Three-Factor Additive Designs More General Than the Latin Square

Richard F. Potthoff

University of North Carolina

A class of designs are described for the case of three factors at levels m, n and p respectively for which the total number of experiments required is a fraction of mnp.

INTRODUCTION

Suppose we have an experiment with m varieties of corn, n types of fertilizer, and p insecticides. Suppose we are assuming an additive model, i.e., one in which there is no interaction among any of the three factors. Suppose that we would prefer to use a design whose number of observations (h) is some fractional part of the full mnp.

In the case where m = n = p, we of course have such a design available to us in the familiar Latin square. For certain other rather specialized cases, there also exist designs in the literature, such as the Youden square [7], or some designs of Shrikhande [6, Chapter 111], for which the total number of observations is some fractional part of mnp. In general, however, for most values of (m, n, p) there is a dearth of designs for additive models which specify some fraction of mnp observations and for which the computation of the results is easily worked out.

This paper is an attempt to partially fill this gap by indicating techniques which can be used to obtain such incomplete designs for a much wider set of values of (m, n, p), and/or for a wider set of values of h in reference to a given (m, n, p). In this sense, the paper is an attempt to get designs more general than the Latin square. We shall call such designs three-dimensional incomplete block (3DIB) designs for additive models, in analogy with the ordinary (two-dimensional) incomplete block design. The "incomplete block" terminology used here for a three-factor design does not necessarily imply that one or more of the three factors in a 3DIB design must be a "blocking" effect, any more than it is required that one of the two factors in a traditional incomplete block design be a blocking effect. Mathematically, the same theory and the same analysis hold for the traditional 2DIB design whether one factor embraces treatment effects and the other blocking effects, or both embrace treatment effects. A similar situation prevails with respect to 3DIB designs. Thus the "incomplete block" terminology might in some experimental situations be considered a slight...

1 This research was supported by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command, under Contract No. AF 49(638)-213. Reproduction in whole or in part is permitted for any purpose of the United States Government.
misnomer for both 2DIB and 3DIB designs, but the nomenclature for the latter was chosen by analogy to the already-established nomenclature of the former.

In any event, the 3DIB designs treated in this paper are essentially three-factor factorial designs which assume that all interactions are absent.

1. Fundamental Theorem for MDIB Designs

This section presents a very general least squares theorem whose applicability extends beyond the 3DIB additive designs considered in this paper. The theorem also has applications to many other types of MDIB (multi-dimensional incomplete block) designs (see [5]). In this paper, the theorem will be referred to in several places. Conclusion (1.5) of the theorem will be appealed to in paragraphs 4(e) and 6(iii) (in obtaining the variance of the estimate of a factor (treatment) contrast); conclusion (1.6-1.7) will be appealed to in paragraphs 4(c) and 6(ii) (in obtaining the s.s. due to the factor). The Fundamental Theorem for MDIB Designs is as follows:

Suppose we have a vector $y$ consisting of $h$ normal, independent, and homoscedastic observations whose expectations are as specified by the linear model

$$E(y) = \gamma_1 v_1 + \gamma_0 v_0,$$

where $\gamma_1$ and $\gamma_0$ contain known constants and $v_1$ and $v_0$ consist of unknown parameters. Suppose there is a vector lying in the estimation space [3] (which consists of $\gamma_1 y$ and $\gamma_0 y$) of the form

$$Q_1 = \gamma_1 y + A_{10} \gamma_0 y,$$

whose expectation is a function of $v_1$ only and is free of $v_0$:

$$E(Q_1) = C_{11} v_1.$$

Let us define

$$\gamma_1' = \gamma_1 + \gamma_0 A_{10}' .$$

Then

$$\gamma_1' \gamma_1 = C_{11} ;$$

and

$$S.S. \quad d u c e o u \quad Q_1 = \gamma_1' Q_1 ,$$

where $\gamma_1$ is any solution $v_1$ of the system

$$C_{11} v_1 = Q_1 .$$

To prove (1.5), we start by combining (1.1), (1.2), and (1.4) to get

$$E(Q_1) = \gamma_1' v_1 + \gamma_0 v_0 .$$

A comparison of (1.3) and (1.8) tells us that

$$\gamma_1' = C_{11}$$

and

$$\gamma_0 = 0_{10} \; (n u l l \; m a t r i x) .$$
Upon pre-multiplying both sides of (1.4) by \( \gamma_1 \) and then using (1.9), we end up with (1.5).

To prove (1.6–1.7), we note that, by a standard theorem on sums of squares, (1.10) \( \text{S.S. due to } Q_1 = x'Q_1 \), where \( x \) is any solution of the system (1.11) \( \gamma_1 x = Q_1 \).

But, noting (1.5), we see that (1.10–1.11) is the same as (1.6–1.7).

2. **Conditional Error Theorem**

We state for later reference the following well-known theorem:

Let \( S_1^2 \) be the error S.S. under the model (2.1) \( E(y) = \gamma t + \epsilon \) \( \gamma \) consists of known constants, and \( t \) contains unknown parameters.) Suppose that (2.2) \( H_0 : B t = 0 \) is some linear hypothesis which can be tested under the model (2.1), and whose S.S. is \( S_1^2 \). Suppose that the model (2.1), when rewritten under the assumption that the hypothesis (2.2) is true, is of the form (2.3) \( E(y) = \gamma^0 t' \).

Let \( S_2^2 \), to be called the conditional error S.S. due to conditional error, be the error S.S. under this new model (2.3). Then (2.4) \( S_2^2 = S_1^2 - S_3^2 \).

In other words, to obtain the conditional S.S. due to a particular hypothesis, we can subtract the error S.S. for the model from the error S.S. as computed under the model which would be in effect if the hypothesis were true.

3. **Notation and Assumptions**

Closely similar notation can be used for all MDIB designs for additive models, whether 2DIB, 3DIB, or of higher dimensions. This section will explain explicitly our 3DIB notation:

(a) We will write the model equation in the form (3.1a) \( E(Y^{(a)}_{i,j,k}) = t_{i,j} + t_{i,k} \), \( a = 1, 2, \cdots, h; i = 1, 2, \cdots, m; j = 1, 2, \cdots, n; k = 1, 2, \cdots, p \) or (3.1b) \( E(y) = \gamma_1 t_1 + \gamma_2 t_2 + \gamma_3 t_3 \).
(We are not using a term \( g \) for the general mean.) The three factors, which we may call the \( I \)-factor, the \( J \)-factor, and the \( K \)-factor, are at \( m \), \( n \), and \( p \) levels respectively. The vector of the \( h \) observations, \( y \), is assumed of course to follow a multivariate normal distribution with mean vector \((3.1b)\) and variance matrix \(\sigma^2 I\). An element \(Y_{ijk}^{(a)}\) of \(y\) represents the \(a\)-th observation (out of \(h_{ijk}\) altogether) in which the \(i\)-th level of the \(I\)-factor, \(j\)-th level of the \(J\)-factor, and \(k\)-th level of the \(K\)-factor appear together. Thus the \(h_{ijk}\)'s (each of which is allowed to be \(0, 1\), or any other positive integer) form what we might call a three-dimensional incidence "matrix", \(h_{ijk}\) being the number of observations made on factor combination \((i, j, k)\). The vector of the \(m\) \(I\)-factor effects is \(t_1 = (t_{1..}, t_{2..}, \cdots t_{m..})'\); similarly, \(t_2 = (t_{.1.}, t_{.2.}, \cdots, t_{.n.})'\) and \(t_3 = (t_{..1}, t_{..2}, \cdots, t_{..p})'\). The elements of the three matrices \(\gamma_1\), \(\gamma_2\), and \(\gamma_3\) are determined in such a way as to make \((3.1b)\) equivalent to \((3.1a)\), which means that each row of each \(\gamma\)-matrix contains all \(0\)'s except for one 1.

(b) A subscript \(1, 2, \) or 3 on any vector or matrix will be associated respectively with the \(I\)-factor (\(m\) elements), \(J\)-factor (\(n\) elements), or \(K\)-factor (\(p\) elements).

(c) We define

\[ h_{..i} = \sum_j \sum_k h_{ijk}, \quad h_{.ij} = \sum_k h_{ijk}, \]

with similar definitions for \(h_{.i.}, h_{..k}, h_{i.k}, \) and \(h_{i..k}\). We will assume that, for any factor, each level appears an equal number of times in the design:

\[ h_{..i} = \frac{h}{m} \quad \text{for all} \quad i, \quad h_{.ij} = \frac{h}{n} \quad \text{for all} \quad j, \quad \text{and} \quad h_{i..k} = \frac{h}{p} \quad \text{for all} \quad k. \]

(d) We define

\[ H_{bb} = \gamma_b \gamma_b' \quad \text{for} \quad b, B = 1, 2, 3. \]

It is then easily verified that, for example, \(H_{1b}\) consists of the \(h_{i..i}\)'s, and \(H_{11}\) is [by virtue of \((3.2)\)] equal to \(h/m\) times the identity matrix. For \(b \neq B\), we will refer to the \(H_{bb}\)'s as the marginal matrices.

(e) We define \(Y_1\) and the \(Y_{.i.}'s\) by

\[ Y_1 = \gamma_1 y = (Y_{1..}, Y_{2..}, \cdots, Y_{m..})'. \]

\(Y_{.1.}, Y_{.2.}, \), and the \(Y_{..i}'s\) are similarly defined. It is then easily verified that, for example,

\[ Y_{.ij} = \sum_i \sum_k \sum_a Y_{a..k}. \]

(f) \(J_{bb}\) or \(j_b (b, B = 1, 2, 3)\) will denote respectively a matrix or vector all of whose elements are 1's. We define a flat matrix to be a matrix all of whose elements are equal.

(g) We shall use \(Q_1 = (Q_{1..}, Q_{2..}, \cdots, Q_{m..})'\) to designate a vector lying in the estimation space of the form

\[ Q_1 = Y_1 + A_{12} Y_2 + A_{13} Y_3 \]

whose expectation is a function of \(t_i\) only:

\[ E(Q_1) = C_{11} t_i \]

(h) The matrix \(C_{11}\) in \((3.4)\) we will assume to be of rank \(m - 1\).
THREE-FACTOR ADDITIVE DESIGNS

(i) $S^2_e$ will denote the error s.s. The s.s.'s for testing the hypotheses that all
I-factor, J-factor, and K-factor contrasts vanish will be denoted respectively
by $S^2_{i1}$, $S^2_{j1}$, and $S^2_{k1}$.

4. 2DIB DESIGNS: ANALYSIS

Before presenting our 3DIB results, we shall by way of introduction briefly
review the standard results for the traditional 2DIB design, using notation and
assumptions along the general lines of Section 3. Doing this will enable us to
see the close analogy between 2DIB and 3DIB analysis for additive models,
and will also get us started off at a familiar starting point. Although this section,
of course, contains no new results, certain techniques of proof presented in this
section will be referred to and used again in Section 6.

The 2DIB additive model is

$$E(Y_{ij}) = \mu + \tau_i + \gamma_j,$$

or

$$E(Y_{ij}) = \gamma_1 t_i + \gamma_2 t_j.$$

The essential results pertaining to 2DIB analysis, all of them well-known, are
as follows:

(a) If we let

$$Q_1 = Y_1 - (n/h)H_{12}Y_2,$$

then

$$E(Q_1) = C_{11}t_i,$$

where

$$C_{11} = (h/m)I_{11} - (n/h)H_{11}H_{21}. $$

(b) Suppose that $\ell_1 t_i$ is a contrast in the I-factor effects (usually called a
treatment contrast); i.e., $\ell_1$ is such that $\ell_1 j_1 = 0$. Then the best estimate of $\ell_1 t_i$
is $l_1 t_i$, where $t_i$ is any vector which satisfies

$$C_{11}t_i = Q_1.$$ 

This result is proved by demonstrating that there must exist a vector $d_i \in \mathbb{R}$
such that

$$C_{11}d_i = l_1,$$

and then by showing that

$$E(l_1 t_i) = E(d_i Q_1) = l_1 t_i.$$ 

(c) The s.s. for testing the hypothesis that all I-factor contrasts vanish is

$$S^2_{i1} = l_1 Q_1, \quad d.f. = m - 1,$$
where \( t_i \), satisfies (4.3). To prove this, note first that, because of 3(h) and because 
\[ C_{ij} t_i = 0 \]  
(null vector),
the statement that "all \( I \)-factor contrasts vanish" is equivalent to the statement that 
\[ C_{ij} t_i = 0. \]  
Hence, by (4.2),
\[ S_A^2 = \text{s.s. due to } Q_1 . \]

The result (4.5) now follows immediately from (1.6).

(d) The two other basic s.s.'s can now be obtained by applying (2.4):
\[ S_1^2 = y'y - (n/h)Y_s Y_s - S_A^2, \quad d.f. = h - m - n + 1; \]
\[ S_2^2 = y'y - (m/h)Y_l Y_l - S_A^2, \quad d.f. = n - 1. \]

(e) The variance of the best estimate \( I_i t_i \) of an \( I \)-factor contrast \( I_i t_i \) is
\[ \text{var} (I_i t_i) = I_i C_i I_i \sigma_i^2, \]
where \( C_i \) is any conditional inverse of \( C_{ij} \). (Note: Bose [3] defines \textit{conditional inverse as follows:} \( C^* \) is a conditional inverse of \( C \) if and only if
\[ C C^* = C. \]
A basic theorem about conditional inverses states that \( C^* \) is a conditional inverse of \( C \) if and only if
\[ C^* x = z \]
is a solution \( x \) of the system
\[ C x = z \]
for every \( z \) for which (4.8) is consistent. To prove (4.6), we start by using (4.4),
the fact that \( C_{ij} \) is symmetric, and (4.3) to write
\[ I_i t_i = d_i C_{ij} t_i = d_i Q_1. \]
From (4.9) and (1.5) we get
\[ \text{var} (I_i t_i) = d_i C_{ij} d_i \sigma_i^2. \]
The result (4.6) now follows by using (4.7), the symmetry of \( C_{ij} \), and (4.4).

5. 2DIB DESIGNS: EXAMPLE

In one aspect the discussion of Section 4 was more general than some of the
standard treatments of incomplete block design theory; we allowed the elements of the incidence matrix \( H_{ij} \) to assume values other than 0 or 1. One example design will suffice to illustrate how this feature can be used to advantage:

\begin{align*}
\text{Design 1.} & \quad m = 4, n = 6, h = 36, n_1 = 27, h/mn = 3/2. \\
& \quad j = 1 \ 2 \ 3 \ 4 \ 5 \ 6 \\
& \quad i = 1 \ h_{ij} = 2 \ 2 \ 2 \ 1 \ 1 \ 1 \\
& \quad 2 \ 2 \ 1 \ 1 \ 2 \ 1 \\
& \quad 3 \ 1 \ 2 \ 1 \ 2 \ 1 \\
& \quad 4 \ 1 \ 1 \ 2 \ 1 \ 2 \\
\end{align*}
THREE-FACTOR ADDITIVE DESIGNS

This simple $4 \times 6$ design with $h = 36$ observations and $n_e = 27$ d.f. for error may be considered as a $3/2$ replication of a complete design having $mn = 24$ observations and having $h_{ij} = 1$ for all $i, j$. The incidence matrix given above is patterned after that of a well-known BIB design [4, p. 329, Plan 11.11], and is obtained by adding 1 to each element of the incidence matrix of this latter design.

Design 1 provides the experimenter some added flexibility in that it allows him to use a halfway point between a design with $mn = 24$ observations and one with $2mn = 48$ observations. Computation for Design 1 is simple; for example, $C_{11}$ has the formula

$$C_{11} = (26/3)I_{11} - (13/6)J_{11}.$$  

6. 3DIB DESIGNS: ANALYSIS

This section will present the 3DIB theory we have developed for the model (3.1). The basic problem is to obtain a vector $Q_1$ of the form (3.3) which satisfies (3.4); once such a vector has been obtained, we are able to apply the theorem of Section 1.

The problem of obtaining such a vector $Q_1$ amounts essentially to finding suitable matrices $A_{12}$ and $A_{13}$. Before we find such matrices $A_{12}$ and $A_{13}$, however, we will have to make some further design assumptions. In fact, there is more than one possible set of design assumptions which we can make in order to be able to get suitable matrices $A_{12}$ and $A_{13}$. Corresponding to each such set of design assumptions we will have a design class; in other words, a given set of design assumptions will be considered to define a class of designs. Two different design classes, to be called Design Class 1 and Design Class 2, will be presented for 3DIB designs with additive models. Each design class will have different formulas for $A_{12}$ and $A_{13}$.

A given 3DIB design may belong to both design classes, to Design Class 1 only, to Design Class 2 only, or to neither design class. Obviously, a design which does not belong to either design class cannot be analyzed by the methods to be presented here.

We should point out, however, that the class(es) to which a design belongs depend on the particular permutation of the factor indices which is being assumed in the model (3.1). There are $6! = 3!$ possible permutations of these indices. The design may belong to different class(es) for different permutations. This is because, for both design classes, the design assumptions are not symmetric in the three indices. Thus the same 3DIB design may be possible to analyze (using our methods) under one permutation but not under a different permutation.

For some designs, we will have a choice between an analysis through Design Class 1 and an analysis through Design Class 2. In such cases, it appears that normally the Class 1 analysis is computationally simpler.

We now present the two design classes and the associated $A_{12}$ and $A_{13}$ matrices which result in (3.4) being satisfied:

**Design Class 1.** This design class consists of all designs which satisfy the following assumptions [in addition to those of (3.2)]:

$$(6.1) \quad H_{12}H_{33} = (h^2/mnp)J_{12} \quad \text{and} \quad H_{13}H_{22} = (h^2/mnp)J_{13}.$$
For any design belonging to Design Class 1, if we put

\[ A_{12} = -(n/h)H_{12} + (1/m)J_{12} \]

and

\[ A_{13} = -(n/h)H_{13} \]

then (3.4) will be satisfied with

\[ C_{11} = (h/m)I_{11} - (n/h)H_{12}J_{12} - (p/h)H_{13}J_{13} + (h/m^2)J_{11} \]

This statement can be proved by direct substitution, using (6.1) along with some of the relations presented in Section 3.

Design Class 2. This design class consists of all designs for which a relation of the following type holds [in addition to (3.2)]:

\[ H_{23}H_{23} = \lambda J_{23} + (s - \lambda)I_{23} \]

Any matrix \( H_{23} \) which satisfies a relation of the form (6.4) [and which also conforms to the restrictions on \( h, i, \) and \( A, k \) indicated by (3.2)] will be defined to be a matrix of RTR structure. Note that the equation (6.4) implies a certain relation between the constants \( s \) and \( \lambda \):

\[ h^2/np = \rho s + s - \lambda. \]

For any design belonging to Design Class 2, if we put

\[ A_{12} = -(n/h)H_{12} - (1/\rho)\left(\frac{n}{h}H_{12}J_{23} - H_{13}\right)H_{23} \]

(6.5) and

\[ A_{13} = (n/p\rho)\left(\frac{n}{h}H_{13}J_{23} - H_{12}\right), \]

then (3.4) will be satisfied with

\[ C_{11} = (h/m)I_{11} + A_{12}H_{21} + A_{13}H_{31} \]

From this point on, the analysis of 3DIB designs will be the same (except for the different formulas for \( C_{11}, \) and \( Q_{1} \)) regardless of whether we are dealing with Design Class 1 or Design Class 2. First, let us give some 3DIB results (relating to both design classes) which are so closely similar to certain 2DIB results that we can refer to Section 4 for the proofs:

(i) The best estimate of an \( I \)-factor contrast \( l^i \) is \( l^i \hat{t}_i \), where \( \hat{t}_i \) is any vector satisfying

\[ C_{11} \hat{t}_i = Q_1. \]

Proof of this is along the lines indicated in 4(b).

(ii) By using a proof similar to that of 4(e), we obtain

\[ S_{23} = \hat{t}_i Q_1, \quad d.f. = m - 1. \]

(iii) The variance of \( l^i \hat{t}_i \) is given by

\[ \text{var} (l^i \hat{t}_i) = \hat{t}_i C_{11} \hat{t}_i \sigma^2. \]

Refer to 4(e) for proof. [Note that the proofs of both (ii) and (iii) utilize the Fundamental Theorem of Section 1.]
The remaining 3DIB results are a little more complicated than the corresponding 2DIB results:

(iv) Let us consider the basic s.s.'s besides $S_{s1}^2$. By using (2.4), we can write

\[ S_{s1}^2 = S_{t1}^2, \quad d.f. = h - m - n - p + 2, \]

\[ S_{s2}^2 = S_{r2}^2 - S_{t2}^2, \quad d.f. = n - 1, \]

\[ S_{s3}^2 = S_{r3}^2 - S_{r1}^2, \quad d.f. = p - 1, \]

where we define $S_{s1}^2$, $S_{s2}^2$, and $S_{s3}^2$ to be the error s.s.'s under the models

\[ E(y) = \gamma_{s1}t_1 + \gamma_{s2}t_2, \]

\[ E(y) = \gamma_{s1}t_1 + \gamma_{s2}t_2, \]

\[ E(y) = \gamma_{s1}t_1 + \gamma_{s2}t_2, \]

respectively. Note that these three s.s.'s due to conditional error can be calculated by applying the 2DIB theory of Section 4; $S_{s1}^2$, for example, is nothing more than the error s.s. based on a 2DIB model with incidence matrix $H_{23}$ (whose elements are the $h_{ik}$'s).

Once $S_{s1}^2$, $S_{s2}^2$, $S_{s3}^2$ have been obtained, we can make $F$-tests in the normal manner to test for equality of $I$-factor, $J$-factor, or $K$-factor effects.

(v) Finally, let us consider the problem of obtaining the best estimate of a $J$-factor contrast $l_jt_2$, or of a $K$-factor contrast $l_kt_3$. One way we might obtain these, of course, is to permute the indices in the model equation (3.1), and to run the same 3DIB analysis as already described using the new model with permuted indices. This will not always work, however, since the design corresponding to a new model with permuted indices may belong to neither Design Class 1 nor to Design Class 2.

An alternative method, which is completely general, consists of solving the normal equations after plugging in the value of the vector $\hat{t}_i$, obtained in solving (6.7). The original normal equations for the model (3.1) are

\[ H_{s1}t_1 + H_{s2}t_2 + H_{s3}t_3 = Y_1 \]

\[ H_{s2}t_1 + H_{s2}t_2 + H_{s3}t_3 = Y_2 \]

\[ H_{s3}t_1 + H_{s2}t_2 + H_{s3}t_3 - Y_3 = Y_3. \]

If we substitute $\hat{t}_i$ [where $\hat{t}_i$ satisfies (6.7)] for $t_i$ in (6.8), and then drop the first equation of the trio (6.8), we end up with the system

\[ H_{s2}t_2 + H_{s2}t_3 = Z_2 \]

\[ H_{s2}t_2 + H_{s3}t_3 = Z_3, \]

where

\[ Z_b = Y_b - H_{s1}t_1, \quad b = 2, 3. \]

Now let $\hat{t}_2$, $\hat{t}_3$ be any solution to (6.9). [Note that the left-hand side of (6.9) is precisely the left-hand side of the normal equations for a 2DIB design with incidence matrix $H_{23}$, so that 2DIB theory can be utilized in solving (6.9).]
It can easily be shown that the vectors $\mathbf{\hat{t}}$, $\mathbf{\hat{t}}_2$, $\mathbf{\hat{t}}_3$ which we have obtained satisfy the full system (6.8). Hence $\mathbf{\hat{t}}_1$, $\mathbf{\hat{t}}_2$, and $\mathbf{\hat{t}}_3$ will be the best estimates of $\mathbf{t}_1$, $\mathbf{t}_2$, and $\mathbf{t}_3$ respectively.

7. 3DIB Designs: Examples and Construction

This section will present several examples of 3DIB designs for additive models selected from over a dozen such example designs appearing in [5]; we will also allude to the methods used to construct the designs we present. The example designs in [5] are all original except for a few which originated as duplicate bridge movements.

To start, we should note that certain simpler types of 3DIB designs can be analyzed easily without resorting to the high-powered methods of Section 6. In particular, designs all of whose marginal matrices are flat would fall in this category. This would include not only Latin squares, but also many other designs, without equal numbers of levels of all three factors, which can be constructed with all marginal matrices flat [5].

Youden squares [7] and Shrikhande's designs [6, Chapter III.] both have specialized theories developed for analyzing them, but alternatively their analyses could be handled (and in a more general manner) by the methods of Section 6.

The first 3DIB example design which we will present is one which may be considered as a 3/2 replication of a $4 \times 4 \times 4$ Latin square:

Design 2. $m = n = p = 4$, $h = 24$, $n_\tau = 14$, $h/mnp = 3/8$.

$$
\begin{align*}
&h_{ijk} = 1 \\
&\text{for } j = 1 \ 2 \ 3 \ 4 \\
&i = 1 \ k = 3, 4 \ 4 \ 1, 2 \ 2 \\
&2 \ 2, 3 \ 3 \ 1, 4 \ 1 \\
&3 \ 1 \ 2, 4 \ 3 \ 2, 4 \\
&4 \ 4 \ 1, 3 \ 2 \ 1, 3 \\
&\text{(and } h_{ijk} = 0 \text{ otherwise}).
\end{align*}
$$

Since the marginal matrices for this design satisfy

\begin{equation}
H_{12}H_{23} = 9J_{13}, \quad H_{13}H_{32} = 9J_{12}, \quad \text{and} \quad H_{23}H_{31} = 9J_{23},
\end{equation}

it belongs to Design Class 1 (6.1) for all 6 permutations of the factor indices. It does not, incidentally, belong to Design Class 2 (6.4) for any permutation of the factor indices.

Computation of $C_{\tau1}$ (6.3) is a simple matter; a conditional inverse $C_{\tau1}^{-1}$ can be easily obtained.

Design 2 was constructed in two stages. First, three marginal matrices with the desired properties (7.1) were written down. Then the $h_{ijk}$'s were selected so as to give these three marginal matrices.

The next example design originated as a duplicate bridge movement, and
appears (although in somewhat different form) in Beynon [2, p. 22, table labelled "Howell Master Sheet for 8 Pairs—4 Tables"). It has the following design plan:

\[ \text{Design 3. } m = 7, n = 8, p = 14, h = 56, n_r = 29, h/mnp = 1/14. \]

\[ h_{ijk} = 1 \]

\[
\begin{array}{cccccccc}
  j = 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
  i = 1 & k = 1B & 6A & 4A & 7B & 7A & 4B & 6B & 1A \\
  2 & 7B & 2B & 7A & 5A & 1B & 1A & 5B & 2A \\
  3 & 6B & 1B & 3B & 1A & 6A & 2B & 2A & 3A \\
  5 & 4B & 4A & 1B & 3B & 5B & 3A & 1A & 5A \\
  6 & 2A & 5R & 5A & 2R & 4R & 6R & 4A & 6A \\
  7 & 5A & 3A & 6B & 6A & 3B & 5B & 7B & 7A \\
\end{array}
\]

Instead of letting \( k \) run from 1 to 14, we are letting \( k = 1A, 2A, \ldots, 7A, 1B, 2B, \ldots, 7B \), for reasons which will become apparent when we discuss the construction of the design.

The marginal matrices of Design 3 conveniently observe the following relations:

\[
\begin{aligned}
H_{12} &= J_{12} \\
H_{12}H_{23} &= \frac{m + 1}{2} J_{13} \\
H_{22}H_{13} &= \frac{m + 1}{2} J_{23} \\
H_{12}H_{23} &= \frac{m + 1}{2} J_{12} \\
H_{13}H_{31} &= (m + 1)J_{11} \\
H_{12}H_{31} &= \frac{m + 1}{2} (I_{11} + J_{11}) \\
H_{21}H_{12} &= mJ_{22} \\
H_{22}H_{33} &= \frac{m + 1}{2} I_{22} + \frac{m - 1}{2} J_{22}.
\end{aligned}
\]

The value of \( m \) for Design 3 is of course 7. The relations (7.2) ensure that the design belongs to Design Class 1 for all permutations of the factor indices; that the \( C_{11} \) matrix (6.3) is quasi-diagonal (where we define a quasi-diagonal matrix to be one which can be represented as the sum of a diagonal matrix and a flat matrix) and thus has a diagonal conditional inverse; and that, if the factor indices are permuted so that the \( J \)-factor replaces the \( I \)-factor, then the (new) \( C_{11} \) matrix will be quasi-diagonal.

The method of constructing Design 3 results in the relations (7.2) being satisfied. The construction is rather tricky but the key to it is the first line of the design plan, which (along with the heading line) is:

\[
\begin{array}{cccccccc}
  j = 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
  i = 1 & k = 1B & 6A & 4A & 7B & 7A & 4B & 6B & 1A \\
\end{array}
\]

Note that the remaining 6 lines of the design plan can be filled in by working diagonally down and to the right, disregarding the last (8th) column, and adding
1 (mod 7) each time to the integral portion of the value of \( k \). The last column is filled in by working straight down the column and adding 1 each time.

Now let us re-write the first line of the design plan along with the heading line (given just above) in a somewhat different form:

\[
\begin{align*}
\text{(Line \( j \))} & \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \\
\text{(Line \( k-A \))} & \quad 6A \quad 4A \quad 7A \\
\text{(Line \( k-B \))} & \quad 1B \quad 7B \quad 4B \quad 6B
\end{align*}
\]

We may form certain differences (mod 7):

\[
\begin{align*}
\text{(Line \( D-A \))} & \quad 6-2=4 \quad 4-3=1 \quad 7-5=2 \\
\text{(Line \( D-B \))} & \quad 1-1=7 \quad 7-4=3 \quad 4-0=5 \quad 0-7=0
\end{align*}
\]

Line \( D-A \) was obtained by subtracting line \( j \) from (the integral portion of) line \( k-A \). Line \( D-B \) was obtained by subtracting line \( j \) from line \( k-B \). Therefore lines \( D-A \) and \( D-B \) amount essentially to the 7th column of the design plan re-written (going from bottom of column to top).

Now observe the following:

(i) The four integers in line \( k-A \) (1, 4, 6, 7) form a difference set (mod 7), as do the four integers in line \( k-B \) (1, 4, 6, 7). This is sufficient to ensure that every row of the design plan has exactly four elements in common with every other row, thereby causing \( \mathbf{H}_{1a} \) to be a matrix of BIB structure. (The difference sets here play the same role as do the difference sets that are used to generate certain ordinary BIB designs.)

(ii) The three integers in line \( D-A \) (1, 2, 4) form a difference set (mod 7), as do the four integers in line \( D-B \) (3, 5, 6, 7). This is sufficient to ensure that every column of the design plan has exactly three elements in common with every other column, thereby causing \( \mathbf{H}_{1a} \) to be a matrix of BIB structure.

(iii) In lines \( D-A \) and \( D-B \) every integer from 1 to 7 appears exactly once. This is sufficient to ensure that every column of the design plan contains each integer exactly once. Since every row of the design plan contains each of four integers with both an \( A \) and a \( B \) suffix, it follows that any row of the design plan must have exactly four elements in common with any column. This ensures then that \( \mathbf{H}_{1a} \mathbf{H}_{1a} \) is a flat matrix. Thus, we see that, if the first line of the design plan can be chosen so that certain relations hold, then the various properties (7.2) will automatically be satisfied.

Design 3 is one of a group of five designs given in [5] which all have similar properties but which have different numbers of levels of the three factors. The relations (7.2) are all satisfied for all five of the designs (each of which has a different value of \( m \)), and, in addition, the following properties also hold for all five designs:

(a) \( n = m + 1 \), \( p = 2m \), \( h = mn = m(m + 1) \), \( n_* = m^2 - 3m + 1 \), and \( h/mnp = 1/2m \). The small value of \( h/mnp \) should be considered as a highly desirable property of these designs.
(b) The variance of the estimate of the difference between any two I-factor effects is \( (2/m)\sigma^2 \).
(c) The variance of the estimate of the difference between any two J-factor effects is \( (2/(m - 1))\sigma^2 \).
(d) Each design has an efficiency of \( m/(m + 1) \) for estimating I-factor contrasts (where a complete design with \( mnp \) observations is used as the basis for comparison).
(e) Each design has an efficiency of \( (m - 1)/m \) for estimating J-factor contrasts.

The four designs which are similar to Design 3 are for \( m = 11, m = 19, m = 9, \) and \( m = 5. \)

The designs for \( m = 11 \) and \( m = 19 \) are constructed in exactly the same fashion as Design 3. The first line of the design plan of the \( m = 11 \) design, which is \( 11 \times 12 \times 22 \), is:

\[ k = 1B, 8B, 8A, 11B, 7B, 3B, 11A, 9A, 7A, 9B, 3A, 1A. \]

The first line of the design plan of the \( m = 19 \) design, which is \( 19 \times 20 \times 38 \), is:


The remainders of both design plans can be filled in by working from the first lines and using the same technique that was indicated for Design 3. The design for \( m = 11 \) is an existing duplicate bridge movement [1, p. 49], but the design for \( m = 19 \) is original.

The design for \( m = 9, \) which will not be given here, was constructed using a more complicated method than that used for the designs for \( m = 7, 11, \) and \( 19 \)—a method which is an extension of the one used for these other three designs. The design for \( m = 5 \) was constructed strictly by trial and error, and is as follows:

**Design 4.** \( m = 5, n = 6, p = 10, h = 30, n_0 = 11, h/mnp = 1/10. \)

\[ h_{ijk} = 1 \]

<table>
<thead>
<tr>
<th>i</th>
<th>j = 1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>k = 1</td>
<td>6</td>
<td>3</td>
<td>4</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>10</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>8</td>
<td>7</td>
<td>2</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>7</td>
<td>1</td>
<td>9</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>2</td>
<td>10</td>
<td>6</td>
<td>3</td>
<td>7</td>
</tr>
</tbody>
</table>

(Here we are letting \( k \) run from 1 to 10 in conventional fashion.)

An attempt was made to construct a design similar to the other five with \( m = 3, \) but this turned out to be impossible.
Our final example design may be considered as a 4/7 replication of a $7 \times 7 \times 7$ Latin square:

**Design 5.** $m = n = p = 7$, $h = 28$, $n_s = 9$, $h/mnp = 4/49$.

$k_{iik} = 1$

for $j = 1\ 2\ 3\ 4\ 5\ 6\ \tilde{i}$

$i = 1\ k = 1\ 3\ 5\ 2$

$2\ 2\ 4\ 6\ 3$

$3\ 3\ 5\ 7\ 4$

$4\ 5\ 4\ 6\ 1$

$5\ 6\ 5\ 7\ 2$

$6\ 3\ 7\ 6\ 1$

$7\ 2\ 4\ 1\ 7$

Design 5 belongs to Design Class 2 (6.4) for all 6 permutations of the factor indices, but does not belong to Design Class 1 for any permutation of the factor indices. The design has been constructed so that the following relations hold:

\[(7.3)\quad H_{12} = H_{13} = H_{23} = H = (say),\quad H_{11} = H_{21} = H_{31} = H',\]

\[H + H' = I + J,\quad HH' = H'H = 2I + 2J.\]

Thus the equations (6.5) become

\[
A_{12} = (1/7)(I - 2H),\quad A_{13} = (\ell/14)(2J - 2I - 3H);
\]

and the $C_{11}$ matrix (6.6) is

\[C_{11} = 3I - (3/7)J.\]

It can be shown that the C-matrix has this same formula for all 6 permutations of the factor indices. However, the formulas for $A_{12}$ and $A_{13}$ will not stay the same from one permutation to another. Design 5 has an efficiency of $3/4$ for estimating $I$-factor, $J$-factor, or $K$-factor contrasts (where the Latin square is used as the standard of comparison).

The key to constructing Design 5 lies in choosing the first line of the design plan properly. Once the first line is determined, the rest of the design plan is filled in by working diagonally down and to the right, adding 1 (mod 7) each time. Note that the first line has been chosen so that the difference set $(1, 2, 3, 5)$ appears in both the first row and the first column of the design plan, and also so that in the first row it appears under $j$-values of $(1, 2, 3, 5)$. All this results in the relations (7.3) being satisfied.

It should be borne in mind that the designs presented in this section constitute just a handful of illustrative examples, and that is is possible to construct many more useful 3DIB designs for additive models based on either Design Class 1 (6.1-6.3) or Design Class 2 (6.4-6.6).
We present finally two numerical examples to illustrate the analysis of 3DIB designs. Although both of the examples are based on live experimental data, the two experiments were run not for their own sake, but purely for the purpose of obtaining non-artificial figures to illustrate 3DIB analysis. Thus these particular experiments, although perhaps otherwise of no use, at least serve the intended function of providing examples.

In the first example, Design 4 was used in an experiment to measure traffic flow at 10 different points in and around the campus and downtown section of Chapel Hill. The observations were made on 5 mornings (Monday, September 18 through Friday, September 22, 1961) between 8 a.m. and 9 a.m., with 6 observations being made each morning. A given “observation” consisted of counting the number of vehicles passing the specified point during a 5-minute period. On a given morning, the first observation was made from 8:00 to 8:05, the second one from 8:10 to 8:15, and so forth, the sixth one being made from 8:50 to 8:55. Thus the 5 levels of the I-factor represented the 5 days of the week, the 6 levels of the J-factor represented the 6 different times of the morning, and the 10 levels of the K-factor represented the 10 locations. Note that the fact that the design is such that \( h_{ij} = 1 \) enabled all 30 observations to be made by a single observer (i.e., the design never required the observer to be in two places at the same time) and also utilized the observer’s time efficiently (since the observer had an observation to make during all 6 times on all 5 mornings).

The raw results of the experiment (i.e., the number of vehicles counted for each specified time-place combination) are given in the following table, where the number in parentheses indicates the location (level of the K-factor):

<table>
<thead>
<tr>
<th>Time</th>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:00</td>
<td>72(1)</td>
<td>49(2)</td>
<td>62(3)</td>
<td>52(4)</td>
<td>57(5)</td>
</tr>
<tr>
<td>8:10</td>
<td>101(6)</td>
<td>50(1)</td>
<td>13(8)</td>
<td>35(7)</td>
<td>55(2)</td>
</tr>
<tr>
<td>8:20</td>
<td>59(3)</td>
<td>98(9)</td>
<td>49(7)</td>
<td>89(1)</td>
<td>89(1)</td>
</tr>
<tr>
<td>8:30</td>
<td>53(4)</td>
<td>92(10)</td>
<td>50(2)</td>
<td>82(9)</td>
<td>46(6)</td>
</tr>
<tr>
<td>8:40</td>
<td>10(8)</td>
<td>38(5)</td>
<td>73(9)</td>
<td>46(6)</td>
<td>67(5)</td>
</tr>
<tr>
<td>8:50</td>
<td>78(10)</td>
<td>12(8)</td>
<td>54(4)</td>
<td>67(5)</td>
<td>48(7)</td>
</tr>
</tbody>
</table>

Since it was guessed that each observation might follow roughly a Poisson distribution, it was decided to analyze the square roots of the observations rather than the raw observations themselves, in order to stabilize the variance:

\[
\begin{array}{ccccccc}
  \text{Time} & 8:00 & 8:10 & 8:20 & 8:30 & 8:40 & 8:50 \\
\text{Monday} & 8.49(1) & 10.05(6) & 7.68(3) & 7.28(4) & 3.16(8) & 8.53(10) \\
\text{Tuesday} & 7.00(2) & 7.07(1) & 9.90(9) & 9.59(10) & 6.16(5) & 3.46(8) \\
\text{Wednesday} & 7.87(3) & 3.61(8) & 7.00(7) & 7.07(2) & 8.54(9) & 7.35(4) \\
\text{Thursday} & 7.21(4) & 5.92(7) & 9.43(1) & 9.06(9) & 6.78(6) & 8.19(5) \\
\text{Friday} & 7.55(5) & 7.42(2) & 10.00(10) & 6.78(6) & 5.83(3) & 6.93(7) \\
\end{array}
\]
An argument might be made that, for purposes of satisfying the additivity assumptions, a log transformation would have been better than this square root transformation. However, the log transformation would probably not stabilize the variance as well. As for the additivity assumptions, it would appear that there are no substantial interactions even with the square root transformation, although location and time of morning might be expected to interact slightly.

In analyzing the results of the experiment, we start by using the table above to obtain

\[
\begin{align*}
Y_1 &= (45.49, 43.18, 41.44, 46.59, 44.51)' \\
Y_2 &= (38.12, 34.07, 44.01, 39.78, 30.47, 34.76)' \\
\end{align*}
\]

Since we will analyze Design 4 through Design Class 1, we use (6.2) and (3.3) to get

\[
Q_1 = Y_1 - (1/3)H_{13}Y_3 = (1/3)(6.00, -4.99, 2.03, .08, -3.12)'
\]

[For example, the third element of \(Q_1\) is obtained by computing
\[
\]

Using (7.2) and (6.3), we obtain

\[
C_{11} = 5I_{11} - J_{11},
\]

of which a conditional inverse is \(\frac{1}{3} I_{11}\). Hence we can write

\[
\hat{t}_1 = (1/5)Q_1 = (.40, -.33, .14, .01, -.21)'
\]

and

\[
S_{11} = \hat{t}_1Q_1 = (1/45)[(6.00)^2 + (-4.99)^2 + \cdots] = 1.66.
\]

By noting 4(a) and 4(b), we see that a solution \(t_2\) to the system (6.9) will be given by

\[
\begin{equation}
(8.1) \quad t_2 = \left( h I_{22} - \frac{p}{h} H_{22}H_{32} \right)^* \left( Z_2 - \frac{p}{h} H_{22}Z_3 \right)
\end{equation}
\]

In the case of Design 4, this formula for \(t_2\) simplifies. Substituting from (7.2), we obtain

\[
\hat{t}_2 = (4I_{22} - \frac{1}{3} I_{22})(Y_2 - \frac{1}{3} Y_3) = (1/4)(Y_2 - \frac{1}{3} Y_3)
\]

Hence

\[
\hat{t}_2 = (1/12)(2.76, 2.04, 9.89, -3.52, -13.21, 2.04)'
\]

\[
= (.23, .17, .82, -.29, -1.10, .17)'
\]
The $t_3$ vector satisfying (6.9) is given by

$$t_3 = (p/h)(Z_3 - H_{32}t_2)$$

Thus

$$t_3 = (7.90, 7.26, 7.03, 7.06, 7.71, 8.21, 6.25, 3.60, 9.42, 9.29)' .$$

Note that the $t_{3..}'s$ and the $t_{.i}'s$ add to zero, but the $t_{i..}'s$ as we have them here add to $p$ times the over-all mean. If we want an estimate of average number of vehicles per five-minute period for each of the 10 locations, we can square each element of $t_3$ to obtain the (perhaps slightly biased) estimate vector

$$(62.4, 52.7, 49.4, 49.8, 59.4, 67.4, 39.1, 13.0, 88.7, 86.3).$$

We have yet to obtain certain s.s.'s. In order to get $S_{i1}^2$, we first need to find $S_{i3}^2$, the error s.s. for the 2DIB design with incidence matrix $H_{22}$. Referring to Section 4, we may write

$$S_{i1}^2 = Y'y - (p/h)Y_3Y_3$$

(8.3)

$$- (Y_3 - \frac{p}{h} H_{32}Y_3)'(I_{22} - \frac{p}{h} H_{32}H_{32})^{-1}(Y_3 - \frac{p}{h} H_{32}Y_3).$$

Hence for Design 4 we have

$$S_{i1}^2 = 1723.93 - \frac{1}{3}(5120.07) - \frac{1}{36}[(2.76)^2 + (2.04)^2 + \cdots]$$

$$= 1723.93 - 1706.69 - 8.35 = 8.89.$$

Thus

$$S_{i3}^2 - S_{i1}^2 = 8.89 - 1.66 = 7.23.$$

For this particular design, it so happens that there is a technique for getting $S_{i3}^2$ which is easier than the one based on the conditional error principle. Referring to Section 7, we see that the design still belongs to Design Class 1 if the factor indices are permuted so that the J-factor trades places with the I-factor. Since the C-matrix associated with that permutation is equal to $4I_{22} - \frac{2}{3}J_{22}$, we have

$$S_{i3}^2 = \hat{t}(4I_{22} - \frac{2}{3}J_{22})\hat{t}.$$
principle, since the error s.s. pertaining to the design with the incidence matrix
$H_{12} = J_{12}$ is just

$$S_{t2} = y'y - (m/h)y_1y_1 - (n/h)y_2y_2 + (1/h)y_3y_3 \ldots$$
$$= 1723.93 - (1/6)(9802.89) - (1/5)(8269.91) + (1/30)(221.21)^2$$
$$= 67.26$$

(where $Y_{...}$ denotes the sum of all $k$ observations). Therefore

$$S_{t2} = S_{t2} - S_t = 67.26 - 7.23 = 60.03.$$

An analysis of variance table we might prepare as follows:

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>D.f.</th>
<th>S.S.</th>
<th>M.S.</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days (I-factor)</td>
<td>4</td>
<td>1.66</td>
<td>.415</td>
<td>.63</td>
</tr>
<tr>
<td>Times of morning (J-factor)</td>
<td>5</td>
<td>8.35</td>
<td>1.67</td>
<td>2.54</td>
</tr>
<tr>
<td>Locations (K-factor)</td>
<td>9</td>
<td>60.03</td>
<td>6.67</td>
<td>10.15</td>
</tr>
<tr>
<td>Error</td>
<td>11</td>
<td>7.23</td>
<td>.657</td>
<td></td>
</tr>
</tbody>
</table>

$S_{t2}$ is highly significant ($\alpha < .0005$), indicating (not surprisingly) that different locations have different traffic volumes. The $F$-test for $S_{t2}$ is significant at the 10% but not at the 5% level; hence we might suspect (looking at the $F$'s) that traffic builds up to a peak around 8:20 to 8:25 and then declines. The sum of squares for days is nowhere close to being significant, and so it would appear that the traffic volume is not much different from one week-day to the next.

In our second example, Design 5, which must be analyzed through Design Class 2, was used in an experiment aimed at determining the effect of certain factors on the estimation of areas. Each of 7 statisticians (a, b, c, d, e, f, g) was asked to estimate the area of each of 4 geometrical figures of variegated shapes and colors. There were 7 shapes (square, U, triangle, L, half-doughnut, rectangle, and circle, or S, U, T, L, H, R, C) and 7 colors (red, turquoise, yellow, black, purple, green, and salmon, or R, T, Y, B, P, G, S). The 7 statisticians represented the 7 levels of the I-factor, the 7 shapes were the 7 levels of the J-factor, and the 7 colors were the 7 levels of the K-factor. If individual differences were not in their own right one of the objects of experimental investigation, the 7 statisticians could be considered as blocks in the traditional sense, i.e., as constituting an unwanted but unavoidable source of variation; however, the results give some information about individual differences which could be interesting for its own sake. Note, though, that this experiment illustrates one of the reasons why incomplete block designs are needed: the block size (if we consider the 7 individuals as blocks) is 4, and probably must be kept small in order to avoid fatigue effects (and in any event cannot be made greater than 7, which is the number of shapes).

As in the case of the previous example, it is difficult to know to what extent the assumptions of the model are satisfied. The additivity assumption probably holds pretty well, although it might not be too surprising, e.g., if there were a little interaction between individuals and shapes. For the homoscedasticity
assumption, an attempt was made to get the variance more nearly uniform by making a transformation of the original data. The raw results of the experiment, in terms of the number of squares inches estimated by the different individuals to be the areas of the figures with the specified shapes and colors, are as follows (where the letter in parentheses denotes the color):

<table>
<thead>
<tr>
<th></th>
<th>S</th>
<th>U</th>
<th>T</th>
<th>L</th>
<th>H</th>
<th>R</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>10.0(R)</td>
<td>8.5(Y)</td>
<td>10.0(P)</td>
<td></td>
<td></td>
<td>10.5(T)</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>10.0(T)</td>
<td>6.9(B)</td>
<td>7.0(G)</td>
<td></td>
<td>7.5(Y)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c</td>
<td></td>
<td>8.0(Y)</td>
<td>10.0(P)</td>
<td>7.5(S)</td>
<td></td>
<td>12.0(B)</td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>17.0(P)</td>
<td></td>
<td>15.0(B)</td>
<td>8.5(G)</td>
<td>16.0(R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e</td>
<td></td>
<td>9.0(G)</td>
<td></td>
<td>10.2(P)</td>
<td>10.5(S)</td>
<td>7.1(T)</td>
<td></td>
</tr>
<tr>
<td>f</td>
<td>24.0(Y)</td>
<td></td>
<td>18.0(S)</td>
<td>15.0(G)</td>
<td>28.0(R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>g</td>
<td>5.0(T)</td>
<td>5.0(B)</td>
<td></td>
<td>5.5(R)</td>
<td></td>
<td>7.0(S)</td>
<td></td>
</tr>
</tbody>
</table>

Each shape was of course presented in 4 different colors during the course of the experiment, and a figure with a given shape had the same actual area each of the 4 times it was presented. The actual areas for the 7 shapes were respectively 12.96, 12.10, 12.16, 12.60, 11.64, 13.50, and 13.86 square inches. The transformation made on the original data in an attempt to stabilize the variance consisted of dividing the estimated area by the actual area, and then taking the logarithm (to the base 10) of this fraction (since it was guessed that the fraction would have a standard deviation roughly proportional to its mean). For example, for the black triangle whose area individual b estimated to be 6.9, we calculate

\[ \log 6.9 - \log 12.16 = -.246. \]

After transforming all our observations, we end up with a table to work with as follows (in which the levels of all three factors are now specified by integers rather than letters):

<table>
<thead>
<tr>
<th></th>
<th>j = 1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-.113(1)</td>
<td>-.153(3)</td>
<td>-.089(5)</td>
<td></td>
<td>-.049(2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>-.083(2)</td>
<td>-.246(4)</td>
<td>-.255(6)</td>
<td></td>
<td>-.255(3)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>.118(5)</td>
<td></td>
<td>-.182(3)</td>
<td>-.100(5)</td>
<td>-.191(7)</td>
<td></td>
<td>-.063(4)</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>-.129(0)</td>
<td>.076(4)</td>
<td>-.137(6)</td>
<td>.074(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.268(3)</td>
<td>.170(7)</td>
<td></td>
<td>.097(5)</td>
<td>-.109(7)</td>
<td>-.290(2)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>- .414(2)</td>
<td>-.384(4)</td>
<td></td>
<td>-.360(1)</td>
<td></td>
<td>.046(6)</td>
<td>.305(1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
To start off, we use this table to obtain

\[
Y_1 = (-.396, -.839, -.536, .131, -.585, .789, -1.455)'
\]
\[
Y_2 = (-.141, -.749, -.343, -.639, -.430, -.244, -.345)'
\]
\[
Y_3 = (-.094, -.832, -.322, -.617, -.124, -.475, -.427)'
\]

Referring to (6.5) and (3.3) and to Section 7, we may write

\[
Q_i = Y_i + (1/7)(1 - 2H)Y_2 + (1/14)(2J - 2I - 3H)Y_3
\]
\[
= (1/14)(14Y_1 + 2Y_2 - 2Y_3 - 4HY_2 - 3HY_3 + 2Y_3... j)
\]

Since

\[
HY_2 = (-1.663, -1.975, -1.757, -1.454, -1.768, -1.073, -1.874)'
\]
\[
HY_3 = (-1.372, -2.246, -1.490, -1.310, -1.858, -1.318, -1.970)'
\]

and \(Y_{...} = -2.891\), we have

\[
Q_1 = (1/14)(-.652, -2.724, -1.830, 5.754, -1.938, 13.972, -12.582)'
\]

Referring to the discussion of Design 5 in Section 7, we see that \((1/3)i_i\) is a conditional inverse of \(C_{11}\). Hence we get

\[
i_i = (1/3)Q_i = (-.016, -.065, -.044, .137, -.046, .333, -.300)'
\]

and

\[S_{i_i}^2 = (1/588)[(-.652)^2 + (-2.724)^2 + \cdots] = .6830.\]

If we want an estimate of the average per cent by which an individual tends to under-estimate or over-estimate area, we can add -.103 \([= (1/28)Y_{...}]\) to the appropriate element of \(i_i\), take the antilog, and subtract 1. E.g., we can estimate that individual \(b\) tends to under-estimate area by 32.1% on the average:

\[\text{antilog} (-.065 - .103) - 1 = -32.1\%\]

The estimates of these average per cents for all 7 individuals are given in the following vector:

\[(-24.0, -32.1, -28.7, +8.1, -29.0, +69.8, -60.5).\]

We now must find \(\bar{i}_1\) and \(\bar{i}_2\). For Design 5, (8.1) becomes

\[
\bar{i}_1 = (\bar{I}_{12} - \bar{J}_{12})^2[ (Y_2 - 4HY_3) + (4HH' - H')i_i]
\]
\[
= (1/14)(4Y_2 - HY_3 + 2I_i - 4H'i_i).
\]

Since

\[H'i_i = (.155, -.426, .209, -.271, .032, .359, -.057)'
\]

we get

\[\bar{i}_1 = (.011, .059, -.058, .008, -.006, -.021, .016)'\].
Using (8.2), we obtain

\[ \hat{t}_i = (1/4)(Y_i - H'\hat{t}_i). \]

Hence

\[ H'\hat{t}_i = (.004, .080, -.019, .025, -.045, .030, -.079)', \]

we have

\[ \hat{t}_i = (-.063, -.122, -.128, -.093, -.028, -.216, -.073)', \]

To obtain \( S^2 \), we first need to find \( S^2_{\text{II}} \). For Design 5, (8.3) becomes

\[ S^2_{\text{II}} = y'y - (1/4)Y_iY_i - (2/7)(Y_i - \frac{1}{7}HY_i)'(Y_i - \frac{1}{7}HY_i). \]

Thus

\[ S^2_{\text{II}} = 1.2188 - (1/4)(1.6088) = .7583 \]

Hence

\[ S^2 = S^2_{\text{II}} - S^2_{\text{II}} = .7583 - .6830 = .0753. \]

As it happens, we will not need to resort to the conditional error formula to get either \( S^2_{\text{II}} \) or \( S^2_{\text{II}} \) for this particular design. Recalling that Design 5 belongs to Design Class 2 for all 6 permutations of the factor indices, and that the C-matrix is equal to \( 3I - (3/7)J \) for every permutation, we can write simply

\[ S^2_{\text{II}} = \hat{t}_i(3I - \frac{2}{7}J)\hat{t}_i = 3\hat{t}_i = .0245 \]

and

\[ S^2_{\text{II}} = \hat{t}_i(3I - \frac{2}{7}J)\hat{t}_i = 3\hat{t}_i - (3/112)Y^2... = .0660. \]

Finally, the analysis of variance table is as follows:

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>D.f.</th>
<th>S.S.</th>
<th>M.S.</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individuals (I-factor)</td>
<td>6</td>
<td>.6830</td>
<td>.1138</td>
<td>13.61</td>
</tr>
<tr>
<td>Shapes (J-factor)</td>
<td>6</td>
<td>.0245</td>
<td>.0041</td>
<td>.49</td>
</tr>
<tr>
<td>Colors (K-factor)</td>
<td>6</td>
<td>.0660</td>
<td>.0110</td>
<td>1.31</td>
</tr>
<tr>
<td>Error</td>
<td>9</td>
<td>.0753</td>
<td>.00837</td>
<td></td>
</tr>
</tbody>
</table>

The F-test for individuals is highly significant (\( \alpha < .0005 \)). Thus there is definite evidence that there were real differences among individuals with respect to the per cent by which they tended to over-estimate or under-estimate, and that some persons were consistent over-estimators while others were consistent under-estimators. Neither \( S^2_{\text{II}} \) nor \( S^2_{\text{II}} \) is significant; hence, if either color or shape had any influence on the estimates of areas, this experiment was too small and/or had too high an error variance to detect such influences. It is possible that the error variance could be substantially reduced by changing certain conditions of the experiment so that it would be less difficult for the subjects to make their estimates.
References