t_{13}^{ik} = 0$, which may seem a bit unrealistic. A similar case could be argued for $f_{.j.} = 0$ or $f_{..k} = 0$.

A more realistic approach to testing the main effects under the assumed presence of three first order interactions would obviously require a little more work. As in the case of the main effects of F_1 in section 3.3., the variances of the main effects for F_1, F_2, and F_3 cannot be defined in terms of a g-inverse for the coefficient matrix in the r.n.e. (i.e. C_{12}^-, C_{13}^-, or C_{23}^-). Therefore, we must proceed as we did in (3.3.7). Let $X = [X_{(12)}X_{(13)}X_{(23)}X_4 \ldots X_m]$ be the design matrix for model (3.3.2). Let $\underline{p}' = (\underline{p}_{12}, \underline{p}_{13}, \underline{p}_{23}, \underline{p}_4, \ldots, \underline{p}_m)$ and let $\underline{L}'\underline{p}$ be a main effect contrast for factor one [i.e. defined in terms of $\underline{p}_{12}$ and $\underline{p}_{13}$]. Then the s.s. due to this contrast is

$$(\underline{L}'\hat{\underline{p}})'(\underline{L}'(X'X)^{-}\underline{L})^{-1}(\underline{L}'\hat{\underline{p}})$$

where $\hat{\underline{p}}$ is a solution to the normal equations and $(X'X)^{-}$ is a g-inverse for $(X'X)$. Further, the variance of $\underline{L}'\hat{\underline{p}}$ is $\underline{L}'(X'X)^{-}\underline{L}\sigma^2$. A similar approach holds then for main effects of F_2 and F_3.

If the d.f. available for $ss_{t_{12}}$ is not $(n_1 - 1)(n_2 - 1)$ or the d.f. for factor one is not $(n_1 - 1)$, we still may be able to make some tests on factor one under the presence of first order interaction. For these tests we may compare factor level combinations as do Elston and Bush (1961). A similar approach would hold for F_2 and F_3.

This essentially covers then the analysis of the effects of F_1, F_2, and F_3 under the assumed presence of three first order (f.o.) interactions. For designs where more than three f.o. interactions are present, the analysis follows sections 3.1, or 3.2, or 3.3. For example, the analysis of the following model

$$E(\underline{y}) = X_{(12)}\underline{p}_{12} + X_{13}\underline{p}_{13} + X_{(23)}\underline{p}_{23} + X_{(14)}\underline{p}_{14} + X_5\underline{p}_5$$

would follow this section, with a few adjustments. Here the parameters $\underline{p}_{12}$, $\underline{p}_{13}$, $\underline{p}_{14}$ would now be defined

$$\underline{p}_{12} = \frac{1}{3}\underline{p}_1 + \underline{p}_2 + \underline{t}_{12}$$

$$\underline{p}_{13} = \frac{1}{3}\underline{p}_1 + \underline{p}_3 + \underline{t}_{13}$$

$$\underline{p}_{14} = \frac{1}{3}\underline{p}_1 + \underline{p}_4 + \underline{t}_{14}$$

Therefore, we shall terminate our approach to the analysis of m-factor models here.

The "Weakness" of Our Approach to the Analysis

A discussion of this approach seems relevant here. In finding the r.n.e. for one parameter, say $\underline{p}_{12}$, the intent is to simplify the solution to the full normal equations, $X'X\underline{p} = X'\underline{y}$, where in general, the design model is $E(\underline{y}) = X\underline{p}$, X is the design matrix, and vector $\underline{p}$ is the parameter of effects. However, for models where F_1 interacts with two or more factors, the s.s. due to this factor one under the assumed presence of the f.o. interactions cannot be found using C_{12}^{-} alone. Using our approach, the main effects of F_1 can be estimated, but again the variances of these effects cannot be found using C_{12}^{-} alone.

However, there are cases where there is little value in investigating such main effect estimates when two-factor interactions are assumed to exist. As mentioned earlier, these are cases where the analysis of the interactions are more important. Further, for those cases where main effects are to be investigated under the assumed presence of two-factor interactions, the special approach developed in this chapter still has some use.

First, it should be said that these problems of main effect analysis only arise for cases where one factor interacts with two or more factors.

Factors that interact only once cause no such problems. Next, it should be said that the special analysis mentioned in Chapter III provides not only a method of solving the r.n.e. but also provides those matrices necessary for calculating the variance-covariance matrix from the full normal equations, a g-inverse for (X'X), $(X'X)^{-}$.

For example, consider model (2) or (3.3.2). In order to find the variances of and s.s. due to contrasts in

$$f_{i..} = \bar{p}^{i.}_{12} + \bar{p}^{.k}_{13} - \bar{p}^{..}_{12} - \bar{p}^{..}_{13}$$

a g-inverse for X'X is required, where X'X is defined

$$X'X = \begin{bmatrix} H_{(12)(12)} & H_{(12)(13)} & \cdots & H_{(12)m} \\ H_{(13)(12)} & H_{(13)(13)} & \cdots & H_{(13)m} \\ \cdot & \cdot & \cdots & \cdot \\ H_{m(12)} & H_{m(13)} & & H_{mm} \end{bmatrix}$$

The matrices in X'X are defined: $H_{(12)(12)} = X'_{(12)}X_{(12)}$, $H_{(12)(13)} = X'_{(12)}X_{(13)}$, and $H_{(12)u}$, $= X'_{(12)}X_u$, $u = 4,\ldots,m$. Also, X'X is symmetric and all the other matrices have been defined in Chapter II and earlier sections of this chapter.

Now refer to Theorem 2.1. Let H be defined as in (3.3.3). Note that this is a submatrix of X'X. Using the notation of Theorem 2.1, let $X'X = B$, $H = B_{22}$, and let the g-inverse H^{-}, calculated for finding $C^{i.}_{12}$ in section 3.3, be $H^{-} = B^{-}_{22}$. Since H^{-} may be a non-singular matrix, the calculation of $M_{11} = (B_{11} - B_{12}B_{22}^{-}B_{21})^{-}$ may be simplified, i.e. $(M_{11})^{-}$ may be a non-singular matrix. Finally, all the submatrices of $X'X^{-} = B^{-}$ may be calculated, and thus a g-inverse for (X'X) may be found from the calculations necessary for the r.n.e. Therefore, if a g-inverse, $(X'X)^{-}$ is really necessary, as for example, in the analysis of certain main effects, this g-inverse may still be found from the analysis developed in this chapter.

Our analysis given in this chapter will provide some insight to the construction of designs to be discussed in Chapter V. The theory for such construction will be provided in Chapter IV.

CHAPTER IV

CONNECTING EXPERIMENTAL DESIGNS

Now that we have established an analysis of some types of incomplete multidimensional designs in Chapters II and III, let us proceed to develop techniques which will aid in the construction of the minimal designs mentioned in Chapter I.

In this chapter we shall establish a concept of connected designs when first order (two-factor) interaction is assumed. We shall use this concept to augment designs with the minimum number of factor level combinations (assemblies) necessary to do so.

The concept of a connected design is known and used most frequently in the analysis of incomplete block designs. A connected incomplete block design has treatments associated with blocks in such a manner that all treatment differences are estimable.

4.1 A Literature Review And Summary of Results

Anderson (1968) and Srivastava and Anderson (1970) defined the concept of connectedness in main effect MD's, where levels of one factor are connected in the presence of the other m-1 factors. Their results will be given in section 4.2. From these results we shall proceed to establish a similar concept for level combinations of two-factors, and thus define a connected design with assumed two factor interactions. Using this definition, we shall establish a procedure for finding assemblies to augment designs which are not completely connected. Further, we shall confine our interest to augmenting designs with only the minimum number of assemblies necessary.

Federer (1956, 1961) presents a procedure for adding new treatments (and thus assemblies) to a design in order to estimate the effects of the new treatments. Federer does use the minimum of necessary assemblies, and is concerned with connecting the levels of one-factor treatments. However, the treatments that he adds to a design are "new" treatments, that did not appear in the original design. We shall be concerned with only connecting factor levels that do appear in the original design.

4.2 Augmenting Designs With Assumed First Order Interactions

Here we shall present the basic definitions and theorems for connected designs. These theorems will be used in developing a procedure for augmenting designs.

Definition 4.1

A multidimensional design (MD), T, consists of h=h(T) assemblies (factor level combinations) and an associated linear model LM(T).

Let T be a MD with m factors $F_1, F_2, \ldots, F_m$, where F_i has n_i levels. Let LM(T) be of the form $E(\underset{\sim}{y}_T) = X_T \underset{\sim}{p}_T$ where y_T is the observation vector (hx1), X_T is the design matrix (hxM(T)), and $\underset{\sim}{p}_T$ is the parameter vector (M(T)x1) with $\underset{\sim}{p}_T' = (\underset{\sim}{p}_{i_1}', \underset{\sim}{p}_{i_2}', \ldots, \underset{\sim}{p}_{i_q}', \underset{\sim}{p}_{j_1 j_2}', \ldots, \underset{\sim}{p}_{j_r j_s}')$ and

$$\underset{\sim}{p}_a' = (p_a^1, p_a^2, \ldots, p_a^{n_a})$$

$$\underset{\sim}{p}_{bc}' = (p_{bc}^{11}, p_{bc}^{12}, \ldots, p_{bc}^{n_b n_c}) .$$

Here $\underset{\sim}{p}_a$ $(a = i_1, i_2, \ldots, i_q)$ refers to the factor F_a that does not interact with any other factor, and $\underline{p}_{bc}$ $((b,c) = (j_1, j_2),\ldots, (j_r j_s))$ refers to pairs of factors that interact with each other. Accordingly, let S_a denote the set of all levels of F_a, and let S_{bc} denote the set of pairs of level combinations of F_b and F_c. M(T) is the number of componets in $\underline{p}_T$. The intention here is to estimate with MD T, parameter contrasts of the form

$$p_a^{\ell_a} - p_a^{\ell'_a} \tag{4.2.1}$$

and

$$p_{bc}^{\ell_b \ell_c} + p_{bc}^{\ell'_b \ell'_c} - p_{bc}^{\ell'_b \ell_c} - p_{bc}^{\ell_b \ell'_c} \tag{4.2.2}$$

Definition 4.2

Contrasts of the form (4.2.1) are called main effect-or type I contrasts, contrasts of the form (4.2.2) are called interaction- or type II- contrasts.

In order to see whether a given MD T allows estimation of all type I- and type II-contrasts, or, if not, how one can generate a design T^* from T that has this property it is useful to introduce the concept of connectedness of a design. We shall simply extend the results of Srivastava and Anderson (1970).

Definition 4.3 (Srivastava and Anderson, (1970)):

(i) Two levels in S_i, say u_i and u'_i, are said to be connected w.r.t. all sets S_a and S_{bc}, if there exists a sequence of assemblies in T, say t_1, $t_2,\ldots,$ t_w, such that

$$\delta^{-1} E\left(\sum_{\rho=1}^{w} (-1)^{\rho} y_{t_\rho} \right) = p_i^{u_i} - p_i^{u_i'} ,$$

where $\delta(\neq 0)$ is a constant and y_{t_ρ} denotes the response y for assembly t_ρ.

(ii) If all levels of S_i are connected then T is said to be connected w.r.t. S_i (or, alternatively, w.r.t. F_i).

(iii) T is said to be type I-connected if it is connected w.r.t. all S_a.

(iv) The sequence of assemblies $(t_1,\ldots,t_w)$ is said to be a chain connecting levels u_i and u_i' of set S_i.

Definition 4.4

(i) Two pairs of level combinations in S_{jk} of the form (u_j,u_k), (u_j',u_k'), with either $u_j=u_j'$, $u_k\neq u_k'$ or $u_j\neq u_j'$, $u_k=u_k'$, are said to be connected w.r.t. all sets S_a and S_{bc}, if there exists a sequence of assemblies t_1', t_2',...,t_z' such that

$$\gamma^{-1} E\left(\sum_{\rho=1}^{z} (-1)^{\rho} y_{t_\rho'} \right) = p_{jk}^{u_j u_k} - p_{jk}^{u_j' u_k'}$$

where $\gamma(\neq 0)$ is a constant.

(ii) If all admissible quadruplets of S_{jk} as given in (i) are connected then T is said to be connected w.r.t. S_{jk} (or, alternatively, w.r.t. $F_j \times F_k$).

(iii) T is said to be type II-connected if it is connected w.r.t.

all S_{bc}.

(iv) The sequence $(t'_1, \ldots, t'_z)$ is said to be a chain connecting level combinations (u_j, u_k), (u'_j, u'_k), set S_{jk}. The above definitions are useful in checking the connectedness of designs.

Definition 4.5 T is said to be completely connected if it is type I- and type II-connected. If T is not completely connected it is said to be partially connected.

To check whether a given design T has the property of type I- and type II-connectedness, the following theorems can be used. Let λ_i^{u} (T) denote the frequency of level u_i (Bose and Srivastava, 1964a) and $\lambda_{jk}^{u_j u_k}$ (T) denote the frequency of level combination (u_j, u_k) in design T.

Theorem 4.1 (Srivastava and Anderson, 1970):

The design T is type I-connected if and only if the following conditions hold: For every set S_i let the sequence $(t_1, t_2, \ldots, t_{w_i})$ with w_i even be such that $T_{1i} = (t_2, t_4, \ldots, y_{w_i})$ and $T_{2i} = (t_1, t_3, \ldots, t_{w_i-1})$ with

(a) $\lambda_r^{u_r}(T_{1i}) = \lambda_r^{u_r}(T_{2i}) \qquad 1 \le u_r \le n_r, \quad r \ne i$

(b) $\lambda_i^{v_i}(T_{1i}) = \lambda_i^{v_i}(T_{2i}) \qquad v_i \ne u_i \text{ or } u'_i$

(c) $\lambda_i^{u_i}(T_{1i}) = \lambda_i^{u_i}(T_{2i}) + \delta$

$\lambda_i^{u'_i}(T_{1i}) = \lambda_i^{u'_i}(T_{2i}) - \delta$

(d) $\lambda^{u_x u_y}(T_{1i}) = \lambda^{u_x u_y}(T_{2i}) \qquad 1 \le u_x \le n_x, \quad 1 \le u_y \le n_y, \quad \text{all } S_{xy}$

where δ is a non-zero integer.

Theorem 4.1 is useful in constructing main effect designs. For designs with two-factor interaction, we have the following theorem.

Theorem 4.2

The design T is type II-connected if and only if the following conditions hold: For every set S_{jk} and every pair of level combinations $(u_j,u_k),(u_j',u_k') \in S_{jk}$ with either $u_j=u_j'$, $u_k\neq u_k'$ or $u_j\neq u_j'$, $u_k=u_k'$, let the sequence of assemblies $(t_1',t_2',\ldots,t_{z_{jk}}')$ with z_{jk} even be such that $T_{1jk} = (t_2', t_4',\ldots,t_{z_{jk}}')$ and $T_{2jk} = (t_1', t_3',\ldots,t_{z_{jk}-1}'$ with

$$\text{(a)}\quad \lambda_r^{u_r}(T_{1jk}) = \lambda_r^{u_r}(T_{2jk}) \qquad 1\leq u_r\leq n_r,\ 1\leq r\leq m$$

$$\text{(b)}\quad \lambda_{xy}^{u_xu_y}(T_{1jk}) = \lambda_{xy}^{u_xu_y}(T_{2jk}) \qquad 1\leq u_x\leq n_x,\ 1\leq u_y\leq n_y$$

$$x\neq y,\ (x,y)\neq(j,k)$$

$$\text{(c)}\quad \lambda_{jk}^{u_ju_k}(T_{1jk}) = \lambda_{jk}^{u_ju_k}(T_{2jk}) + \gamma$$

$$\lambda_{jk}^{u_j'u_k'}(T_{1jk}) = \lambda_{jk}^{u_j'u_k'}(T_{2jk}) + \gamma$$

where γ is a non-zero integer

$$\text{(d)}\quad \lambda_{jk}^{v_jv_k}(T_{1jk}) = \lambda_{jk}^{v_jv_k}(T_{2jk}) \qquad (v_j,v_k)\neq(u_j,u_k) \text{ and } (u_j',u_k')$$

Proof: Suppose the conditions (a), (b), (c) and (d) hold. Consider the expression

$$E\left(\sum_{\rho=1}^{z} (-1)^{\rho} y_{t_\rho} \right).$$

By condition (a) all main effects $p_r^{u_r}$, and by (b), interaction effects $p_{xy}^{u_x u_y}$ where $S_{xy} \neq S_{jk}$, will each appear in this expression the same number of times with a plus sign as with a minus sign. Now consider the interaction effects of level combinations of S_{jk}. All effects of level combinations of S_{jk}, other than the two mentioned, $(u_j u_k)$, $u_j' u_k')$, will then appear in the expression with a plus and a minus sign the same number of times as indicated by (d). Then by (c), it is seen that

$$E\left(\sum_{\rho=1}^{z} (-1)^{\rho} y_{t_\rho}\right) = \left(p_{12}^{u_j u_k} - p_{12}^{u_j' u_k'}\right)$$

Conversely, suppose that $t_1', t_2', \ldots, t_z'$ is a chain connecting level combinations (u_j, u_k) and (u_j', u_k'), of set S_{jk}. It then follows directly from Definition 4.4 that conditions (a), (b), (c), and (d) must hold.

The relationship between connectedness and estimability is given by the following theorems.

Theorem 4.3 (Srivastava and Anderson, 1970).

The type I-contrast $p_i^{u_i} - p_i^{u_i'}$ is estimable if and only if u_i and u_i' in S_i are connected w.r.t. all sets S_a and S_{bc}.

Theorem 4.4

The linear function $p_{jk}^{u_j u_k} - p_{jk}^{u_j' u_k'}$ is estimable if and only if (u_j, u_k) and (u_j', u_k') in S_{jk} are connected with respect to all sets S_a and S_{bc}.

Proof: If there is a chain connecting these two level combinations of S_{jk}, then by Definition 4.4, the contrast is estimable. Conversely, if the contrast is estimable then there exists some linear combination of $E(y_{t_i})$, such that

$$\sum_i a_i E(y_{t_i}) = p_{jk}^{u_j u_k} - p_{jk}^{u_j' u_k'}$$

where the coefficients a_i are rational numbers. To see this, we observe that since the contrast is estimable, we know that its best linear unbiased estimator is

$$\hat{p}_{jk}^{u_j u_k} - \hat{p}_{jk}^{u_j' u_k}$$

where the $\hat{p}_{jk}^{u_j u_k}$ are obtained as the solutions to the normal equations; viz. $\hat{p}_T = (X_T'X_T)^- X_T'y_T$ A matrix with integer elements, Γ, may be chosen such that a choice of $(X_T'X_T)^-$ could be $(X_T'X_T + \Gamma)^{-1}$, since $(X_T'X_T + \Gamma)$ is a non-singular matrix. (See Theorem 2.3).

Knowing then that $(X_T'X_T + \Gamma)$ and X_T have integer elements, we see that the elements of $(X_T'X_T + \Gamma)^{-1}$ are rational, and have the solutions to the normal equations as linear functions of the observations with rational coefficients. The proof is then straightforward.

Following the proof of Anderson (1968), we may let δ be an integer such that $a_i = c_i \delta^{-1}$, where c_i is an integer for all i. Then

$$\sum_i a_i (y_{t_i}) = \delta^{-1} \sum_i c_i E(y_{t_i})$$
$$= p_{jk}^{u_j u_k} - p_{jk}^{u'_j u'_k}$$

Then we may arrange the assemblies, with t_i each appearing c_i times, into a sequence $(t_1, \ldots, t_z) = T^*$, such that $(t_2, t_4, \ldots, t_z) = T_1^*$ is formed by all assemblies with a plus sign and $(t_1, t_3, \ldots, t_{z-1}) = T_2^*$ is formed by all assemblies with a minus sign. Since (a),(b),(c), and (d) of theorem 4.2 must hold for this sequence T^*, T^* is a chain connecting the level combinations (u_j, u_k), $(u'_j u'_k)$, of S_{jk}.

This completes the proof.

In order to show then that all type I and type II contrasts are estimable, we may use theorem 4.4 and its corollaries which follow directly.

Corollary 4.4.1 Type II-contrasts are estimable if either $p_{jk}^{u_j u_j} - p_{jk}^{u_j u'_k}$ and $p_{jk}^{u'_j u_k} - p_{jk}^{u'_j u'_k}$ or $p_{jk}^{u_j u_k} - p_{jk}^{u'_j u_k}$ and $p_{jk}^{u_j u'_k} - p_{jk}^{u'_j u'_k}$ are estimable.

Corollary 4.4.2 If the design T is type II-connected, then all contrasts of the form $\sum_{u_j, u_k} c_{jk}^{u_j u_k} p_{jk}^{u_j u_k}$ with $\sum_{u_j, u_k} c_{jk}^{u_j u_k} = 0$ are estimable.

Corollary 4.4.3 If the design T is type I-and type II-connected, i.e. completely connected, then all type I-and type II-contrasts are estimable.

To check whether a given MD, T, is completely connected we use the transitivity property of connectedness and the following definition and theorem.

Definition 4.6

Let T be a MD with $M^*(T)$ sets S_a, S_{bc}. Let T_ν be a MD design obtained from T by ignoring ν sets, and let S_α be one of the remaining sets, where S_α can refer to a set of type S_a or S_{bc}; such as S_α is said to belong to T_ν. Then S_α is said to be connected w.r.t. the remaining $M^*(T) - \nu - 1$ sets in the original T if T_ν is connected w.r.t. S_α.

Theorem 4.5

An MD, T, is completely connected if and only if there exists a sequence of designs T_ν, $\nu = 0,1,\ldots, M^*(T) - 2$ such that T_ν is connected w.r.t. to at least one S_α, where S_α belongs to T_ν.

Proof: We shall verify the first steps in this sequential procedure. Without loss of generality suppose that S_1 is connected with respect to every other set S_a and S_{bc} in the design $T_o = T$. Then for any contrast

$$\sum_{u_1} \phi_1^{u_1} p_1^{u_1}$$

where the $\phi_1^{u_1}$ are constants with $\sum_{u_1} \phi_1^{u_1} = 0$, there exists a linear function of the observations, $\underline{\psi}'\underline{y}_T$ say, such that

$$(4.2.3) \qquad E\,(\underline{\psi}'\underline{y}_T) = \sum_{u_i} \phi_i^{u_i}\, p_i^{u_i}.$$

Next suppose that S_{jk} is connected with respect to T, which is formed by ignoring S_i (or F_i) in T. Then according to Definition 4.2 there exists a sequence of assemblies $\sum_{\rho=1}^{z} (-1)^{\rho} t_{\rho}$ say, with z even such that

$$(4.2.4) \qquad E\,(\delta^{-1} \sum_{\rho=1}^{z} (-1)^{\rho} t_{\rho}) = \sum_{u_i} \phi_i^{u_i} p_i^{u_i} + p_{jk}^{u_j u_k} - p_{jk}^{u_j' u_k'} \quad \text{for all}$$

pairs $(u_j u_k)$ and $(u_j' u_k')$ in S_{jk} as specified earlier. Obviously $\sum_{u_i} \phi_i^{u_i} p_i^{u_i}$ in (4.2.4) is a contrast in the $p_i^{u_i}$ and hence estimable by virtue of (4.2.3). Hence S_{jk} is connected in T and hence by Corollary 4.4.1 all type II-contrasts in S_{jk} are estimable.

Next suppose without loss of generality that S_ℓ is connected in T_2 which is obtained from T by omitting S_i and S_{jk}. Again, there exists a sequence of assemblies, $\sum_{\rho=1}^{w} (-1)^{\rho} t_{\rho}$ say, with w even such that

$$(4.2.5) \qquad E(\delta^{-1} \sum_{\rho=\phi}^{w} (-1)^{\rho} t_{\rho}) = \sum_{u_i} \phi_i^{u_i} p_i^{u_i} + \sum_{u_j, u_k} \xi_{jk}^{u_j u_k}\, p_{jk}^{u_j u_k} + p_\ell^{u_\ell} - p_\ell^{u_\ell'}$$

for all pairs u_ℓ, u_ℓ' in S_ℓ. Obviously, since w is even, $\sum \phi_i^{u_i}\, p_i^{u_i}$ is estimable because of (4.3.3) and $\sum \xi_{jk}^{u_j u_k}\, p_{jk}^{u_j u_k}$ is estimable because of (4.3.4) and Corollary (4.4.2). Hence, S_ℓ is connected in T. We proceed in this way until we reach $T_{M*(T)-1}$.

The converse is obvious. If T is completely connected, then each T_i is connected. This completes the proof.

Definition 4.7

Two MD's T and T' are said to be model equivalent, denoted by $LM(T) = LM(T')$ if $p_T = p_{T'}$.

Theorem 4.6

If T is a partially connected MD, then there exists a sequence of assemblies $D^* = (d_1^*,\ldots,d_n^*)$ such that $T^* = (T+D^*)$ is a completely connected design with T and T^* being model equivalent.

Proof (Augmentation Procedure): The following proof is constructive and hence provides an augmentation procedure for augmenting partially connected designs.

Let T be a partially connected MD and suppose that T is not connected with respect to S_{α_1}, $S_{\alpha_{2_1}}$, ..., S_{α_ℓ}, where S_{α_i} is either of the form S_a or S_{bc}. Let $\omega(S)$ denote the number of different factors in S; e.g., $\omega(S_a) = 1$, $\omega(S_{bc}) = 2$, $\omega(S_a U S_{bc}) = 3$, $\omega(S_{ab} U S_{bc}) = 3$, etc. Then choose a set of assemblies not in T, say D_1, such that $T_{(1)} = T+D_1$ is connected with regard to S_{α_1}. Now ignore all elements of S_{α_1} in $T_{(1)}$ to obtain the $(m-\omega(S_{\alpha_1}))$- dimensional design $T_{(1)1} = (T+D_1)_1$. Next choose a set of assemblies not in $T_{(1)1}$, say D_2, such that $T_{(1,2)} = T_{(1)1}+D_2$ is connected with regard to S_{α_2}. Again, ignore all elements of S_{α_2} in $T_{(1,2)}$ and denote the resulting $(m-\omega(S_{\alpha_1} U S_{\alpha_2}))$-dimensional design by $T_{(1,2)1}$. Continue this process with $S_{\alpha_3},\ldots,S_{\alpha_\ell}$. Notice that the sets of assemblies $D_2,\ldots,D_\ell$ contain only $m_2,\ldots,m_\ell$ factors with $m > m_2 > \ldots > m_\ell$. In order to make them compatible with T, add to D_i $(i=2,\ldots,\ell)$ arbitrary level combinations of the remaining $m - m_i$ factors not in D_i. Call this set of assemblies then D_i^* and let $D^* = \sum_{i=1} D_i^*$ with $D_1^* = D_1$. Using Theorem 4.5, this completes the proof.

The augmentation procedure is, of course, not unique; different D^* may differ in their numbers of assemblies $h(D^*)$ as well as in the actual assemblies. One subclass of augmented designs that is of interest is the class of designs T^* that are obtained from T with the minimal number of additional assemblies.

Definition 4.8

The Md, T^*, is said to be a minimal augmented multidimensional design (MAMD) with regard to the MD, T, if D^* has the minimum number of assemblies such that $T^* = T+D^*$ is completely connected.

Theorem 4.7

For a given MD, T, there exists a MAMD, $T^* = T + D^*$ with the number of different assemblies in T^*, equal to $h(T^*)$, where $h(T^*) = h(T) + h(D^*)$, and $h(D^*) = \nu(p_T) - r(X_T)$, $\nu(p_T)$ is the number of d.f. for p_T, and $r(X_T)$ is the rank of $\mathbf{X}_T$.

Proof: The fact that a MAMD, T^*, does exist follows from Theorem 4.6. Let the d.f. due to S_α (independent of the design) be $\nu_\alpha(p_T)$ and the d.f. available for S_α in T be $\nu_{\alpha_T}(p_T)$. Then $\nu_\alpha(p_T) - \nu_{\alpha_T}(p_T)$ represents the additional d.f. required to connect S_α. Since S_α is not connected in T, not all assemblies required to connect S_α are present in T. Therefore, $\nu_\alpha(\underline{p}_T) - \nu_{\alpha_T}(\underline{p}_T)$ simply represents the number of new assemblies needed to connect S_α. If we proceed in like manner with k such sets then $\nu(\underline{p}_T) = \sum_{i=1}^{k} \nu_i(\underline{p}_T)$ and $r(X_T) = \sum_{i=1}^{k} \nu_{i_T}(\underline{p}_T)$. It follows that $h(D^*) = \nu(\underline{p}_T) - r(X_T)$.

An illustration of theorem 4.7 is found in Example 4.1.

Example 4.1

$M(T) = 3$, $m = 4$, $n_1 = n_2 = n_3 = 3$. The factors F_i, $i=1,\ldots 4$, are identified by F_1 (1,2,3), F_2 (I,II,III), F_3 (A,B,C), F_4 $(\alpha,\beta,\gamma,)$

The Design T

A 3^{-2} Replicate of 3^4 Factorial
h(T) = 9 Assemblies

$(1,I,A,\alpha)$	(1)	$(2,I,B,\beta)$	(4)	$(3,I,C,\gamma)$	(7)
$(1,II,C,\beta)$	(2)	$(2,II,A,\gamma)$	(5)	$(3,II,B,\alpha)$	(8)
$(1,III,B,\gamma)$	(3)	$(2,III,C,\alpha)$	(6)	$(3,III,A,\beta)$	(9)

The above assemblies may be represented by t_i, $i=1,\ldots,9$. For the assumed model

$$E(\underline{y}) = X_{12}\underline{p}_{12} + X_{13}\underline{p}_{13} + X_4\underline{p}_4$$

the design T is not connected in sets S_{12}, S_{13}, or S_4.

A sequence to connect T, D^*, may be expressed in terms of the D_j, $j=1,2,3$, that follow.

Sequence	Assemblies	Set Connected
D_j		
D_1:	$(1,I,A,\beta)$, $(1,I,A,\gamma)$	S_4
D_2:	$(2,I,A)$ $(1,II,A)$ $(1,III,A)$ $(3,I,A)$ $(2,III,A)$ $(3,II,A)$	S_{12}
D_3:	Empty	S_{13}

Therefore $h(D^*) = 8$ and $h(T^*) = 17$.

In order to see that $T^* = T + D^*$ is connected in sets S_{12}, S_{13}, and S_4, observe that if for the purpose of illustration one observation is taken per assembly and if $(1,I,A,\beta) = y_{1,I,A,\beta}$, then for set S_4

$$(1,I,A,\alpha) - (1,I,A,\beta) = \hat{p}_4^{\alpha} - \hat{p}_4^{\beta}$$

$$(1,I,A,\alpha) - (1,I,A,\gamma) = \hat{p}_4^{\alpha} - \hat{p}_4^{\gamma}$$

Now that S_4 is connected we may ignore it for the purpose of finding assemblies to connect S_{12}. We need only to find assemblies in three factors. (Any levels of S_4 may be appended to these assemblies). Then, for set S_{12}

$$(2,I,A) + (1,II,A) - (2,II,A) - (1,I,A) =$$

$$\hat{p}_{12}^{2I} + \hat{p}_{12}^{1II} - \hat{p}_{12}^{2II} - \hat{p}_{12}^{1I}$$

$$(3,I,A) + (1,II,A) - (3,II,A) - (1,I,A) =$$

$$\hat{p}_{12}^{3I} + \hat{p}_{12}^{1II} - \hat{p}_{12}^{3II} - \hat{p}_{12}^{1I}$$

$$(1, III, A) + (2, II, A) - (2, III, A) - (1, II, A) =$$

$$\hat{p}_{12}^{1III} + \hat{p}_{12}^{2II} - \hat{p}_{12}^{2III} - \hat{p}_{12}^{1II}$$

By theorem 4.5, S_{13} must be connected.

The augmented design of Example 4.1 would have the minimum number of assemblies needed to estimate contrasts of the sets S_{12}, S_{13}, and S_4, since $h(T^*) = 17$. This need not always be the case. We may add the minimum number of assemblies to augment a design T and still have $h(T^*) > \nu(\underline{p}_T) + 1$, as indicated by Example 4.2 which follows.

In Example 4.2, we shall use standard notation to represent the factor level combinations, as the notation given in Example 4.1 was given for the purpose of illustration.

Example 4.2

The Design T

(Addelman and Kempthorne, 1961)
An Orthogonal Main Effect Plan For A
$4 \times 3^4 \times 2$ Factorial
$h(T) = 25$ Assemblies

000000 (1)	101111 (6)	202221 (11)	302221 (16)	000000 (21)
011220 (2)	112200 (7)	212001 (12)	310011 (17)	010121 (22)
022011 (3)	122020 (8)	220120 (13)	320201 (18)	021201 (23)
022101 (4)	120201 (9)	220210 (14)	321020 (19)	022021 (24)
000221 (5)	100021 (10)	201001 (15)	302100 (20)	002210 (25)

The model: $E(\underset{-}{y}) = X_1\underset{-}{p}_1 + X_{23}\underset{-}{p}_{23} + X_{24}\underset{-}{p}_{24} + X_5\underset{-}{p}_5 + X_6\underset{-}{p}_6$

The old assemblies are t_i, $i = 1,\ldots,25$.

Sequence D_i	New Assemblies	Assemblies Used To connect S_α	Set Connected
D_1:	111220 (26) 211220 (27)	$(t_2, t_{26}, t_{27}, t_9, t_{18})$	S_1
D_2:	Empty	(t_8, t_{24})	S_6
		(t_1, t_{10}, t_3, t_8)	S_5
D_4:	110211 (28)	$(t_{26}, t_4, t_{22}, t_{28}, t_1,$ $t_{23}, t_9, t_{15})$	S_{23}
D_5:	Empty	Empty	S_{24}

$h(D^*) = 3$, $h(T^*) = 28$, $\nu(\underset{-}{p}_T) = 20$.

Note here that using assemblies t_2 and t_{26} we may obtain an estimate of $p_1^0 - p_1^1$. In like manner assemblies t_{27}, t_9, and t_{18} may be used to demonstrate that all type I-contrasts of S_1 are estimable, and hence S_1 is conencted. "Assemblies Used To Connect S_α" then connect S_6, S_5, S_{23}, and S_{24} in a similar manner.

Corollary 4.7.1

For a MAMD, T^*, the number of assemblies in D^*, $h(D^*)$, is independent of the ordering of the sets $S_{\alpha_1}, S_{\alpha_2},\ldots,S_{\alpha_\ell}$ in the augmentation procedure.

Proof: Obvious from the augmentation procedure.

CHAPTER V

OPTIMALITY OF MINIMAL AND MINIMAL AUGMENTED DESIGNS

5.1 Introduction

We established in Chapter IV a procedure for constructing MAMD's. For a given design T, a sequence D^* may be chosen such that $h(D^*)$ is a minimum number. In this chapter we shall present a procedure for determining which D^* is best or "optimal" for augmenting T. This procedure also gives rise to a method of constructing a design, T^*, such that $h(T^*)$ is a minimum number; i.e. $h(T^*) = \nu(\underline{p}_{T^*}) + 1$.

Note that if T^* is a MAMD, $h(T^*)$ is not necessarily a minimum number. As an illustration, in Example 4.2, $h(T^*) = 28$, and

$$\nu(\underline{p}_{T^*}) = 3 + 4(2) + 1 + 2\cdot 2 + 2\cdot 2 = 20.$$

Therefore another class of designs whose total number of assemblies are such that $h(T^*) = \nu(\underline{p}_{T^*}) + 1$ shall be investigated in this chapter.

Due to the difficulty that arises in constructing MD's subject to having $h(T^*)$ or $h(D^*)$ a minimum number, the usual optimal properties such as balance, symmetry or orthogonality may not apply. For this reason we shall be led to an approach to these designs which is limited in scope. However, this approach shall be readily applicable.

There is no previous work on MAMD's or designs which require $h(T^*)$ to be a minimum number. Addelman and Kempthorne (1961) present orthogonal main effect plans for k^n factorials, some of which require the minimum number of assemblies, $k(n-1) + 1$, but these designs are very few, since their goal was the construction of orthogonal designs, not minimal designs. Raktoe and Federer (1968) give an approach to non-orthogonal main effect plans in k^n factorials. This method requires k n assemblies for k even and $k(n-1) + 2$ assemblies for k odd. We shall present designs for k even or odd that require $k(n-1) + 1$ assemblies. Also, our approach will not be limited to k^n factorials.

5.2 Summary of Chapter V

Section 5.3 covers the relationship between MAMD's and designs which require a minimum number of assemblies. Section 5.4 discusses optimality criteria that are used on main effect designs and introduces the A-optimality criterion that we will use for our minimal designs. Section 5.5. outlines a procedure for determining a set of "optimum" designs, Ψ, from which one finds the A-optimal design with respect to Ψ. This procedure is discussed further in section 5.6 and methods for determining a set Ψ are given in sections 5.7 and 5.8. Examples will be given in 5.7 and 5.8 in order to help the reader understand these methods. Since 5.7. and 5.8 deal specifically with main effect designs, particular examples of first order interaction designs developed from these main effect designs are given in section 5.9.

5.3 MMD's and MAMD's

Before considering optimality criteria, let us clearly understand the difference between the MAMD's defined in Chapter IV and MD's which require a minimum number of assemblies.

Definition 5.1

If T is a completely connected MD with d.f. $\nu(p_T)$ for p_T in h=h(T) assemblies, then T is said to be a minimal multidimensional design (MMD) if $h = \nu(p_T) + 1$.

Example 5.1

Let T be a completely connected main effect MMD of m factors. Then $\nu(p_T) = \sum_{i=1}^{m} (n_i - 1) = \sum_{i=1}^{m} n_i - m$, and $h = \sum_{i=1}^{m} n_i - m + 1$.

Some MAMD's are MMD's. This is illustrated by the design $T^* = T+D^*$ in Example 4.1, where $h(T^*) = 17 = 16 + 1$ and $\nu(p_{T^*}) = 4(2) + 4 + 4 = 16$. As we will see later, optimal MMD's may be found by using a method similar to the method used to find optimal MAMD's. Therefore the construction of optimal MAMD's shall be considered first (section 5.5). Before doing this we need to discuss some optimality criteria to be used on MAMD's and MMD!s.

The Use Of Main Effect Criteria To Determine Optimal MAMD's and MMD's.

The criteria we shall discuss in section 5.4 are used most frequently for main effect designs and are concerned with minimizing the variance or average variance of $\hat{p}_a^{u_a} - \hat{p}_a^{u_a'}$. One of these criteria shall thus serve as a guide to some MD's where two factor interactions are assumed, since the parameter $p_{bc}^{u_b u_c}$ can be handled in a manner similar to the parameter $p_a^{u_a}$. For example, in order to show that the set of factor levels, S_{bc} is connected in T, simply demonstrate that all differences, $p_{bc}^{u_b u_c} - p_{bc}^{v_b v_c}$, $(u_b u_c) \neq (v_b v_c)$, are estimable, since all interaction contrasts of the form $p_{bc}^{u_b u_c}$ +

$p_{bc}^{u'_b u'_c} - p_{bc}^{u'_b u_c} - {}_{bc}^{u_b u'_c}$ can be expressed in terms of the "two-way" differences,

$$p_{bc}^{u_b u_c} + p_{bc}^{u'_b u'_c} - p_{bc}^{u'_b u_c} - p_{bc}^{u_b u'_c} = (p_{bc}^{u_b u_c} - p_{bc}^{u'_b u_c}) + (p_{bc}^{u'_b u'_c} - p_{bc}^{u_b u'_c}) .$$

Therefore, it is possible to use main effect optimality criteria on the paramerters $p_{bc}^{u_b u_c}$ in order to compare designs where a two-factor interaction is assumed.

However, in future discussion we shall develop a general construction of designs in $M^*(T)$ sets, S_α, $\alpha = 1,\ldots,M^*(T)$. These sets are to be understood sets of factor levels, S_a or sets of level combinations, S_{bc}. Generally we speak of sets S_α. However, we shall not consider sets of the form S_{bc}, S_{bd}, say; i.e., we shall limit our general developments to designs where if first order interaction is assumed, no factor shall interact with two or more factors. Although we give a general approach to these latter designs (see Case 3 and 4 of section 5.9), the particular procedures to be mentioned in section 5.6, 5.7, and 5.8 will not apply in general. More shall be said on these types of designs in section 5.9.

5.4 A Review of Some Main Effect Optimality Criteria

Previous work in optimal designs by Plackett and Burman (1946) Chernoff (1953), Ehrenfeld (1955), Kempthorne (1956), and Kiefer (1959a, 1959b, 1961) will be summarized here.

A design may be A-, E-, or D-optimal. An A-optimal design is one where the average variance of differences $\hat{p}_a^{u_a} - \hat{p}_a^{u'_a}$ is least. An E-optimal design has the variances of thse differences bounded by the least upper bound. A D-optimal design estimates these differences with a minimum generalized variance. Quite often A-optimal designs are both E and D-optimal. (Kiefer, (1959b). Therefore we shall concern ourselves with A-optimality.

The following theorem gives a relationship between the latent roots of the c.m. for $\underline{p}_\alpha$ and A-optimality. This theorem will apply to all sets S_α, for parameter $\underline{p}_\alpha$, $\alpha = 1,\ldots,M(T^*)$. In the following theorem if $p_\alpha = p_{bc}$, then $n_\alpha = n_b n_c$.

Theorem 5.1 (Kempthorne, 1956):

Let C_α be the $n_\alpha \times n_\alpha$ coefficient matrix of the reduced normal equations, $C_\alpha \hat{\underline{p}}_\alpha = \underline{Q}_\alpha$. Further, let the rank of C_α be $n_\alpha - 1$, and denote by λ_i, $i = 1,2,\ldots,n_\alpha - 1$, the $n_\alpha - 1$ positive latent roots of C_α. Then

$$\frac{1}{n_\alpha(n_\alpha - 1)} \sum_{\substack{u_\alpha, u'_\alpha \\ u_\alpha \neq u'_\alpha}} V(\hat{p}_\alpha^{u_\alpha} - \hat{p}_\alpha^{u'_\alpha}) = \frac{2}{n_\alpha - 1} \sum_{i=1}^{n_\alpha - 1} \frac{1}{\lambda_i} \sigma^2$$

where $V(\hat{p}_\alpha^{u_\alpha} - \hat{p}_\alpha^{u'_\alpha})$ is the variance of $\hat{p}_\alpha^{u_\alpha} - \hat{p}_\alpha^{u'_\alpha}$.

Thus, a design T^* will be A-optimal for set S_α if the c.m. for $\hat{\underline{p}}_\alpha$ has roots λ_i such that $\Sigma \frac{1}{\lambda_i}$ is a minimum.

Ordinarily, for experimental designs, one approaches the concept of optimality in terms of the variance-covariance matrix involved. Since C_α is a singular we have no variance-covariance matrix in the strict sense. However, we shall relate A-optimality to the g-inverse for C_α, C_α^-. If C_α is an $n_\alpha \times n_\alpha$ c.m. for $\hat{\underline{p}}_\alpha$ of rank $n_\alpha - 1$ and $(C_\alpha + J)^{-1}$ (see theorem 2.2) is used as a g-inverse for C_α, then

$$tr(C_\alpha^-) = \sum_{i=1}^{n_\alpha - 1} \frac{1}{\lambda_i} + \frac{1}{n_\alpha}$$ where n_α is known and independent of the design (see Hinkelmann, 1968a). Therefore, a design that is A-optimal for S_α, has a c.m. for $\hat{\underline{p}}_\alpha$, C_α, such that there exists a g-inverse, C_α^-, where $tr(C_\alpha^-)$ is a minimum.

5.5 Finding Optimal MAMD's

The Method

We will indicate here how A-optimality can be applied to the problem of augmenting a partially connected design T for one set S_α. (The problem of augmenting T for more than one set shall be discussed later). As indicated in Chapter IV there may be many sequences, D^* that will connect S_α with the minimum number of assemblies. In order to find an optimal D^*, and thus an optimal $T^* = T + D^*$, our method is as follows: 1) consider a set of possible sequences that will connect S_α in T^*, i.e. determine an optimum set of designs, Ψ (defined by a property to be mentioned in this section and constructed by methods of section 5.7 and 5.8). This optimum set is to be generated either from the given design to be augmented, T_{Ψ_a} (which we will use for MAMD's) or a constructed design, T_{Ψ_b} (which we will use for MMD's). Next,

(2) in the set Ψ, find the A-optimal design for set S_α.

Before we apply this method let us first make a change in notation. For a set Ψ of optimum designs, an optimal MAMD will now be designated by T^* and an optimal sequence of assemblies by D^*. The members of Ψ will be MAMD's, and will be designated by T_i with corresponding sequence D_i, $i=1,\ldots,z$, say.

Definition 5.2

Let $\Psi = (T_1, T_2, \ldots, T_z)$ be a set of MD's. Let T^* be a member of this set and C_α^* the coefficient matrix for $\underline{p}_\alpha$ in T^*. T^* is said to be A-optimal for $\underline{p}_\alpha$ with respect to Ψ if for any other design T_i in Ψ whose corresponding c.m. has roots λ_j,

$$\sum_{j=1}^{n_\alpha - 1} \frac{1}{\lambda_j^*} \leq \sum_{j=1}^{n_\alpha - 1} \frac{1}{\lambda_j}$$

where λ_j^* are the non-zero roots of the c.m. for $\hat{\underline{p}}_\alpha$ in T^*. Since the only optimality criterion we are considering is A-optimality, we shall say in what follows that T^* is Ψ-optimal for set S_α.

The Set Ψ

We next consider the problem of constructing the set Ψ. The members of Ψ are already partially defined from the design to be augmented, T_{Ψ_a}. Each member of Ψ starts with the design to be augmented, T_{Ψ_a}, and differs only in the assemblies which are added to connect the set S_α. The members of Ψ are then $T_i = T_{\Psi_a} + D_i$, where D_i is constructed according to Theorem 4.5 and so that S_α is connected in T_i.

There are three ways the D_i may be determined. There may be an orthogonal or balanced MAMD, T_i^*, that results from T_{Ψ_a} by adding a certain set of assemblies, D_i^*. Certainly we should not ignore such a possibility or such a set D_i.

Next, due to the experimental conditions, the factor level combinations and thus the sequences, D_i may be limited to a very few. However, these two latter cases are rare and in this section we shall concern ourselves with a general approach, the third possibility, to the construction of D_i and a particular set Ψ.

<u>A General Approach To Finding MAMD's For Ψ</u>

Consider the c.m. in the r.n.e. for $\hat{\underline{p}}_\alpha$,

(5.5.1) $$C_\alpha = H_{\alpha\alpha} - (H_{\alpha 1} \ldots H_{\alpha m})H^- (H_{\alpha 1}, \ldots, H_{\alpha m})'.$$

An investigation of the composition of C_α may provide some information on Ψ-optimal MAMD's. The construction of $T_i = T_{\Psi_a} + D_i$ and thus the calculation of C_α for each T_i, is limited by two things over which we have no choice. These are (1) the actual construction of T_{Ψ_a} and (2) the number of assemblies in $D_i, h(D_i)$. We shall acknowledge these two constraints when constructing Ψ using a general approach.

Since T_{ψ_a} is already fixed, many elements of the matrices used to calculate C_α of the T_i in (5.5.1) are already fixed. Without loss of generality, now let $\alpha = 1$. Let us concern ourselves with the matrix $H = (H_{uv})$, $u,v = 2,\ldots,M^*(T)$, whose g-inverse shall be used in (5.5.1); i.e., we are ignoring how the levels of the set S_1, the set to be connected in T_i, appear in T_i. Trying to describe any given T_{ψ_a} in general terms is almost impossible, and thus trying to describe the matrix $H = (H_{uv})$, $u,v = 2,\ldots, M^*(T)$, in general terms may be impossible. Therefore, let us direct our attention to the diagonal elements of H (the level frequencies of each factor set S_β, $\beta \neq 1$).

There are cases where, if for each factor all levels of the same factor appear with equal frequency, a balanced, or symmetrical design, or one with like properties results. For example, Bose and Srivastava (1964b) indicate that this property is one requirement of an MPBD. This property, that all levels of the same set, S_β, $\beta = 2,\ldots,M^*(T)$, appear with equal frequency then may be viewed as a desirable characteristic. (However, this is not a rule, Addelman and Kempthorne(1961) give some examples of orthogonal designs where factor levels appear with proportional frequencies.) Further, in accordance with the definition of a good fraction given in Chapter I, this attempt at achieving equal frequencies may establish some uniformity of variance on effects involving sets S_β, $\beta = 2,\ldots,M^*(T)$.

Therefore, from the calculation of C_α, we have some idea of what might be the preferred sequences, D_i. If possible, the level combinations of all sets S_β, $\beta \neq 1$, in $T_i = T_{\Psi_a} + D_i$, should be such that the level (or level combinations) frequencies of each set S_β be equal (to k_β for example).

In constructing the sequences D_i, the next difficulty we face is that $h(D_i)$ is to be a minimum. Thus is may not be possible to have the level combinations of S_β appear with equal frequency in $T_i = T_{\Psi_a} + D_i$. Let $h(T_i) = a_\beta (\text{mod } n_\beta)$. Then it is possible to have a_β levels of S_β occur one more time than $n_\beta - a_\beta$ levels of S_β. If $h(T_i) = k_\beta n_\beta + a_\beta$, then we have

$$
\begin{aligned}
(5.5.2) \qquad h_{\cdots \ell_\beta \cdots} &= k_\beta, && \text{for } n_\beta - a_\beta \text{ levels of } S_\beta \\
&= k_\beta + 1, && \text{for } a_\beta \text{ levels of } S_\beta
\end{aligned}
$$

Although it will not always be possible, we recommend that the experimenter try to establish (5.5.2) for sets $S_\beta, \beta \neq 1$ in constructing the designs T_i of Ψ.

Federer (1961) used this type of an approach in adding new treatments to block designs. He required r treatments to appear ν_r times and ℓ treatments to appear ν_ℓ times, where $r+\ell$ is the total number of treatments. As mentioned, due to a possibly unusual nature of T_{Ψ_a} (5.5.2) may not be obtainable. Then it may be possible to have for set S_β, r_j levels occurring ν_{r_j} times, $j=1,2,\ldots,$ with $\Sigma\, r_j = n_\beta$.

However, those cases where the frequencies of each level in S_β are not equal in T_{Ψ_a} are unusual. The MD's most frequently used are those with balance, partial balance, or orthogonality that require the level frequencies of S_β to be the same. (The exception is again given in some cases of Addelman and Kempthorne, 1961). Therefore, most T_{Ψ_a}'s to be augmented will have equal level frequencies in each S_β.

Now all members of our set Ψ, $T_i = T_{\Psi_a} + D_i$, are to be the same w.r.t. level combinations of all sets S_β, $\beta \neq 1$. This means that ignoring S_1, all D_i are to be the same. Further in determining the D_i such that (5.5.2) holds for all T_i, we also determine the off diagonal matrices of H, H_{uv}, $u \neq v$, $u,v=2,\ldots,M^*(T)$. In summary then, all designs in Ψ will use the same H matrix for calculating the c.m. C_1.

To achieve different designs T_i, in Ψ, we vary the levels of S_1 in combination with levels of all other sets S_β. Exactly how this can be done and still achieve designs T_i that are connected in S_1 shall be discussed in sections 5.7 and 5.8. Let us for the moment assume that this can be done. Then all the members of Ψ will differ only in the occurrence of levels of S_1. One of these T_i will then be Ψ-optimal for set S_1.

A Summary of Our General Approach To Constructing Ψ

Let us review our procedure that we have just discussed for finding a Ψ-optimal MAMD. One begins with a design T_{Ψ_a}, not connected in set S_α. A set Ψ of possible MAMD's is considered and an Ψ-optimal design for S_α is found.

A general approach to construcing members of Ψ is follows: (1) The sequences D_i are to be constructed using levels of sets S_β, $\beta \neq \alpha$, by the augmentation procedure of Theorem 4.6 and in light of (5.5.2) (Procedure I, section 5.7 will demonstrate exactly how this can be done); (2) next, the sequences, D_i, are to have levels of S_α occurring with different frequencies (Procedure II, section 5.8 will demonstrate exactly how this can be done).

An example of this method is seen in Appendix II, where a 1/3 replicate of a 3^3 factorial is augmented for the presence of a two-factor interaction. There, $S_\alpha = S_{12}$ and $S_\beta = S_3$. For S_3 in T^*

(5.5.3) $$h_{..\ell} = 3 \quad \text{for } \ell = \text{level III}$$
$$= 4 \quad \text{for } \ell = \text{levels I, II} .$$

The set Ψ then consists of all possible level combinations of S_{12} with S_3, such that (5.5.3) holds for S_3.

For More Than One Set S_α

Now let us consider the problem of augmenting design T_{Ψ_a} for more than one set S_α, say $S_{\alpha_1}, S_{\alpha_2}, \ldots, S_{\alpha_k}$. The procedure mentioned in this section still applies, where from Ψ we determine a design T^* optimal for one set, say S_{α_1}. For the design T_{Ψ_a}, sequences D_i can be constructed such that one attempts to maintain (5.5.2) for all sets S_β, $\beta \neq \alpha_1$. (Note, $S_{\alpha_2},\ldots,S_{\alpha_k}$ are included here). Then the members of Ψ, T_i, will differ only in the manner that levels of S_{α_1} occur with the factor level combinations from the other sets S_β, $\beta \neq \alpha_1$.

Some reasons why we attempt to find a MAMD that is optimal for one set S_{α_1} shall follow. A common case of a MAMD may be a design T_{Ψ_a} that is augmented for only one set S_α. For example, the case where a main effect design is augmented for the presence of one two-factor interaction. Common main effect plans will not have many suspected interactions, since if many are suspected, a first-order interaction design perhaps would have initially been suggested. More reasons why we give a method of find Ψ-optimal MAMD's for one set are given in the next section.

5.6 Reasons For Finding MMD's and MAMD's That Are Ψ-optimal For Only One Set, S_α.

The method of finding Ψ-optimal MAMD's outlined in section 5.5 shall also apply to MMD's. More will be said on this method in the following sections, but first let us present some reasons for finding a design that is Ψ-optimal for only one set, S_α.

Given a fixed number of assemblies, it may be preferred to seek a design which is "best" for one factor (or pair of factors) rather than seek a design that is "semi-best" for two or more factors. For example, it may be preferred to choose a design T^* in ψ that has minimum average variance for estimate of $p_a^{u_a} - p_a^{u_a'}$, rather than a design T^{**} that has the minimum variance on all such sets $p_\alpha^{u_\alpha} - p_\alpha^{u_\alpha'}$, $\alpha=1,\ldots,M^*(T)$, or on a set of estimable function $b_1, b_2,\ldots,b_k$. However, T^* should have some degree of uniform variance on effects not involving S_a.

We shall take this view toward finding optimal MMD's (and MAMD's) for two reasons. The first is that with the tools at hand (previous developments), a procedure for finding an A-optimal MMD (or MAMD) is made easier. A unified approach to finding MAMD's (and as it will be seen for MMD's) is made possible through the use of the general form of the matrix C_α. The relationship of C_α and a g-inverse C_α^- was seen in section 5.4. Using the latent roots of C_α, Ψ-optimality was defined. In constructing Ψ using the form of C_α as a guide, we attempted some "protection" for the sets S_β, $\beta = 1,\ldots,M^*(T)$, $\beta \neq \alpha$, by maintaining (5.5.2). Thus using previous developments, we can establish one approach to finding these designs, and this fact is important, since MMD's and MAMD's may be difficult to construct.

The second reason that we take such a view is that in order to use a criterion as A-optimality on all estimates $\hat{b}_1, \hat{b}_2,\ldots,\hat{b}_k$, one has to be able to describe the design in terms of the full rank model, $E(y) = Ub$, $b' = (b_1,\ldots,b_k)$. If λ_i, $i=1,\ldots,k$, are the latent roots of $(U'U)$, then the design that is A-optimal for the set of estimable functions $b_1,\ldots,b_k$ is the design with the minimum sum, $\Sigma \frac{1}{\lambda_i}$.

However, this approach requires some description of the construction of $(U'U)$. Given that $h(T^*)$ or $h(D^*)$ is to be a minimum number, the description of $(U'U)$ is difficult in general terms. Therefore, the construction of a set Ψ from which one identifies an A-optimal design for estimable functions $b_1,\ldots,b_k$ is not apparent.

Let us next investigate some methods for constructing Ψ as indicated in section 5.5.

5.7 Procedure I

In this section we will present a method for constructing members of the set Ψ used to find optimal MAMD's. This method also presents a way of constructing a set Ψ to be used for finding optimal MMD's. In constructing the set Ψ for MAMD's, the starting point T_{Ψ_a}, the design to be augmented, is given. In constructing Ψ for MMD's a starting point T_{Ψ_b} has to be created.

In creating T_{Ψ_b}, we shall simultaneously demonstrate how the D_i of T_i in the set Ψ for MAMD's may be constructed in levels from the sets S_β, $\beta \neq \alpha$. Then in section 5.8, we will indicate ways that levels of S_α may be applied to the D_i for MAMD's or to the design T_{Ψ_b} for MMD's.

For the sake of notation, in Procedure I (given in this section) and in Procedure II (given in section 5.8) we will assume we are constructing main effect designs; i.e. we have sets of levels (instead of level combinations) S_β, $\beta = 1,\ldots,m$, only. Without loss of generality, the construction of designs with sets S_β, $\beta =1,\ldots,M^*(T)$, follows in the same manner, as we will demonstrate by the examples given in section 5.9. For the following definition, the reader should recall that $h_{\ldots \ell_\beta, \ldots}$ is the frequency of level ℓ_β of set S_β.

Design Class I

A MMD in h assemblies with factors $F_1, \ldots, F_m$ with factor F_β having n_β levels, is said to belong to Design Class I if, for $\beta = 1, \ldots m$,

$$(5.7.1) \qquad \begin{aligned} &(1) \quad h = k_\beta n_\beta + a_\beta \qquad \text{and} \\ &(2) \quad h_{\ldots \ell_\beta \ldots} = k_\beta \qquad \text{for } n_\beta - a_\beta \text{ levels of } S_\beta \\ &\qquad\qquad\quad\; = k_\beta + 1 \quad \text{for } a_\beta \text{ levels of } S_\beta \end{aligned}$$

Procedure I, though not too difficult to apply is difficult to describe in general terms. Therefore, it may be helpful to the reader to refer to Example 5.2 and other examples as often as necessary in this and in the upcoming section.

A Summary of Procedure I

1. In order to construct a MMD in m factors, we first construct an MMD in two factors, then an MMD in three factors,..., and finally an MMD in m factors.

2. In each case we use the same method:

   a) To construct a MMD of k factors we "augment" a design of k-1 factors.

   b) Starting with a design of k-1 factors $F_1, F_2, \ldots, F_{k-1}$, in $h^{(k-2)}$ assemblies, $n_k - 1$ additional assemblies in k - 1 factors are to be added to the original $h^{(k-2)}$ such that (5.7.1) holds for all k - 1 factors.

c) These $n_k - 1$ additional assemblies should be repeats of $n_k - 1$ assemblies of the original $h^{(k-2)}$. There are two reasons for this

1) In doing so, (5.7.1) is easier to maintain

2) Levels of S_k can be appended to the $h^{(k-2)} + n_k - 1$ factors in a simple manner.

d) If c) can not be attained, the some of the $n_k - 1$ assemblies have to be new assemblies in $k - 1$ factors. These new assemblies have to be chosen so that (5.7.1) is attained and levels of S_k can be easily connected in the design of $h^{(k-2)} + n_k - 1$ assemblies of k factors.

Procedure I For Constructing Designs of Class I

$m=2$

We will first construct a main effect plan for a $n_1 \times n_2$ factorial in $h^{(1)}$ assemblies, where $h^{(1)} = n_1 + n_2 - 1$. A design in levels of S_1 and S_2 is constructed accordingly:

(a) For the purposes of describing Procedure I, we will first choose the set S_ℓ, $\ell = 1$ or 2 with the largest n_ℓ. Without loss of generality, let $n_1 \geq n_2$. Also let $\phi_i, i=1,\ldots,n_1$ be the levels of set S_1. We may then list ϕ_i in a column of $h^{(1)}$ elements, where $h^{(1)} = a_1^{(1)} (\text{mod } n_1)$. For this case of two factors $a_1^{(1)} = n_2 - 1$:

$$(5.7.2)\qquad \left.\begin{array}{l} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_{n_1} \\ \phi_1 \\ \vdots \\ \phi_{a_1} \end{array}\right] h^{(1)} \quad \text{Levels of } S_1$$

Note that $h_{i.} = k_1^{(1)} + 1$ for $a_1^{(1)}$ levels of S_1

$= k_1^{(1)}$ for $n_1 - a_1^{(1)}$, levels of S_1

with $k_1^{(1)} = 1$.

(b) Next we associate levels of S_2, θ_j, $j=1,\ldots,n_2$, with the $h^{(1)}$ levels of S_1, in such a manner that all differences $\theta_j - \theta_j{}'$ are estimable. This may be achieved in many ways. We shall present an insight to some of these methods.

(i) If $n_1 = n_2$, then one approach may be as follows

$$(5.7.3)\qquad \left.\begin{array}{ll} \phi_1 & \theta_1 \\ \phi_2 & \theta_2 \\ \vdots & \vdots \\ \phi_{n_1} & \theta_{n_1} \\ \phi_1 & \theta_2 \\ \phi_2 & \theta_3 \\ \vdots & \vdots \\ \phi_{a_1} & \theta_{a_1+1} \end{array}\right] h^{(1)} \quad \begin{array}{l}\text{Assemblies In Levels} \\ \text{Of } S_1 \text{ and } S_2\end{array}$$

Note that all contrasts in levels of set S are now estimable (i.e. assemblies $\Phi_1\theta_1$ and $\Phi_1\theta_2$ can be used to estimate $\theta_1 - \theta_2$, etc.) and therefore by theorem 4.5, so are all contrasts in levels of S_1. Note

$$\text{also } h_{\cdot j} = k_2^{(1)} + 1 \quad \text{for } (a_2^{(1)} = a_1^{(1)}) \text{ levels of } S_2$$
$$= k_2^{(1)} \quad \text{for } n_2 - a_2^{(1)} \text{ levels of } S_2$$

(And as a consequence of $m = 2$, $k_2^{(1)} = k_1^{(1)} = 1$)

(ii) If $n_1 > n_2$: here we will introduce a method that one can use for later cases of $m > 2$ factors. We shall also use this method on factor level combinations in establishing Procedure II in section 5.8.

First it should be mentioned again that $h^{(1)} = n_1 + n_2 - 1$; and $h^{(1)} = n_2 + a_2^{(1)}$. The levels of S_2 will be appended to levels of S_1 so as to keep

$$h_{\cdot j} = k_2^{(1)} + 1 \quad \text{for } a_2^{(1)} \text{ levels of } S_2$$
$$= k_2^{(1)} \quad \text{for } n_2 - a_2^{(1)} \text{ levels of } S_2$$

(Note, once more, that here $k_2^{(1)} = 1$).

One way of doing this may be as follows:

$$(5.7.4)\qquad \left.\begin{array}{ll} \Phi_1 & \theta_1 \\ \Phi_2 & \theta_2 \\ \cdot & \cdot \\ \Phi_{n_2} & \theta_{n_2} \\ \Phi_{n_2+1} & \theta_{\omega_1} \\ \cdot & \cdot \\ \Phi_{n_1} & \theta_{\omega_{n_1-n_2}} \\ \Phi_1 & \theta_2 \\ \Phi_2 & \theta_3 \\ \Phi_3 & \theta_4 \\ \cdot & \cdot \\ \Phi_{a_1^{(1)}} & \theta_{a_1^{(1)}+1} \end{array}\right] - h^{(1)} \text{ assemblies in levels of } S_1 \text{ and } S_2$$

A discussion of $\theta_{\omega_1}, \ldots, \theta_{\omega_{n_1-n_2}}$ is relevant here. These levels of S_2 will be chosen by the experimenter so that $h_{.j} = k_2^{(1)} + 1$ or $k_2^{(1)}$. The last $a_1^{(1)} = n_2 - 1$ assemblies in S_1 and S_2 were chosen so that differences, $\theta_j - \theta_{j'}$ of S_2 are estimable. Set S_1 is then connected by Theorem 4.5.

In constructing designs of m factors, this last method will be used frequently. In place of Φ_i (the levels of S_1) we will be discussing Φ_i (level combinations of m-1 factors). Then the levels θ_j will be of set S_m and our concern will be connecting set S_m. More shall be said on this later, but notice here that since the levels ϕ_i, $i=1,\ldots,a^{(1)}$ are repeated, set S_2 is easily connected. Therefore in later discussions, we would like the level combinations Φ_i, to be repeated, so that S_m is easily connected. We shall use this idea for m = 3.

m = 3

For a minimal effect plan for a $n_1 \times n_2 \times n_3$ factorial, $h^{(2)} = n_1 + n_2 + n_3 - 2 = h^{(1)} + n_3 - 1$ assemblies are required. For three factors we proceed as if we were construcing a MMD in two factors. For the $h^{(1)}$ assemblies of (5.7.4) let $\Phi_1 = (\phi_1 \theta_1)$, $\Phi_2 = (\phi_2 \theta_2),\ldots,$ $\Phi_{h^{(1)}} = (\phi_{a_1^{(1)}} \theta_{a_1^{(1)}+1})$. We will first get $h^{(2)}$ assemblies in the first two factors. The $h^{(2)}$ assemblies in S_1 and S_2 must be such that (5.7.1) holds and levels of set S_3 may be easily appended and connected. Therefore we must add $n_3 - 1$ assemblies in levels of sets S_1 and S_2 to the $h^{(1)}$ assemblies such that the last statement holds. Let these $n_3 - 1$ new assemblies in S_1 and S_2 be represented by Φ_j^*, $j=1,\ldots,n_3 - 1$. As just discussed, it would be highly desirable if we could achieve (5.7.1) by merely repeating the first $n_3 - 1$ assemblies in Φ_i, i.e. $\Phi_1^* = \Phi_1$, $\Phi_2^* = \Phi_2,\ldots,$ etc. Then we could connect S_3 as we connected S_2 in the m = 2 factor case. However, it may not always be possible that the Φ_i level combinations be repeated in all the $n_3 - 1$ additional assemblies of S_1 and S_2.

(Example 5.3 indicates this for sets S_1, and S_2, and S_3, in (5.8.2).)

Therefore some of the Φ_j^* may not be equal to any of the Φ_i, $i=1,\ldots h^{(1)}$, assemblies given before. If this is the case, some adjustments have to be made in connecting S_3. (S_3 is not as easily connected as in the case where all Φ_j^*'s are repeated Φ_i's.) In Table III, section 5.8, we shall give an example, of such a case for m=4. For this type of situation, one merely makes use of the Φ_j^*'s that are repeated Φ_i's and thus connects some levels of S_3, and uses these connected levels of S_3 to connect levels of S_2 and S_1. More shall be said of this in Procedure II, section 5.8.

Now the $h^{(2)}$ assemblies in S_1 and S_2 may be expressed, using Φ_j^*, $j=1,\ldots,n_3 - 1$ to represent the new assemblies in S_1 and S_2:

$$(5.7.5)\qquad \left.\begin{array}{l} \Phi_1 \\ \Phi_2 \\ \cdot \\ \Phi_{h^{(1)}} \\ \Phi_1^* \\ \cdot \\ \Phi_{n_3-1}^* \end{array}\right] \text{— } h^{(2)} \text{ assemblies in } S_1 \text{ and } S_2 \text{ only}$$

As mentioned, Φ_j^* are such that (5.7.1) holds for S_1 and S_2 and Φ_j^* are chosen such that S_3 can be connected. Then the levels of $S_3, \alpha_1, \alpha_2, \ldots, \alpha_{n_3}$ may be appended to (5.7.5) in the following manner.

$$
(5.7.6) \qquad \left.\begin{array}{ll}
\Phi_1 & \alpha_1 \\
\Phi_2 & \alpha_2 \\
\cdot & \cdot \\
\Phi_{h^{(1)}} & \alpha_q \\
\Phi_1^* & \alpha_2 \\
\cdot & \cdot \\
\Phi_{n_3-1}^* & \alpha_{n_3}
\end{array}\right] \text{— } h^{(2)} \text{ assemblies in } S_1,\ S_2, \text{ and } S_3
$$

A discussion of $\alpha_1, \ldots, \alpha_q$ is relevant here. These levels of S_3 are to be chosen by the experimenter so that (5.7.1) holds for S_3. The choosing of levels $\alpha_1, \ldots, \alpha_q$ in conjunction with assemblies Φ_j^* determines how S_3 is connected. (See 5.7.10 in Example 5.2 for an illustration). In this way we can construct a main effect plan for a $n_1 \times n_2 \times n_3$ factorial in $h^{(2)} = \sum_{i=1}^{3} n_i - 2$ assemblies and still retain (5.7.1) for S_1, S_2 and S_3. If one proceeds in a similar manner for m factors, a main effect design $h = h^{(m-1)} = \sum_{i=1}^{m} n_i - m + 1$ assemblies can be developed.

The case for $m = 3$ is illustrated in Example 5.2.

Example 5.2

A plan of investigation requires a main effect plan for a $2 \times 3 \times 4$ factorial in $2 + 3 + 4 - 2 = 7$ assemblies.

We first establish a main effect design for a 2×3 factorial in $h^{(1)} = 2 + 3 - 1 = 4$ assemblies. In order to agree with (5.7.4) and for the purpose of illustration, we shall consider the plan for a $3 \times 2 \times 4$ factorial, and consider the $h^{(1)} = 4$ assemblies of a 3×2 factorial first.

Let the levels of S_1 be 0, 1, 2, and of S_2 be 0, 1. Then according to (5.7.3) we have

$$\left.\begin{matrix} 0 \\ 1 \\ 2 \\ 0 \end{matrix}\right\} h^{(1)} = 4 \text{ levels of } S_1 \tag{5.7.7}$$

Then according to (5.7.4)

$$\left.\begin{matrix} 00 \\ 11 \\ 20 \\ 01 \end{matrix}\right\} h^{(1)} = 4 \text{ level combinations of } S_1 \text{ and } S_2 \tag{5.7.8}$$

Note that if we had followed (5.7.4) exactly we would have

$$\begin{matrix} 00 \\ 11 \\ 20 \\ 00 \end{matrix}$$

where the (00) combination is repeated, and this is not desired, since the level 0 of set S_2 occurs 3 times and level 1 occurs once, which is not in accordance with (5.7.1).

Since $h^{(1)} = 0 \pmod 2$, we find $h_{\cdot j} = 2 = k_2^{(1)}$ for 2 levels of S_2, and $h_{\cdot j} = 3$ for no level of S_3; and $h^{(1)} = 1 \pmod 3$, implies $h_{i\cdot} = 1 = k_1^{(1)}$ for 2 levels of S_1 and $h_{i\cdot} = 2$ for 1 level of S_1.

The design in (5.7.8) is also connected since if we represent an observation on assembly (01) as y_{01}, then

$$E[y_{01} - y_{11}] = p_1^0 - p_1^1 ,$$

$$E[y_{00} - y_{01}] = p_2^0 - p_2^1 ,$$

$$E[y_{00} - y_{20}] = p_1^0 - p_1^2 .$$

For a 3 x 2 x 4 factorial, $h^{(2)} = 7$ assemblies are required. Therefore to (5.7.8) we add three more assemblies as we did in (5.7.6): here $\Phi_1^* = \Phi_1$, $\Phi_2^* = \Phi_2$ and $\Phi_3^* = \Phi_3$,

$$(5.7.9) \qquad \left.\begin{array}{l} \left.\begin{array}{c} 00\\ 11\\ 20\\ 01 \end{array}\right] h^{(1)} = 4 \\ \left.\begin{array}{c} 00\\ 11\\ 20 \end{array}\right] n_3 - 1 = 3 \end{array}\right] h^{(2)} = 7 \text{ assemblies in } S_1 \text{ and } S_2 .$$

Now append all the n_3 levels of S_3 to the first n_3 level combinations of (5.7.9) and then append $n_3 - 1$ levels of S_3 to the last $n_3 - 1$ level combinations of (5.7.9). From this we have a design connected in S_3 and thus connected in S_1 and S_2:

$$(5.7.10)\qquad \left.\begin{array}{c}000\\111\\202\\013\\001\\112\\203\end{array}\right] \text{— } h^{(2)} = 7 \text{ assemblies in } S_1, S_2 \text{ and } S_3 \text{: A minimum main effect plan for a 3x2x4 factorial.}$$

According to (5.7.6), $\alpha_1 = 0$, $\alpha_2 = 1$, $\alpha_3 = 2$, $\alpha_4 = 3$, $q = 4$.

Note that for (5.7.10) we have $h^{(2)} = 1 \pmod 2$, $h^{(2)} = 1 \pmod 3$, and $h^{(2)} = 3 \pmod 4$, which implies $k_1^{(2)} = 2$, $k_2^{(2)} = 3$, $k_3^{(2)} = 1$.

Hence

$$\begin{aligned} h_{i\cdot\cdot} &= 3 \text{ for } i = 0\\ &=2 \text{ for } i = 1,2 \\[6pt] h_{\cdot j\cdot} &= 4 \text{ for } j = 0\\ &= 3 \text{ for } j = 1 \\[6pt] h_{\cdot\cdot\ell} &= 2 \text{ for } \ell = 1,2,3\\ &= 1 \text{ for } \ell = 0 \quad . \end{aligned}$$

which satisfies the characteristics of Class I designs.

5.8 Procedure II

Using Procedure I we may now construct a design T_{ψ_b} (or a sequence of assemblies, D_i) in $m - 1$ factors. We will next consider methods to append levels of S_α to the assemblies of T_{ψ_b} (or to the assemblies of D_i). Without loss of generality, and to ease the presentation of our results, we shall consider appending levels of S_m to the design

T_{Ψ_b} constructed in m - 1 factors.

One method of appending levels of S_m to T_{Ψ_b}, is Procedure I. Using Procedure I, a_m levels of S_m would occur $k_m + 1$ times and $n_m - a_m$ levels of S_m would occur k_m times. Another method of appending levels of S_m to T_{Ψ_b} will be Procedure II (soon to be described). Procedure II will require $n_m - 1$ levels of S_m to occur once and 1 level of S_m to occur $h - n_m + 1$ times.

Given designs constructed by Procedure I and Procedure II, we may then vary the association of levels of S_m in combination with the levels of the other m - 1 factors and in this way, construct still other members of set Ψ. In each case, the levels of S_m will have a different occurrence and frequency of occurrence with the level combinations of the other factors.

Designs to be attained by Procedure II will be identified by Design Class II.

Design Class II

A MMD in h assemblies with factors $F_1, F_2, \ldots, F_m$ with factor F_β having n_β levels is said to belong to Design Class II if for $\beta = 1, \ldots, m-1$, the design is a member of Design Class I, and

$$h_{\ldots \ell m} = h - n_m + 1 \qquad \text{for 1 level of } S_m$$
$$= 1 \qquad \text{for } n_m - 1 \text{ levels of } S_m.$$

Procedure II, though not too difficult to apply, is difficult to describe in general terms. Therefore, it may help the reader to refer to Example 5.3 and other examples as often as necessary in this section.

A Summary of Procedure II

1. From a MMD of $h - n_m + 1$ assemblies of $m - 1$ factors, T_{ψ_b}, a MMD of Design Class II in h assemblies is to be constructed.

2. To T_{ψ_b} add $n_m - 1$ assemblies in $m - 1$ factors such that (5.7.1) holds for all $m - 1$ factors in h assemblies.

3. Append one level of S_m to the first $h - n_m + 1$ assemblies and the remaining $n_m - 1$ levels of S_m to the remaining $n_m - 1$ assemblies.

4. The set S_m is then automatically connected in the h assemblies which form a MMD of Class II.

Procedure II For Constructing Designs of Class II

We are now going to describe Procedure II in more detail. Assume a design of $m - 1$ factors constructed by Procedure I in $h^{(m-2)} = \sum_{i=1}^{m-1} n_i - m + 2$ assemblies. Denote the $h^{(m-2)}$ assemblies by Φ_i, $i=1,\ldots,h^{(m-2)}$. We will need $n_m - 1$ additional assemblies Φ_j^*'s in $m - 1$ factors in order to connect S_m in $h^{(m-1)} = h^{(m-2)} + n_m - 1$ assemblies. As before, we would like to repeat as many Φ_i's as possible in the Φ_j^*'s. If not all Φ_j^* can be repeated Φ_i''s ,then some of the Φ_j^*'s need to be constructed.

How To Construct "New" Φ_j^*'s

One approach is as follows: Ignore, for the moment, levels of S_{m-1}. It may be possible to repeat level combinations of m - 2 factors of the $h^{(m-2)}$ assemblies in the necessary Φ_j^*'s so as to retain the characteristics of (5.7.1) for the m - 2 factors. Assume that this is true. Then levels of S_{m-1} have to be chosen according to the frequencies of (5.7.1) and appended to the repeated level combinations of m - 2 factors. (See assemblies (xi) and (xii) of (5.8.2).) Thus n_m - 1 additional assemblies of m - 1 factors can be constructed according to (5.7.1)

If it is not possible to repeat level combinations of m - 2 factors in the Φ_j^*'s, then we would ignore levels of sets S_{m-2}, S_{m-1}, and seek to repeat level combinations of m - 3 factors while maintaining (5.7.1). We would then have to append according to (5.7.1) levels of sets S_{m-2} and S_{m-1} to the level combinations of m - 3 factors. This method could be repeated again. Once the Φ_j^*'s are constructed, the next step is to append levels of S_m.

Connecting Levels of S_m

Let $\theta = 1,\ldots,n_m$, be the levels of S_m. Let us append these levels to Φ_i, and the n_m - 1 added assemblies, Φ_i^*, in the following way:

$$\underline{h^{(m-1)} \text{ assemblies in } m \text{ factors}}$$

(5.8.1)
$$\left.\begin{array}{ll} \left.\begin{array}{ll} \Phi_1 & \theta_1 \\ \Phi_2 & \theta_1 \\ \vdots & \\ \Phi_{h^{(m-2)}} & \theta_1 \end{array}\right\} & h\ldots_z = h^{(m-2)} \quad \text{for } z=1, \\ \left.\begin{array}{ll} \Phi^*_1 & \theta_2 \\ \cdot & \cdot \\ \Phi^*_{n_{m-1}} & \theta_{n_m} \end{array}\right\} & h\ldots_z = 1 \quad \text{for } z=2,\ldots,n_m \end{array}\right.$$

The reason behind this construction is obvious. The first $h^{(m-2)}$ assemblies serve as a connected design in the first m-1 factors, since the same comparisons in $h^{(m-2)}$ assemblies can be made as before. Therefore S_m is automatically connected. Also, because of the particular way in which the Φ^*_i are chosen, for $h^{(m-1)} = a_\beta \pmod{n_\beta}$, we have

$$\begin{aligned} h\ldots_{\ell_\beta}\ldots &= k_\beta + 1 && \text{for } a_\beta \text{ levels of } S_\beta \\ &= k_\beta && \text{for } n_\beta - a_\beta \text{ levels of } S_\beta \end{aligned}$$

Notice that all that is required of the Φ^*_j's is that they satisfy (5.7.1) . In constructing the Φ^*_j's for an m-factor design according to Procedure I, the Φ^*_j's must not only retain the requirements of (5.7.1), but they must also help in connecting S_m. This is easily done if all the Φ^*_j's are repeated Φ_i's as mentioned earlier. However, when this is not true, we could refer to the method just discussed where one ignores sets $S_{m-1}, S_{m-2},\ldots,S_{m-k}$, repeats assemblies of m - k - 1 factors such that (5.7.1) holds, and then (1) append levels of S_{m-1}, $S_{m-2}, \ldots, S_{m-k}$ each according to (5.7.1), but (2) with some intent of using the assemblies, Φ_j^*, $j=1,\ldots,n_m - 1$, to eventually connect S_m. (See Tables II and III and design T_4 of Example 5.4). This method is not too difficult to perform, since in "sequentially" constructing m factor designs according to Procedure I, we repeat as many assemblies as possible and as often as possible.

Now that we have Procedures I and II we may use them to construct a set Ψ from which we identify a Ψ- optimal MAMD . The following examples shall illustrate this.

Example 5.3

Consider Example 5.2. If another factor, F_4, at 6 levels is required in the experiment, and if particular interest is centered on F_4, a MMD may be constructed using h = 12 assemblies or 5 more than in (5.7.10). We first need to construct a design of three factors in 12 assemblies. To the 7 assemblies of (5.7.10) we add 5;

(5.8.2)

Number	Assembly	
(i)	(000)	$h^{(2)}$ = 7 assemblies in S_1, S_2 and S_3,
(ii)	(111)	
(iii)	(202)	
(iv)	(013)	
(v)	(001)	
(vi)	(112)	
(vii)	(203)	
(viii)	(202)	n_4 - 1 = 5 assemblies in S_1, S_2, and S_3,
(ix)	(013)	
(x)	(111)	
(xi)	(110)	Φ_j^*'s not repeated Φ_i's.
(xii)	(200)	

Table I

A Main Effect Plan For a 3x2x4x6

Factorial in 12 Assemblies: T_1

Assembly Number	The Assemblies
(i)	(0 0 0 0)
(ii)	(1 1 1 0)
(iii)	(2 0 2 0)
(iv)	(0 1 3 0)
(v)	(0 0 1 0)
(vi)	(1 1 2 0)
(vii)	(2 0 3 0)
(viii)	(2 0 2 1)
(ix)	(0 1 3 2)
(x)	(1 1 1 3)
(xi)	(1 1 0 4)
(xii)	(2 0 0 5)

Assemblies (xi) and (xii) are not given in (5.7.10), but were constructed in agreement with

$$h_{i}.. = 6 \qquad \text{for all } i$$

$$h._{j}. = 4 \qquad \text{" " } i$$

$$h..k = 3 \qquad \text{" " } k$$

since $h^{(3)} = 0$ (mod 3, 2, 4.)

Also, assemblies (xi) and (xii) do repeat (11) and (20), factor level combinations that were used to construct the 7 assemblies in (5.7.10). In repeating (11) and (20), we make the construction of other designs in Ψ easier, as we shall see in constructing T_4 in Table III of Example 5.4. To finish the illustration of Procedure II, Table I now follows by repeating level 0 of S_4.

Example 5.4

We shall present a set $\Psi = [T_1, \ldots, T_6]$ in this example. T_1 is given in Table I. T_2 will be developed by Procedure I. T_3 will follow from T_2 by interchanging levels of S_4 in T_2. T_4 will have a unique construction. T_5 and T_6 then follow from T_4, by changing levels of S_4.

We shall use (5.8.2) which was essentially developed by Procedure I in order to illustrate the construction of T_2 and T_4. For the given level combinations of S_1 and S_2, let A = (00), B = (11), C = (20), and D = (01). Then (5.8.2) becomes

	Number	Assembly
	(i)	A0
	(ii)	B1
	(iii)	C2
(5.8.3)	(iv)	D3
	(v)	A1
	(vi)	B2
	(vii)	C3
	(viii)	C2
	(ix)	D3
	(x)	B1
	(xi)	B0
	(xii)	C0

Next, in order to assign levels of S_4 to (5.8.3), we note that if we connect any two of the sets S_Φ (A,B,C,D), S_3(0,1,2,3) or S_4(0,1,...,5), the third will be connected by Theorem 4.5.

T_2 will be the most difficult to construct for it is of Design Class I. From (5.8.3), consider assemblies [(i), (xii)], [(ii),(v)] and [(vii), (ix)]. To each pair, append one different level of S_4. All levels of the S_Φ(A,B,C,D) are then connected. We now can ignore S_Φ. To each pair of assemblies, [(vi), (xi)], [(iii), (x)], and [(viii), (iv)], append one different level of the remaining three levels of S_4. Now S_3 is connected, and therefore so is S_4. Ignoring sets S_3, S_4, it then follows that sets S_1 and S_2 are now connected by (5.7.8).

T_4 is constructed a little differently. First, some levels of S_4 are connected, then these levels are used to connect S_Φ and S_3. Then, by Theorem 4.5, all levels of S_4 will be connected. By making use of assemblies (ii), (iii) and (iv) and the repetition of B,C,D, in(x) (viii), and (ix) respectively, we connect levels 0,1,2, and 4 of set S_4, as shown in Table II and use this relationship to connect S_Φ

Table II

The Construction of T_4

Assembly Number	Levels 0,1,2,4 of S_4 Connected	All levels of S_Φ, S_5 Connected
(i)	.............(A0)	(A02)
(ii)	.............(B10)	(B10)
(iii)	.............(C22)	(C22)
(iv)	.............(D30)	(D30)
(v)	.............(A1)	(A14)
(vi)	.............(B2)	(B2)
(vii)	.............(C3)	(C32)
(viii)	.............(C20)	(C20)
(ix)	.............(D34)	(D34)
(x)	.............(B11)	(B11)
(xi)	.............(B0)	(B0)
(xii)	.............(C0)	(C00)

and S_3.

In order to demonstrate how we connect S_Φ and S_3 we now refer to Table III. The primary assemblies of Table III give the comparisons necessary to estimate the differences in levels of each set, and once a set is connected, its levels then may be ignored. The secondary assemblies are those comparisons involving the four connected levels of S_4 that support the primary assemblies. For example, assemblies (ii) and (v) estimate the difference in the levels of A and B of S_Φ, because the difference in levels 2 and 4 of S_4 is made estimable by assemblies (iv) and (ix).

Note that without appending any level of S_4 to (vi) and (xi), T_4 is connected. The levels 3 and 5 of S_4 may be appended now to BO and B2 abritrarily. Thus all sets S_Φ, S_3 and S_4 are connected in T_4. Now once again ignore levels of S_3 and S_4. Then in Table IV, sets S_1 and S_2 are connected by (5.7.8), and T_4 is connected.

Design T_3 follows from T_2 by interchanging levels of S_4 in T_2 in the following manner, $0\to1$, $1\to3$, $2\to4$, $3\to2$, $4\to5$, $5\to0$. T_5 follows from T_4 by replacing the levels of S_4 in (v), (i), and (vii) with 0. T_6 follows from T_5 by changing the first assembly to (0002).

Now that we have constructed a set of designs, Ψ, we may identify a Ψ-optimal design. If λ_i, $i=1,\ldots,5$ are the non-zero latent roots for each matrix C_4 in design T_j, $j=1,\ldots,6$, then we see in Table V that T_2 and T_3 are Ψ-optimal for S_4.

Table III

Assemblies That Connect S_Φ, S_3

Set	Levels Connected	Primary Assemblies	Secondary Assemblies
S_Φ	A, B	(ii), (v)	(iv), (ix)
S_Φ	A, C	(i), (xii)	(iii),(viii)
S_Φ	C, D	(vii),(ix)	(iii),(viii), (iv), (ix)
S_3	0, 1	(i), (iii)	(iii), (vii)
S_3	1, 2	(iii), (x)	(ii), (x), (iii),(viii)
S_3	2, 3	(iii), (vii)	——

Table IV

A Set $\Psi^{\#}$ For a 3x2x4x6 Factorial

Assembly Number	T_1	T_2	T_3
(i)	(0000)	(0005)	(0000)
(ii)	(1110)	(1110)	(1111)
(iii)	(2020)	(2021)	(2023)
(iv)	(0130)	(0132)	(0134)
(v)	(0010)	(0010)	(0011)
(vi)	(1120)	(1124)	(1125)
(vii)	(2030)	(2033)	(2032)
(viii)	(2021)	(2022)	(2024)
(ix)	(0132)	(0133)	(0132)
(x)	(1113)	(1111)	(1113)
(xi)	(1104)	(1104)	(1105)
(xii)	(2005)	(2005)	(2000)

\# with level frequencies of S_4 :

ℓ		$h_{\cdots\ell}$	
0	7	2	2
1	1	2	2
2	1	2	2
3	1	2	2
4	1	2	2
5	1	2	2

Table IV (continued)

Assembly Number	T_4	T_5	T_6
(i)	(0002)	(0000)	(0002)
(ii)	(1110)	(1110)	(1110)
(iii)	(2022)	(2022)	(2022)
(iv)	(0130)	(0130)	(0130)
(v)	(0014)	(0010)	(0010)
(vi)	(1125)	(1125)	(1125)
(vii)	(2032)	(2030)	(2030)
(viii)............	(2020)	(2020)	(2020)
(ix)	(0134)	(0134)	(0134)
(x)	(1111)	(1111)	(1111)
(xi)	(1103)	(1103)	(1103)
(xii)	(2000)	(2000)	(2000)

with level frequencies of S_4 :

ℓ		$h_{\cdot\cdot\cdot\ell}$	
0	4	7	6
1	1	1	1
2	3	1	2
3	1	1	1
4	2	1	1
5	1	1	1

Table V

The Sum, $\Sigma\frac{1}{\lambda_i}$, For Members Of, Ψ, Example 5.4

T_i	T_1	T_2	T_3	T_4	T_5	T_6
$\Sigma\frac{1}{\lambda_i}$	13.34	12.00	12.00	18.36	13.34	15.92

We have just presented an example of how one may construct a set Ψ. More examples shall be given in section 5.9 where Ψ-optimal designs with assumed two factor interactions are presented.

5.9 Examples of Minimal Designs With Two-Factor Interactions

Some MMD's with two-factor interactions present are easy to construct. The process follows the construction of main effect designs indicated in sections 5.7 and 5.8 Other MMD's with interaction do not have ease of construction, but may be constructed using some guidelines of optimality. Examples given in this section will be for different cases of 3^n factorials, since these plans are easy to illustrate and understand. However, these methods extend easily to more complex cases, i.e. mixed factorials.

Case (1): A 3^3 factorial under the model $E(\underline{y}) = X_{12}\underline{p}_{12} + X_3\underline{p}_3$

For this plan a 9x3 main effect plan is first constructed, since there are 9(=3x3) combinations of S_{12} and 3 levels of S_3. For this design, 11 assemblies are required. We need to first construct members of the set Ψ. Then find a Ψ-optimal design, T^*. This is done in Appendix II, and the Ψ-optimal design for S_{12} is given in Table VI. Once a Ψ-optimal MMD for set S_{12} is found, assemblies in sets S_1, S_2 and S_3 follow by the correspondence indicated in Table VI.

Coincidentally, the design indicated by Table VI may also be viewed as a Ψ-optimal MAMD (as indicated in Appendix II). The first nine

Table VI

A Ψ-Optimal Design For Case (1)

Assembly Number		(a) Assemblies of S_{12}, S_3	(b) Assemblies of S_1, S_2, S_3 #
(i)	....	(0 0)	(0 0 0)
(ii)	....	(1 0)	(0 1 0)
(iii)	....	(2 0)	(0 2 0)
(iv)	....	(3 1)	(1 0 1)
(v)	....	(4 1)	(1 1 1)
(vi)	....	(5 1)	(1 2 1)
(vii)	....	(6 2)	(2 0 2)
(viii)	....	(7 2)	(2 1 2)
(ix)	....	(8 2)	(2 2 2)
(x)	....	(1 1)	(0 1 1)
(xi)	....	(1 2)	(0 1 2)

Column (b) follows from (a) using the correspondence: $S_{12} \to S_1 \times S_2 : 0 \to (00),\ 1 \to (01),\ 2 \to (02),\ 3 \to (10)$, etc.

assemblies of column (b) form a 1/3 replicate of a 3^3 factorial that may serve as an orthogonal main effect plan. The last two assemblies then "augment" this plan for the presence of a factor one and two interaction.

Also, column (a) of Table VI was constructed (using Procedure II) so that estimates of all differences $p_{12}^{u_1u_2} - p_{12}^{v_1v_2}$, $(u_1,u_2) \neq (v_1,v_2)$ can be found. Although these differences may not be meaningful, they do permit one to apply Theorem 4.4. Thus set S_{12} can be handled as if it were a set of levels rather than level combinations. This idea is used again in Case (2) and later in Case (5).

Case (2): A 3^4 factorial under the model $E(y) = X_{12}p_{12} + X_{34}p_{34}$

For this case let contrasts of $p_{34}^{\ell k}$ be considered "more important" than those of p_{12}^{ij}. We will then determine a set Ψ and from Ψ find a design which is Ψ-optimal for the set S_{34}. Using Procedure I and Procedure II we need to find 17 assemblies of levels of S_{12} and S_{34}, i.e. $h^{(1)} = 9 + 9 - 1 = 17$. A set Ψ for Case (2) is given by Table VII. T_1 is constructed using Procedure II and T_2, T_3, and T_4 follow from Procedure I and by varying levels of S_{34}.

For T_ℓ, $\ell = 1,\ldots,4$

$$\begin{aligned} h_{i\cdot} &= 1, \qquad i = 8 \\ &= 2, \qquad \text{otherwise} \end{aligned}$$

and for T_1

$$h_{\cdot j} = 9, \qquad j = 8$$

Table VII

A Set Ψ For Case (2)

Assembly Number	T_1	T_2	T_3	T_4
(i)	(0 0)	(0 0)	(0 0)	(0 0)
(ii)	(1 1)	(1 1)	(1 1)	(1 1)
(iii)	(2 2)	(2 2)	(2 2)	(2 2)
(iv)	(3 3)	(3 3)	(3 3)	(3 3)
(v)	(4 4)	(4 4)	(4 4)	(4 4)
(vi)	(5 5)	(5 5)	(5 5)	(5 5)
(vii)	(6 6)	(6 6)	(6 6)	(6 6)
(viii)..........	(7 7)	(7 7)	(7 7)	(7 7)
(ix)	(8 8)	(8 8)	(8 8)	(8 8)
(x)	(0 8)	(0 7)	(0 1)	(0 7)
(xi)	(1 8)	(1 8)	(1 2)	(1 8)
(xii)	(2 8)	(2 7)	(2 3)	(2 6)
(xiii)..........	(3 8)	(3 8)	(3 4)	(3 7)
(xiv)	(4 8)	(4 7)	(4 5)	(4 8)
(xv)	(5 8)	(5 8)	(5 6)	(5 6)
(xvi)	(6 8)	(6 7)	(6 7)	(6 7)
(xvii)..........	(7 8)	(7 8)	(7 8)	(7 8)

$h._j = 1$ otherwise

For T_2

$h._j = 5$, $j = 7, 8$

$= 1$, otherwise

For T_3 $h._j = 1$, $j = 0$

$= 2$, otherwise

and for T_4 $h._j = 4$, $j = 7, 8$

$= 3$, $j = 6$

$= 1$, otherwise

By using Table VIII we find that T_3 is Ψ-optimal for set S_{34}. The design T_3 of sets S_1, S_2, S_3 and S_4 follows from the correspondence given in Table IX.

Given that set S_{12} is connected, we can ignore it. Then we are left with 18 assemblies of sets S_3 and S_4. Set S_3 is obviously connected and therefore so is set S_4. We may do the same for showing that sets S_1 and S_2 are connected.

Case (3): A 3^3 factorial under the model $E(\underset{\sim}{y}) = X_{12}\underset{\sim}{p}_{12} + X_{13}\underset{\sim}{p}_{13}$

Here is a case where because of the above model, the main effect approach of Procedures I and II are not applicable. However, note that in section 3.3 the matrices, $C_{12}^{i.}$, $i=1,...,n_1$, are elements of a diagonal matrix, C_{12} in the r.n.e. for the parameter $\underset{\sim}{p}_{12}$. Therefore a g-inverse for C_{12} may be found from finding a g-inverse for

Table VIII

The Sum, $\frac{1}{\lambda_i}$, For Each Member of Ψ , Case (2)

T_i	T_1	T_2	T_3	T_4
$\Sigma\frac{1}{\lambda_i}$	20.18	22.70	18.94	24.89

Table IX

The Design T_3 of Case (2)

	(a)	(b)
Assembly Number	Assemblies in S_{12}, S_{34}	# Assemblies in S_1, S_2, S_3, S_4
(i)	(0 0)	(0000)
(ii)	(1 1)	(0101)
(iii)	(2 2)	(0202)
(iv)	(3 3)	(1010)
(v)	(4 4)	(1111)
(vi)	(5 5)	(1212)
(vii)	(6 6)	(2020)
(viii)	(7 7)	(2121)
(ix)	(8 8)	(2222)
(x)	(0 1)	(0001)
(xi)	(1 2)	(0102)
(xii)	(2 3)	(0210)
(xiii)	(3 4)	(1011)
(xiv)	(4 5)	(1112)
(xv)	(5 6)	(1220)
(xvi)	(6 7)	(2021)
(xvii)	(7 8)	(2122)

Column (b) follows from (a) using the correspondence: $S_{12} \rightarrow S_1 x S_2$, ; 0 $\rightarrow$(00), 1 $\rightarrow$(01), etc., and $S_{34} \rightarrow S_3 x S_4$; 0 $\rightarrow$(00), 1 $\rightarrow$(01), etc.

each $C_{12}^{i.}$. If $\lambda_j^{i.}$, $j = 1,\ldots,n_2 - 1$, are the non-zero latent roots for each $C_{12}^{i.}$, $i = 1,\ldots,n_1$, then

$$\sum_{i=1}^{n_1} \sum_{j=1}^{n_2-1} \frac{1}{\lambda_j^{i.}}$$

can be used to find a Ψ-optimal design from a given set Ψ.

The constructing of a set Ψ of minimal designs similar to Case (3) may be very difficult. Because of this difficulty, the actual construction of a minimal design, may be in itself an optimal criterion. We shall present a particular example of Case (3) and demonstrate by this example how one could generate a set Ψ.

In order to construct an MMD for Case (3) fifteen assemblies are required, i.e. $h = 3(2) + 2(2)(2) + 1 = 15$. Eleven assemblies can be chosen that lead to four independent contrasts in the p_{12}^{ij} and four independent contrasts in the p_{13}^{ik}. Once these assemblies are constructed, four more assemblies can be constructed such that all constrasts of one set are estimable. Then by Theorem 4.5, the contrasts of the other set are also estimable.

To show this, let (i,j,k) denote the observation and the assembly. For the purposes of illustration, let $i = 1,2,3$, and $j = A,B,C$, and $k = I,II,III$, and let $p_{12}^{ij} = ij$ and $p_{13}^{ik} = ik$. Then 11 assemblies (summarized in Table X) can be constructed such that

E[(1AI) + (2BII) = (1BII) - (2AI)] = [1A + 2B - 1B - 2A] +
[1I + 2II - 1II - 2I]

E[(3AII) + (2CIII) - (3CIII) - 2AII)] = [3A + 2C - 3C - 2A] +
[3II + 2III - 3III - 2II]

E[(2CIII) + (3BI) - 2BI) - (3CIII)] = [2C + 3B - 2B - 3C] +
[2III + 3I - 2I - 3III]

E[(1CIII) + (2BII) - (1BII) - (2CIII)] = [1C + 2B - 2C - 1B] +
[1III +2II - 1II - 2III]

where E[ ] denotes the expected value.

The design in Table X is connected by the addition of (2BI), (1CII), (2CII), and (3CII), since

E[(1AI) + (2BI) - (1BI) - (2AI)] = 1A + 2B - 1B - 2A

E[(3AII) + (2CII) - (3CII) - (2AII)] = 3A + 2C - 3C - 2A

E[(2CII) + (3BII) - (2BII) - (3CII)] = 3B + 2C - 3C - 2B

E[(1CII) + (2BII) - (2CII) - (1BII)] = 1C + 2B - 2C - 1B

The fifteen assemblies are now summarized in Table XI. A set ψ now can be constructed by using different interaction contrasts (other than those given) to be estimated. Then from such a set Ψ, an A-optimal design for S_{12} can be found.

An extension to the model of Case (3) where three interactions are present will be given next.

Case (4) A 3 factorial under the model $E(\underline{y}) = X_{12}\underline{p}_{12} + X_{13}\underline{p}_{13} + X_{23}\underline{p}_{23}$

This case is similar to Case (3) in that the main effect approach of Procedure I and II is not applicable. We shall present a particular

Table X

11 Assemblies of The Required 15 For Case (3)

	A	B	C
1	I	II	III
2	I, II	I, II	III
3	II	I	III

Table XI

One Design of 15 Assemblies For Case (4)

	A	B	C
1	I	I, II	II, III
2	I, II	I, II	II, III
3	II	I	II, III

example of Case (4), and a set of similar designs Ψ may then be constructed as in Case (3), and a Ψ-optimal design for S_{12}, say, can then be identified.

For Case 4, nineteen assemblies are required with all contrasts in S_{12}, S_{13}, and S_{23} being estimable. We shall use the notation of Case (3). To the fifteen assemblies given in Case (3) we add (1AII), (1AIII), 1BIII), and (1CI) so that contrasts in S_{23} are estimable. Let p_{23}^{jk} = jk . Then

$$E[(1AIII) + (1BII) - (1AII) - (1BIII)] = AIII + BII - AII - BIII$$

$$E[(1BI) + (1CII) - (1CI) - (1BII)] = BI + CII - CI - BII$$

$$E[(1AI) + (1CII) - (1CI) - (1AII)] = AI + CII - CI - AII$$

$$E[(1AII) + (1CIII) - (1CII) - (1AIII)] = AII + CIII - CII - AIII$$

These assemblies are summarized in Table XII.

By using different contrasts in S_{12}, S_{13}, and S_{23} to be estimated, one may construct other minimal designs in these sets and compare then on the basis of Ψ-optimality for set S_{12}.

In summary then for models of the type

Case (1) $E(\underline{y}) = X_{12}\underline{p}_{12} + X_{3}\underline{p}_{3}$

Case (2) $E(\underline{y}) = X_{12}\underline{p}_{12} + X_{34}\underline{p}_{34}$

our procedures for constructing a set Ψ may be easily applied. For Case (1) we were able to vary all levels of S_{12} in order to obtain Ψ. In Case (2) we varied some levels of S_{34} and identified the Ψ-opti-

TABLE XII

One Design Of 19 Assemblies For Case (4)

	A	B	C
1	I,II,III	I,II,III	III,II,I
2	I,II	I,II	III,II
3	II	I	III,II

mal design. However, for cases (3) and (4)

$$\text{Case (3):} \quad E(\underset{\sim}{y}) = X_{12}\underset{\sim}{p}_{12} + X_{13}\underset{\sim}{p}_{13}$$

$$\text{Case (4):} \quad E(\underset{\sim}{y}) = X_{12}\underset{\sim}{p}_{12} + X_{13}\underset{\sim}{p}_{13} + X_{23}\underset{\sim}{p}_{23}$$

we were not able to define a set Ψ by varying levels of one set, S_{12}, for example. Ψ must be defined by those contrasts one is able to estimate. For Case (3), one finds a set of assemblies involving two groups of independent interaction contrasts, then "augments" this set to estimate one group. The other group is then estimable. To achieve Ψ for Case (4), again groups of independent contrasts have to be estimated. A member of set Ψ for Case (4) was found by "augmenting" Case (3) for the presence of one more two-factor interaction. Other members of Ψ may be constructed using different contrasts and then an A-optimal member for one set chosen.

CHAPTER VI

PROPOSED EXTENSIONS

Since no previous literature is available that gives specific direction to finding designs requiring a minimum number of assemblies, it is hoped that this work has provided some grounds for future research. In this chapter we shall briefly mention some avenues of investigation.

Our approach to the analysis of incomplete designs was used because it does simplify the work involved. However, a weakness of this approach arises in analysis of designs where one factor interacts with two or more factors. This weakness appears once again in the construction of MAMD's and MMD's (as in Cases (3) and (4), Chapter V). Another problem that appears in the construction of the set from which we determine a Ψ-optimal design. The construction of ψ is within a limited framework, and a particular design T^* is designated optimal with respect to other particular designs. Further, the design T^* is optimal for only one set, S_α .

Due to the extreme limitations (i.e., a minimum number of assemblies) one finds working with the general case, the difficulties just mentioned arise. Our approach was intended to remain as general as possible, applicable to symmetric and asymmetric factorials, However, solutions to the above problems may be found through the investigation of certain classes of MMD's and MAMD's. It may be possible to eliminate these restrictions (but then restrict application) by considering 2^{k_1} factorials, for example.

By using a full rank model, $E(y) = Ub$, the problems of analysis would not appear. The construction of $U'U$ for 2^{k_1} factorials may give some insight into the construction of a set Ψ, if one has to limit oneself to such a set. The latent roots of $U'U$, θ_i could then be related to the sum $\Sigma \frac{1}{\theta_i}$, the $tr(U'U)^{-1}$ and an A-optimal design for all estimable functions could then be identified.

Similar developments for 3^{k_2} factorials may follow. From this work, optimum plans for $2^{k_1} 3^{k_2}$ factorials may result. For these types of special cases, and using the matrix $U'U$ as a guide, all effects will be considered in terms of optimality. Then for very large experiments involving $2^{k_1} 3^{k_2}$ factorials, plans could result requiring the very minimum number of necessary assemblies. For $k_1 =$ 30 and $k_2 = 10$ (i.e. 40 factors), only 51 assemblies would be necessary for a main effect MMD.

The next problem one may consider is the relationship of the composition of C_α, the c.m. in the r.n.e. for $\underline{p}_\alpha$, to its latent roots. One approach is to investigate $tr(C_\alpha) = \sum_{i=1}^{n_\alpha - 1} \lambda_i$.

A design with maximum $tr(C_\alpha)$ permits maximum power on tests of p_α. With

$$C_\alpha = H_{\alpha\alpha} - (H_{\alpha 1}, \ldots, H_{\alpha m}) H^- (H_{\alpha 1}, \ldots, H_{\alpha m})'$$

and with $tr(H_{\alpha\alpha}) = h$, the number of assemblies in T, $tr(C_\alpha)$ depends upon the diagonal elements of $[(H_{\alpha 1}, \ldots, H_{\alpha m}) H^- (H_{\alpha 1}, \ldots, H_{\alpha m})'$.

These diagonal elements are functions of H^- which indirectly by Theorem 2.1 can be expressed as functions of the diagonal matrices of H. A further investigation may reveal how the level frequencies of sets S_β, $\beta \neq \alpha$, effect $tr(C_\alpha)$.

CHAPTER VII

SUMMARY

In this work we considered the design and analyses of special types of m-factor designs, MMD's and MAMD's, that use a minimum number of experimental units. For cases where due to the number of required assemblies or due to the construction of these assemblies, available fractional factorials do not apply, MMD's or MAMD's may be of use. This is particularly true if there is an extreme limitation on experimental units and many factors, and thus many effects, are to be examined. However because of the nature of these designs, a previous estimate of error may be required. For this case, Cochran and Cox (1957) comment:

> If the use of an estimate of error from previous experiments is contemplated, some preliminary questions should be answered. Were the previous experiments conducted on similar experimental material, under similar conditions of experimentation and with treatment effects of about the same order of magnitude? Were the previous error variances homogeneous from one experiment to another, as indicated by a test of homogeneity of variances? If these questions are answered in the affirmative, the risk in taking a external estimate of error is decreased.

Also, another source for an estimate of error may be other experimental units not required by the design but available to the experimenter.

Due to the limitation of factor level combinations, the analysis of these designs requires some discussion. For this analysis we used a reduced form of the normal equations and thereby generalized the results of Potthoff (1958). This analysis is applicable to all incomplete designs where two-factor interactions may or may not be assumed.

Once a form of analysis for MMD's and MAMD's is established, we proceeded to investigate methods for constructing them. An investigation into the concept of connected MD's where no interaction is assumed (Srivastava and Anderson, 1970) led us to the development of a similar concept for MD's where first order interaction is assumed. Using this latter concept we proceeded to develop a method of adding the minimum of necessary assemblies to a given design in order to estimate effects previously not estimable. In this way we developed a procedure for constructing MAMD's. This augmentation procedure enables one, for example, to construct first order interaction designs from main effect plans.

Given such a procedure, the question then arose on how to construct optimal MAMD's. Because of the limitation on the number of assemblies to be added to a design, optimality for MAMD's is difficult to define. However, using our approach to the analysis of incomplete designs we were able to present one method of finding preferred MAMD's.

This method involved the construction of a set of "optimum" MAMD's, Ψ. This set may be limited to a few MAMD's by the experimental situation or may be an almost infinite set. If no balanced or orthogonal MAMD's are available, we limit our investigation to a special class of MAMD's. Using this class we are then able to construct some designs for a particular set, Ψ, from which we identify one design which is A - optimal for one factor.

Design optimality is further discussed for the construction of MMD's which require the total number of assemblies used to be a minimum. Again, because of the constraint on the number used, optimality for MMD's is difficult to define. However, one approach to finding preferred MMD's similar to that of finding preferred MAMD's can be given. A set Ψ of MMD's can be generated using the procedures developed, and an A-optimal design chosen.

Using this approach to constructing main effect MMD's, examples of 3^n factorials with first order interaction were given. One of these examples is both a MMD and a MAMD, and the analysis of data generated for this example is given in Appendix I. The optimality of this particular design is discussed in Appendix II and its efficiency is compared to a more common fraction of a 3^3 factorial.

Discussions on optimality and efficiency bring us to part of the philosophy upon which this dissertation is based. An optimal design or efficient design is usually a preferred design according to a criterion that is similar to the criterion a consumer uses in buying a product. This criterion may be expressed accordingly, "Am I getting the most return for my investment?" Let us say that the cost of a product that a consumer is buying varies with size or quantity (the analogy here is with experimental units). As in the case of the more common fraction of the 3^3 factorial given in Appendix II, there may be a better investment or "bargain" in "buying" a larger size or quantity. However, a greater quantity requires a greater cost, and if this expense cannot be met by the consumer, the true bargain may be in what he can afford.

POSTSCRIPT

As mentioned in Chapter I, many experimental plans do not apply when the number of experimental units available to the experimenter are extremely limited or when certain factor level combinations are not available. For these cases we draw upon a well known expression dealing with affairs of the heart, and comment that it is better to have experimented once than not to have experimented at all.

BIBLIOGRAPHY

1. Addelman, S. (1962a)."Orthogonal Main-Effect Plans For Asymmetrical Factorial Experiments." Technometrics 4, 11-46.

2. Addelman, S. (1962b)."Symmetrical and Asymmetrical Fractional Factorial Plans." Technometrics 4, 47-58.

3. Addelman, S. (1963). "Techniques For Constructing Fractional Replicate Plans." Journal American Statistical Association 58, 45-71.

4. Addelman, S. and Kempthorne, O. (1961). Orthogonal Main Effect Plans. Aeronautical Research Laboratory, Office of Aerospace Research, United States Air Force, Wright-Patterson Air Force Base, Ohio.

5. Anderson, R. (1968). "On the Construction and Analysis of Multi-Factor Partially Balanced Designs." University of Nebraska, Department of Statistics, unpublished Ph.D. dissertation.

6. Banerjee, K. S. and Federer, W. T. (1966). "On Estimation and Construction in Fractional Replication." Annals Mathematical Statistics 37, 1033-1039.

7. Bose, R. C. (1939). "On the Construction of Balanced Incomplete Block Designs." Annals Eugenics 9, 353-399.

8. Bose, R. C. (1958). "Mimeographed Class Notes on Analysis of Variance, Linear Estimation, and Testing of Linear Hypothesis." University of North Carolina, Department of Statistics, unpublished.

9. Bose, R. C. and Srivastava, J. N. (1964a). "Analysis of Irregular Factorial Fractions." Sankhyã,(Series A) 26,(parts 2 and 3) 117-144.

10. Bose, R. C. and Srivastava, N. N. (1964b). "Multidimensional Partially Balanced Designs and Their Analysis, With Applications To Partially Balanced Factorial Fractions." Sankhyã,(Series A) 26, (parts 2 and 3) 145-168.

11. Box, G. E. P. and Hunter, J. S. (1957). "Multifactor Experimental Designs." Annals Mathematical Statistics 28, 195-241.

12. Causey, B. D. (1967). Some Multidimensional Incomplete Block Designs. University of Chicago Library, Department of Photoduplication, Ph.D. dissertation, unpublished.

13. Chernoff, H. (1953). "Locally Optimum Designs For Estimating Parameters." Annals Mathematical Statistics 24, 586-602.

14. Cochran, W. G. and Cox, G. M. (1957). Experimental Designs (2nd ed.), John Wiley and Sons, New York.

15. Cornish, R. (1938). "Factorial Treatments in Incomplete Randomized Blocks." Journal Australian Institute Agricultural Science 4, 199-203.

16. Daniel, C. (1962). "Sequences of Fractional Replicates in the 2^{p-q} Series." Journal American Statistical Associaton 57, 403-429.

17. Daniel, C. (1971). "One-At-A-Time Plans." The Fisher Memorial Lecture, American Statistical Association Meetings, Colorado State University, August 25, 1971.

18. Ehrenfeld, S. (1955). "On the Efficiency of Experimental Designs," Annals Mathematical Statistics 26, 247-255.

20. Elston, R. C. and Bush, N. (1964). "The Hypothesis That Can Be Tested When There Are Interactions in an ANOVA Model." Biometrics 20, 681-699.

21. Federer, W. T. (1961). "Augmented Designs With One-Way Elimination of Heterogeneity." Biometrics 17, 447-473.

22. Federer, W. T., Hedayat, A., and Raktoe, B. L. (1972). "Econostat, Optimal Fractional Factorial Designs: A New Approach." Institute Mathematical Statistics Meetings, Iowa State University, April 28, 1972.

23. Finney, D. J. (1945). "The Fractional Replication of Factorial Experiments." Annals Eugenics 12, 291-301.

24. Fisher, R. A. (1935). The Design and Analysis of Experiments. Oliver and Boyd LTD, London.

25. Freedman, G. H. (1964). "The Addition of Further Treatments to Latin Square Designs." Biometrics 20, 713-729.

26. Hinkelmann, K. (1968a). "Design of Experiments (Lecture Notes)." Virginia Polytechnic Institute & S.U., Department of Statistics, unpublished.

27. Hinkelmann, K. (1968b). "Analysis of Variance (Lecture Notes)." Virginia Polytechnic Institute & S.U., Department of Statistics, unpublished.

28. Hoke, A. T. (1971). "Economic Second Order Designs in Sequences of Partially Balanced Fractions of the 3^n Factorial." American Statistical Association Meetings, Colorado State University, August 24, 1971.

29. Kiefer, J. (1958). "On the Nonrandomized Optimality and Randomized Nonoptimality of Symmetrical Designs." Annals Mathematical Statistics 29, 675-699.

30. Kiefer, J. (1959). "Optimum Experimental Designs." Journal Royal Statistical Society, (ser. B) 21, 272-379.

31. Kiefer, J. (1961). "Optimum Designs In Regression Problems, II." Annals Mathematical Statistics 32, 298-325.

32. Kempthorne, O. (1952). The Design and Analysis of Experiments. John Wiley and Sons, New York.

33. Kempthorne, O. (1956). "The Efficiency Factor of an Incomplete Block Design." Annals Mathematical Statistics 27, 846-851.

34. Kramer, C. Y. and Bradley, R. A. (1957a). "Intra-Block Analysis For Factorials in Two-Associate Class Group Divisible Designs." Annals Mathematical Statistics 28, 349-361.

35. Kramer, C. Y. and Bradley, R. A. (1957b). "Examples of Intra-Block Analysis For Factorials in Group Divisible, Partially Balanced, Incomplete Blocks." Biometrics 13, 197-224.

36. Margolin, B. H. (1969). "Orthogonal Main-Effect Plans Permitting the Estimation of All Two-Factor Interactions For the $2^n 3^m$ Factorial Series of Designs." Technometrics 11, 747-761.

37. Plackett, R. L. and Burman, J. R. (1946). "The Design of Optimum Multifactorial Experiments." Biometrika 33, 305-325.

38. Potthoff, R. F. (1958). "Multi-Dimensional Incomplete Block Designs." Institute of Statistics Mimeograph Series No. 211, University of North Carolina, Chapel Hill.

39. Potthoff, R. F. (1962a). "Three-Factor Additive Designs More General Than the Latin Square." Technometrics 4, 187-208.

40. Potthoff, R. F. (1962b). "Four-Factor Additive Designs More General Than the Graeco-Latin Square." Technometrics 4, 361-366.

41. Potthoff, R. F. (1963). "Three Dimensional Incomplete Block Designs For Interaction Models." Biometrics 19, 229-263.

42. Raktoe, B. L. and Federer, W. T. (1968). "A Unified Approach For Constructing a Useful Class of Non-Orthogonal Main Effect Plans in k^n Factorial " Journal Royal Statistical Society, Series B, 30, 371-380.

43 Rao, C. R. (1967). "Calculus of Generalized Inverses of Matrices; Part 1, General Theory." Sankhya, Ser. A, 29, 319-342.

44. Rhodes, C. A. (1964). "Contributions to the Theory, Computation, and Application of Generalized Inverses." University of North Carolina at Raleigh, Department of Statistics, unpublished Ph.D. dissertation.

45. Shah, B. V. (1959). "A Generalization of Partially Balanced Incomplete Block Designs," Annals Mathematical Statistics 30, 1041-1050.

46. Srivastava, J. N. and Anderson R. (1970). "Some Basic Properties of Multidimensional Partially Balanced Designs." Annals Mathematical Statistics 41, 1430-1445.

47. Tukey, J. W. (1949). "One Degree of Freedom For Non-Additivity." Biometrics 1, 232-242.

48. Tukey, J. W. (1955). "Answer to Query on One-Degree of Freedom Test," , Biometrics 11, 372-375.

49. Vajda, S. (1967). "The Mathematics of Experimental Design." Griffin's Statistical Monographs and Courses, no. 23, Hafner, New York.

50. Yates, F. (1936). "Incomplete Randomized Blocks." Annals Eugenics 7, 121-140.

51. Youden, W. J. (1937). "Use of Incomplete Block Replication in Estimating Tobacco Mosaic Virus." Contributions Boyce Thompson Institute 9, 41-48.

52. Youden, W. J. and Hunter, J. S. (1955). "Partially Replicated Latin Squares." Biometrics 11, 278-89.

APPENDIX I

A 1/3 REPLICATE OF A 3^3 FACTORIAL AUGMENTED FOR AN ASSUMED TWO-FACTOR INTERACTION: THE ANALYSIS OF DATA

Presented in Appendix I will be an application of the analysis of Model (2) given in section 3.2. For the preliminary test on fruit juices introduced in section 1.1, we shall assume a 1/3 replicate of a 3^3 factorial is available for a test design. However, the fruit juice salesmen have indicated that a price-juice interaction may be present. In order to allow for this interaction, the nine assemblies of the 1/3 replicate given by (1AI), (2BI), (3CI), (1CII), (2AII), (3BII), (1BIII), (2CIII), and (3AIII), have been augmented by two assemblies, (1BII) and (1BI). For an estimate of error, it is assumed that these eleven assemblies are to be replicated. Thus twenty-two homogeneous test periods are to be used, and it is assumed that the error involved in the application of these assemblies is normally distributed. An observation on a test period will represent the total number of 3-oz. juice cans sold during that period. The model used for this design will be

$$E(y_{ijk\omega}) = \mu + p_1^i + p_2^j + p_3^k + t_{12}^{ij} = p_{12}^{ij} + p_3^{k*}$$

where $p_{12}^{ij} = p_1^i + p_2^j + t_{12}^{ij}$ and $p_3^{k*} = p_3^k + \mu$. (For the r.n.e. for p_3^k, μ is associated with p_{12}^{ij}, i.e. $p_{12}^{ij} = p_{12}^{ij*} + \mu$)

We shall assume, for the purpose of illustration, that the experiment has been performed. The data generated for this experiment shall come from values given for the parameters and error terms. We shall assume μ = 500

and for the juices,

$$p_1^A = 30 \qquad p_1^B = 50 \qquad p_1^C = -80$$

for the prices,

$$p_2^1 = 50 \qquad p_2^2 = -30 \qquad p_2^3 = -20$$

for the days,

$$p_3^{I} = 40 \qquad p_3^{II} = -20 \qquad p_3^{III} = -20$$

and for the price-juice interaction,

t_{12}^{ij}

	j =	1	2	3
	A	-20	20	0
i =	B	-10	-40	50
	C	30	20	-50

For the data, error terms have been generated from a normal distribution with mean 0 and variance 9. Since the data is to be in integer form, these error terms are given in the form of integers (decimals have been rounded to whole numbers.) The error terms are then for the first replication, respectively

(0,4,5,3,-2,-1,-3,2,-1,1,2),

and for the second replication, respectively

(-4,-2,0,5,-3,1,-4,-1,2,0,1).

With these parameters and error terms we have the following data in Table XIII:

TABLE XIII

Sales Data* From Preliminary Test On Frozen Fruit Juices

Price-Juice-Day Level combination	Total Sales Observation 1	Observation 2	Total
(1AI)	600	596	1196
(2BI)	524	518	1042
(3CI)	395	390	785
(1CII)	483	485	968
(2AII)	498	497	995
(3BII)	559	561	1120
(1BIII)	567	566	1133
(2CIII)	392	389	781
(3AIII)	489	492	981
(1BII	571	570	1141
(1BI)	632	631	1263

* Data in units of 3-oz, cans

Source: Data generated from given parameters

Following the methods of section 3.2, the r.n.e. for $\underset{\sim}{p}_{12}$ are found, $C_{12}\,\hat{\underset{\sim}{p}}_{12} = \underset{\sim}{Q}_{12}$, where C_{12} and $\underset{\sim}{Q}_{12}$ are given by

$$C_{12} = \begin{bmatrix} 1.50 & -0.50 & 0.00 & 0.00 & -0.50 & 0.00 & 0.00 & 0.00 & -0.50 \\ -0.50 & 4.33 & -0.50 & -0.50 & -0.50 & -0.67 & -0.67 & -0.50 & -0.50 \\ 0.00 & -0.50 & 1.50 & 0.50 & 0.00 & 0.00 & 0.00 & -0.50 & 0.00 \\ 0.00 & -0.50 & -0.50 & 1.50 & 0.00 & 0.00 & 0.00 & -0.50 & 0.00 \\ -0.50 & -0.50 & 0.00 & 0.00 & 1.50 & 0.00 & 0.00 & 0.00 & -0.50 \\ 0.00 & -0.67 & 0.00 & 0.00 & 0.00 & 1.33 & -0.67 & 0.00 & 0.00 \\ 0.00 & -0.67 & 0.00 & 0.00 & 0.00 & -0.67 & 1.33 & 0.00 & 0.00 \\ 0.00 & -0.50 & -0.50 & -0.50 & 0.00 & 0.00 & 0.00 & 1.50 & 0.00 \\ -0.50 & -0.50 & 0.00 & 0.00 & -0.50 & 0.00 & 0.00 & 0.00 & 1.50 \end{bmatrix}$$

$\underset{\sim}{Q}'_{12}$ = (124.5, 444.5, -88.0, -61.0, -29.5, -184.0, 16.0, 64.0, -286.5)

A g-inverse, C^{-}_{12} is then calculated and $\hat{p}_{12} = C^{-}_{12} Q_{12}$ follows:

$$C^{-}_{12} = \begin{bmatrix} 0.753 & -0.025 & -0.247 & -0.247 & 0.253 & -0.191 & -0.191 & -0.247 & 0.253 \\ -0.025 & 0.198 & -0.025 & -0.025 & -0.025 & 0.031 & 0.031 & -0.025 & -0.025 \\ -0.247 & -0.025 & 0.753 & 0.253 & -0.247 & -0.191 & -0.191 & 0.253 & -0.247 \\ -0.247 & -0.025 & 0.253 & 0.753 & -0.247 & -0.191 & -0.191 & 0.253 & -0.247 \\ 0.253 & -0.025 & -0.247 & -0.247 & 0.753 & -0.191 & -0.191 & -0.247 & 0.253 \\ -0.191 & 0.031 & -0.191 & -0.191 & -0.191 & 0.864 & 0.364 & -0.191 & -0.191 \\ -0.191 & 0.031 & -0.191 & -0.191 & -0.191 & 0.364 & 0.864 & -0.191 & -0.191 \\ -0.247 & -0.025 & 0.253 & 0.253 & -0.247 & -0.191 & -0.191 & 0.753 & -0.247 \\ 0.253 & -0.025 & -0.247 & -0.247 & 0.253 & -0.191 & -0.191 & -0.247 & 0.753 \end{bmatrix}$$

For notation purposes, let $\hat{p}^{1A}_{12} = 1A$. Then

$\hat{p}_{12} = (1A, 1B, 1C, 2A, 2B, 2C, 3A, 3B, 3C)'$

The values of $\hat{p}^{ij}_{12}$ are given below:

$\hat{p}^{ij}_{12}$

	j =	1	2	3
	A	55.9	16.4	13.5
i=	B	89.4	-21.0	79.0
	C	2.9	-86.5	-149.6

Now, in order to define contrasts in main effects and interactions, we refer to the "f" parameters defined in (3.1.3) and (3.1.5)

$$f_{i.} = p^{i.}_{12} - p^{..}_{12}$$

$$f_{.j} = p^{.j}_{12} - p^{..}_{12}$$

and

$$f_{ij} = p^{ij}_{12} - p^{..}_{12} - p^{.j}_{12} + p^{..}_{12}$$

If we applied (3.1.4) and $p^{..} = 0$, we have

$$\hat{f}_{i.} = \hat{p}^{i.}_{12} = \hat{p}^{i}_{1} = \frac{1}{3} \sum_{j} \hat{p}^{ij}_{12}$$

$$\hat{f}_{.j} = \hat{p}^{.j}_{12} = \hat{p}^{j}_{2} = \frac{1}{3} \sum_{i} \hat{p}^{ij}_{12}$$

and

$$\hat{f}_{ij} = \hat{p}^{ij}_{12} - \hat{p}^{i.}_{12} - \hat{p}^{.j}_{12} = \hat{t}^{ij}_{12}$$

Note that $t^{1A}_{12} + t^{2B}_{12} - t^{1B}_{12} - t^{2A}_{12}$ is estimated by $\hat{p}^{1A}_{12} + \hat{p}^{2B}_{12} - \hat{p}^{1B}_{12} - \hat{p}^{2A}_{12}$.

Thus all contrasts on main effects and interactions can be estimated by contrasts on $\hat{p}_{12}$. For the following contrasts, $\ell_i = L_i' \hat{p}_{12}$, the corresponding variance of ℓ_i, $V(\ell_i)$ is given by

$$\underline{L}_i' \, C_{12}^{-} \, \underline{L}_i \, \sigma^2$$

			$V(\ell_i)$
For prices			
	$\ell_1 = \hat{p}_1^1 - \hat{p}_1^3$	68.5	$0.33 \, \sigma^2$
	$\ell_3 = \hat{p}_1^1 - 2\hat{p}_1^2 + \hat{p}_1^3$	91.2	$1.00 \, \sigma^2$
For juices			
	$\ell_3 = \hat{p}_2^A - \hat{p}_2^C =$	106.3	$0.22 \, \sigma^2$
	$\ell_4 = \hat{p}_2^A - 2\hat{p}_2^B + \hat{p}_2^C =$	-147.3	$0.89 \, \sigma^2$

For price-juice interaction, using our adopted notation,

$\ell_5 = 1A + 2B - 1B - 2A =$	-71.0	$4.0 \, \sigma^2$
$\ell_6 = 3A + 2C - 2A - 3C =$	60.0	$5.0 \, \sigma^2$
$\ell_7 = 2C + 3B - 2B - 3C =$	163.0	$5.0 \, \sigma^2$
$\ell_8 = 1C + 2B - 1B - 2C =$	-21.0	$3.0 \, \sigma^2$

The s.s. due to each contrast is obtained by

$$(\underline{L}_i' \hat{\underline{p}}_{12})' \left[\underline{L}_i' \, C_{12}^{-} \, \underline{L}_i \right]^{-1} (\underline{L}_i' \hat{\underline{p}}_{12})$$

and the s.s. due to main effects and interactions found in the following ANOVA:

Source	df		s.s
Due to p_{12}^{ij}	8	$\hat{p}'_{12}\ Q_{12}$	110,123.50
Due to p_{1}^{j}	2	Prices	22,387.71
Due to p_{2}^{j}	2	Juices	55,630.10
Due to t_{12}^{ij}	4	Interaction	6,697.67
Due to p_{3}^{k*}	3	By Subtraction	5,923,334.00 †
Due to error	11	$\sum_{\omega} (y_{ijk\omega} - y_{ijk.})^2$	53.50
Total	22	$\underset{\sim}{y}'\ \underset{\sim}{y}$	6,033,511.00

† uncorrected s.s. (referred to, as "unadjusted s.s.")

The unadjusted s.s. here can be received by $Y_3^{*\prime}\ H_{33}^{-1}\ Y_3^{*}$ or, as given, by subtraction. As one would expect from the parameters used to generate the data, all main effects and interactions are significant at the 0.01 level.

For tests on the differences in days (I, II, III), we find the r.n.e. for p_3, $C_3\ \grave{\underset{\sim}{p}}_3 = \underset{\sim}{Q}_3$ and calculate a g-inverse, C_3^- to find $\grave{p}_3$:

$$C_3^- = \begin{bmatrix} 1.333 & -0.667 & -0.667 \\ -0.667 & 1.333 & -0.667 \\ -0.667 & -0.667 & 1.333 \end{bmatrix}$$

$$Q_3 = (84.0 \quad -38.0 \quad -46.0)$$

$$C_3^- = \begin{bmatrix} 0.50 & 0.0 & 0.0 \\ 0.0 & 0.50 & 0.0 \\ 0.0 & 0.0 & 0.50 \end{bmatrix}$$

$$\hat{p}_3 = (\hat{p}_3^{I}, \quad \hat{p}_3^{II}, \quad \hat{p}_3^{III}) =$$

$$(42.0, \quad -19.0, \quad -23.0)$$

Then for contrasts in p_3^k (on the main effects of days) we have

	ℓ_i		$V(\ell_i)$
ℓ_9	$= \hat{p}_3^{I} - \hat{p}_3^{II}$	$= 65.0$	$1.0\ \sigma^2$
ℓ_{10}	$= \hat{p}_3^{I} - 2\hat{p}_3^{II} + \hat{p}_3^{III}$	$= 57.0$	$3.0\ \sigma^2$

Again, the s.s. due to these contrasts is found and summarized:

Source	df	s.s.
Days (due to p_3^k)	2	5,308.15
Due to p_{12}^{ij*}	9	6,028,149.35†
Error	11	53.50
Total	22	6,033,511.00

† uncorrected s.s. (referred to as "unadjusted s.s.)

The unadjusted s.s. due to p_{12}^{ij*} can be found in this case either by

$Y^{*'}_{(12)} H^{-1}_{(12)(12)} Y^{*}_{(12)}$ or, as given by subtraction. As expected, due to the input parameters, the main effects of days are significant at the 0.01 level.

"The presentation of results depends on their nature and the audience for whom they are intended," (Cochran and Cox, 1950) and since the purpose here is to illustrate the analysis of section 3.2, we could terminate our presentation here. However, more information on the main effects may be requested. For example, the mean responses of price 1 may be desired. To obtain an estimate of this mean response, $\hat{p}^{1}_{1} + \hat{\mu}$, a value for must be found. We may obtain this value from our work; however, as mentioned in section 3.4, "a weakness" of our approach is that in finding a g-inverse for C_{12}, we find a variance-covariance matrix for contrasts involving p^{ij}_{12} only (and not μ, for example). The variance of $\hat{p}^{1}_{1} + \hat{\mu}$ would require a variance-covariance matrix for all parameters.

A value for $\hat{\mu}$ is obtained form the adjusted n.e.,

$$X'_3 X_3 \hat{p}^*_3 = X'_3 y - X'_3 X_{(12)} \hat{p}_{12},$$

where

$$X_3'X_3 = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 6 \end{bmatrix} \qquad X_3'y = \begin{bmatrix} 4 & 2 & 8 & 6 \\ 4 & 2 & 2 & 4 \\ 2 & 8 & 9 & 5 \end{bmatrix}$$

$$X_3'X_{(12)}p_{12} = \begin{bmatrix} 50.4 \\ -375.6 \\ -32.7 \end{bmatrix}$$

$$\text{and } X_3'y - X_3'X_{(12)}p_{12} = \begin{bmatrix} 4336.4 \\ 3848.4 \\ 2862.3 \end{bmatrix}$$

From this equation we obtain $\hat{\underline{p}}_3^*$, and a value for $\hat{\mu}$ is $\mu = \hat{p}_3^{*I} - p_3^I$. From the above we have $\hat{p}_3^{*I} = 542.05$, and from previous work we have $\hat{p}_3^I = 42.0$. Therefore $\hat{\mu} = 500.05$. Then

for prices

$$\ell_{11} = \hat{p}_1^1 + \hat{\mu} = 549.49$$

$$\ell_{12} = \hat{p}_1^2 + \hat{\mu} = 469.67$$

$$\ell_{13} = \hat{p}_1^3 + \hat{\mu} = 481.00$$

for juices

$$\ell_{14} = \hat{p}_2^A + \hat{\mu} = 528.66$$

$$\ell_{15} = \hat{p}_2^B + \hat{\mu} = 549.16$$

$$\ell_{16} = \hat{p}_2^C + \hat{\mu} = 422.33$$

for days

$$\ell_{17} \quad \hat{p}_3^I + \hat{\mu} = 542.05$$

$$\ell_{18} \quad \hat{p}_3^{II} + \hat{\mu} = 481.05$$

$$\ell_{19} \quad p^{III} + \hat{\mu} = 477.05$$

APPENDIX II

A 1/3 REPLICATE OF A 3 FACTORIAL AUGMENTED FOR AN ASSUMED TWO-FACTOR INTERACTION: OPTIMALITY

For Case 1, Section 5.9, and Appendix I, the design mentioned shall now be investigated for its optimality. The assemblies of a 1/3 replicate can be given as the following Latin square, with days (I,II,III) juices (A,B,C) and prices (1,2,3) in nine assemblies:

A 1/3 Replicate of a 3^3 Factorial in 9 Assemblies: T_{ψ_a}

	1	2	3
I	A	B	C
II	C	A	B
III	B	C	A

If we let the factors day (D), juices (J) and prices (P) be represented symbolically, the given design has the following alias structure obtained from the defining contrast $I = DJP^2$:

$$D = D^2J\,P^2 = JP^2$$

$$J = D\,J^2P^2 = DP^2$$

$$P = DJ = DJP$$

$$DP = DJ^2 = JP$$

In order to estimate main effects, the assumption of no interaction is made. However, given a two-factor interaction between J and P, the main effects are no longer estimable and the set of days S_3, and the juice-price level combinations, S_{12}, are not connected in T_{ψ_a}. Since there are 2 d.f. to be associated with set S_3 and 8 d.f. to be associated with S_{12}, and, there are only 8 d.f. available for estimating main effects and interactions in T_{Ψ_a}, at least 2 or more assemblies are needed to connect S_{12} (and thus S_3). We shall assume that interaction contrasts are of particular interest. Thus we are confronted with constructing a MAMD that is optimal for S_{12}. We shall determine a set Ψ and then choose a design T* that is Ψ-optimal for S_{12}.

The set Ψ will be determined by MAMD's, $T_i = T_{\Psi_a} + D_i$, and thus the sequences D_i. The sequences D_i, shall be constructed according to (5.5.2) for set S_3, such that in T_i

$$h_{..\ell} = 3, \quad \ell = III$$
$$= 4, \quad \ell = I, II$$

Here $h(T_i) = 11 = 3 \cdot 3 + 2$ or $h(T_i) = 2 \pmod 3$ and $k_3 = 3$, $a_3 = 2$. In order to augment T_{Ψ_a} with 2 assemblies only two different types of D_i are possible: one where a level combination of S_{12} is repeated and one where such a combination is not repeated. For example, let D_1 = [(1BI), (1BII)] and D_2 = [(2cII), (3AIII)] be the two possible sequences used to augment T_{Ψ_a}. Let y_{1AI} be an observation for assembly (1AI). Then for T_1.

$$E\,[y_{1BI} - y_{1BIII}] = p_3^{I} - p_3^{III}$$

$$E\,[y_{1BII} - y_{1BIII}] = p_3^{II} - p_3^{III}$$

and for T_2,

$$E\,[y_{3AI} - y_{3AIII}] = p_3^{I} - p_3^{III}$$
$$E\,[y_{2CII} - y_{2CIII}] = p_3^{II} - p_3^{III}$$

Therefore the sets S_{12} and S_3 are connected in T_1 and T_2. A c.m., C_{12}, is calculated for T_1 and for T_2 and the corresponding sum, $\Sigma\frac{1}{\lambda_i}$, is 12.67

for T_1 and 14.67 for T_2. Therefore $T_1^* = T_{\psi_a} + D_1$ is Ψ-optimal for S_{12}:

The Augmented 1/3 Replicate
in 11 Assemblies: T_1^*

	1	2	3
I	A,B	B	C
II	C,B	A	B
III	B	C	A

Thus, in augmenting any 1/3 replicate of a 3^3 for a set S_{12}, repeating one level combination of S_{12}, will produce a Ψ-optimal MAMD (and MMD) for S_{12}.

Having found a design that is Ψ-optimal for S_{12}, we shall now investigate the variances of contrasts of the price-juice interaction. We shall compare T_1^*, a 11/27 replicate of a 3^3 factorial, with a 2/3 replicate to determine an efficiency factor for T_1^*.

A 2/3 Replicate of a 3^3 Factorial
in 18 Assemblies: $T_{2/3}$

	1	2	3
I	A,C	B,C	A,B
II	B,C	A,B	A,C
III	A,B	A,C	B,C

Let us consider the nine different interaction contrasts involving four level combinations of J and P. (These 9 are all such possible contrasts.) For purpose of presentation let $\hat{p}_{12}^{1\Lambda} = {}^1\Lambda$. Then the nine estimates, ℓ_i, and their variances $V(\ell_i)$ in T_1^*, given one observation per assembly, are:

i	ℓ_i	$V(\ell i)$
1	1A + 2B - 1B - 2A	$8.0\ \sigma^2$
2	1A + 3B - 1B - 3A	$6.0\ \sigma^2$
3	1A + 2C - 1C - 2A	$10.0\ \sigma^2$
4	1A + 3C - 1C - 3A	$10.0\ \sigma^2$
5	1B + 2C - 1C - 2B	$6.0\ \sigma^2$
6	1B + 3C - 1C - 3B	$8.0\ \sigma^2$
7	2A + 3B - 2B - 3A	$10.0\ \sigma^2$
8	2A + 3C - 2C - 3A	$10.0\ \sigma^2$
9	2B + 3C - 3B - 2C	$10.0\ \sigma^2$
	Average of $V(\ell_i)$ =	$8.67\sigma^2$

The variance of ℓ_i for the 2/3 replicate is $2.33\ \sigma^2$.

Then for an efficiency factor we have

$$E.F._{(1)} = \frac{\text{Average variance of 9 contrasts in } T_{2/3}}{\text{Average variance of 9 contrasts in } T_1^*} \cdot \frac{18}{11}$$

$$= \frac{2.33}{8.67}\ \frac{\sigma^2}{\sigma^2} \cdot \frac{18}{11} = 0.440$$

Another efficiency factor can be given (Hoke, 1971) as

$$E.F._{(2)} = \frac{\text{tr } (\ C_{12}^{-}) \text{ for } T_{2/3}}{\text{tr } (\ C_{12}^{-}) \text{ for } T_1^*} \cdot \frac{18}{11} = \frac{4.41}{12.67}\ \frac{18}{11}$$

$$= 0.570$$

In calculating $E.\ F._{(2)}$ one compares the average variance of $\hat{p}^{u_\alpha u_\beta}_{12} - \hat{p}^{v_\alpha v_\beta}_{12}$, $(u_\alpha u_\beta) \neq (v_\alpha v_\beta)$ for T_1^* with respect to $T_{2/3}$.

(For tr (C_{12}^{-}), the sum $\Sigma \frac{1}{\lambda_i}$, λ_i a root of C_{12}, is used.)

From $E.F._{(1)}$ and $E.F._{(2)}$ it may be argued that the efficiency of

T_1^* is somewhat low, but there are other desirable properties of T_1^* to consider. First, only 11 different assemblies are required for T_1^* as opposed to 18 for $T_{2/3}$.

Given 17 experimental units, $T_{2/3}$ could not be performed, but T_1^* could be, with 6 d.f. available for error. Further, the 17 - 11 = 6 assemblies used for an estimate of error do not have to be different,i.e. one assembly could be replicated 7 times.

Next, given the 11 assemblies, one could provide a test for interaction using the method of Tukey (1949, 1955) where the 4 d.f. for interaction are separated into 3 d.f. for interaction and 1 d.f. for error. Should a test on interaction prove significant, one would then have a design to estimate the interaction contrasts.

Finally, by recalling the C_3^- matrix given in Appendix I, it is seen that T_1^* remains orthogonal for days (I,II,III) and that the variance of $\hat{p}_3$ contrasts are essentially the same for T_1^* and $T_{2/3}$. (The E.F.$_{(1)}$ for contrasts in p_3 is 1.)

The vita has been removed from
the scanned document

MINIMAL MULTIDIMENSIONAL DESIGNS

By

John T. Sennetti

(ABSTRACT)

Given is an investigation of experimental plans that require a minimum number of assemblies, minimal multidimensional designs (MMD's) and minimal augmented multidimensional designs (MAMD's). Such designs are incomplete factorials which allow for only some two-factor interactions. An analysis of incomplete factorials with and without interactions is developed from a reduced form of the normal equations.

$$C_i p_i = Q_i$$

where p_i is a vector of estimates of factor i effects and C_i is the coefficient matrix for p_i and Q_i is a vector of transformed observations. The general forms for C_i and Q_i are presented.

The construction of MMD's and MAMD's is made possible from results obtained on connected designs. A definition of a connected design where two-factor interactions are assumed leads to a procedure for "connecting" experimental plans. This procedure provides a way of adding assemblies to a design in order to estimate contrasts not originally estimable in the design. Using this augmentation procedure and the minimum number of assemblies to be added, MAMD's may be constructed. MMD's follow by sequentially augmenting with the minimum number of assemblies, m-factor designs, $m = 1,2,\ldots,m^*$, whose total number of factor level combinations are a minimum. A method for finding MMD's and MAMD's which are optimal for one factor or for a set of two-factors

is then presented as well as some examples of MMD's and MAMD's with and without two-factor interactions. Data is generated and analyzed for a particular design which is both a MAMD and a MMD, and a discussion of this design's optimality is also given.