Chapter 11

Incomplete-Block Designs

11.1 Introduction

In Chapter 4 we saw that natural blocks may have a fixed size irrespective of the number of treatments. In that chapter we considered block designs where each treatment occurred at least once per block. Here we deal with block designs where the blocks are not large enough to contain all treatments.

Assume that there are b blocks of k plots each, and t treatments each replicated r times. Thus

$$N = bk = tr. \tag{11.1}$$

Also assume that blocks are *incomplete* in the sense that (i) $k < t$ and (ii) no treatment occurs more than once in any block.

Definition For distinct treatments i and j, the *concurrence* λ_{ij} of i and j is the number of blocks which contain both i and j.

Example 11.1 (Concurrence) The four blocks of a design with $t = 6$, $r = 2$, $b = 4$ and $k = 3$ are as follows.

$$\{1,2,3\}, \{1,4,5\}, \{2,4,6\}, \{3,5,6\}.$$

Here $\lambda_{12} = 1$ and $\lambda_{16} = 0$.

This chapter gives a very brief introduction to the topic of incomplete-block designs, covering the classes of design which are most useful in practice. Proofs are deliberately omitted.

11.2 Balance

Definition An incomplete-block design is *balanced* if there is an integer λ such that $\lambda_{ij} = \lambda$ for all distinct treatments i and j.

The name 'balanced incomplete-block design' is often abbreviated to BIBD.

Theorem 11.1 *In a balanced incomplete-block design,*

$$\lambda(t-1) = r(k-1)$$

and therefore $t-1$ divides $r(k-1)$.

Proof Treatment 1 occurs in r blocks, each of which contains $k-1$ other treatments. So the total number of concurrences with treatment 1 is $r(k-1)$. However, this number is the sum of the λ_{1j} for $j \neq 1$, each of which is equal to λ, so the sum is $(t-1)\lambda$. ■

Here are some methods of constructing balanced incomplete-block designs.

Unreduced designs There is one block for each k-subset of $\mathcal{T}$. Thus

$$b = {}^{t}\mathrm{C}_{k} = \frac{t!}{k!\,(t-k)!},$$

which is large unless $k=2$, $k=t-2$ or $k=t-1$. Also,

$$r = \frac{bk}{t} = \frac{(t-1)!}{(k-1)!\,(t-k)!} = {}^{t-1}\mathrm{C}_{k-1}$$

and

$$\lambda = \frac{r(k-1)}{t-1} = \frac{(t-2)!}{(k-2)!\,(t-k)!} = {}^{t-2}\mathrm{C}_{k-2}.$$

For example, the unreduced design for 5 treatments in blocks of size 2 has the following ten blocks.

$$\{1,2\},\ \{1,3\},\ \{1,4\},\ \{1,5\},\ \{2,3\},\ \{2,4\},\ \{2,5\},\ \{3,4\},\ \{3,5\},\ \{4,5\}$$

Cyclic designs from difference sets This method gives designs with $b=t$. Identify the treatments with the integers modulo t. Choose an *initial* block $B = \{i_1, i_2, \ldots, i_k\}$. The other blocks of the *cyclic* design are $B+1$, $B+2$, …, $B+(t-1)$, where

$$B+1 = \{i_1+1, i_2+1, \ldots, i_k+1\},$$

and so on, all arithmetic being modulo t.

$\{1,2,4\}$, $\{2,3,5\}$, $\{3,4,6\}$, $\{4,5,0\}$, $\{5,6,1\}$, $\{6,0,2\}$, $\{0,1,3\}$

Figure 11.1: Cyclic design generated from the initial block $\{1,2,4\}$ modulo 7.

For example, when $t = 7$ the initial block $\{1,2,4\}$ gives the design in Figure 11.1.

Not every choice of initial block will lead to a balanced design. To check whether an intial block B is suitable, construct its *difference table*.

	i_1	i_2	$\dots$	i_k
i_1	0	$i_2 - i_1$	$\dots$	$i_k - i_1$
i_2	$i_1 - i_2$	0	$\dots$	$i_k - i_2$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
i_k	$i_1 - i_k$	$i_2 - i_k$	$\dots$	0

All subtraction in this table is done modulo t. It can be shown that λ_{ij} is equal to the number of occurrences of $i - j$ in the body of the table. Of course, all the diagonal entries are equal to 0. If every non-zero integer modulo t occurs equally often in the non-diagonal entries of the table then B is said to be a *difference set* and the cyclic design is balanced.

The difference table for the initial block in Figure 11.1 is as follows.

	1	2	4
1	0	1	3
2	6	0	2
4	4	5	0

Every non-zero integer modulo 7 occurs once in the body of the table, so the cyclic design should be balanced with $\lambda = 1$. This can be verified directly from Figure 11.1.

Designs from Latin squares If q is a power of a prime number and either $t = q^2$ and $k = q$ or $t = q^2 + q + 1$ and $k = q + 1$, there are constructions of balanced incomplete-block designs which use mutually orthogonal Latin squares of size q. These are given at the end of Section 11.3.

Balanced designs are obviously desirable, and in some sense fair. They are optimal, in a sense to be defined in Section 11.8. However, there may not exist a balanced incomplete-block design for given values of t, r, b and k. Equation (11.1) and Theorem 11.1 give two necessary conditions. A more surprising necessary condition is given by the following theorem.

Theorem 11.2 (Fisher's Inequality) *In a balanced incomplete-block design,* $b \geq t$.

Practical constraints such as costs often force us to have b less than t, so balanced incomplete-block designs are not as useful as they seem at first sight.

11.3 Lattice designs

Definition An incomplete-block design is *resolved* if the blocks are grouped into larger blocks and the large blocks form a complete-block design.

When natural conditions force us to use small blocks, and hence an incomplete-block design, use a resolved design if possible so that large blocks can be used for management. Of course, this is not possible unless k divides t.

In Example 1.5, the strips form a resolved incomplete-block design because the fields are large blocks.

Theorem 11.3 (Bose's Inequality) *In a resolved balanced incomplete-block design,* $b \geq t + r - 1$.

Most resolved designs used in practice are not balanced, because too many blocks would be needed for balance.

A *lattice design* is a special sort of resolved incomplete-block design in which $t = k^2$. The construction method is as follows.

(i) Write the treatments in a $k \times k$ square array.

(ii) For the first large block, the rows of the square are the blocks.

(iii) For the second large block, the columns of the square are the blocks.

(iv) For the third large block (if any), write down a $k \times k$ Latin square and use its letters as the blocks.

(v) For the fourth large block (if any), write down a $k \times k$ Latin square orthogonal to the first one and use its letters as the blocks.

(vi) And so on, using $r - 2$ mutually orthogonal Latin squares.

Example 11.2 (Lattice design) Suppose that $t = 9$ and $k = 3$. We can take the treatment array

1	2	3
4	5	6
7	8	9

.

Then the blocks in the first large block are

$$\{1,2,3\},\ \{4,5,6\},\ \{7,8,9\}$$

and the blocks in the second large block are

$$\{1,4,7\},\ \{2,5,8\},\ \{3,6,9\}.$$

For a third large block we can use the Latin square

A	B	C
B	C	A
C	A	B

,

which gives the blocks

$$\{1,6,8\},\ \{2,4,9\},\ \{3,5,7\}.$$

Finally, for a fourth large block we can use the following Latin square, orthogonal to the previous one:

α	β	γ
γ	α	β
β	γ	α

.

The blocks in the fourth large block are

$$\{1,5,9\},\ \{2,6,7\},\{3,4,8\}.$$

Thus we obtain a design with two, three or four large blocks.

In a lattice design, every concurrence λ_{ij} is equal to 0 or 1. For each treatment i, the number of other treatments j for which $\lambda_{ij} = 1$ is equal to $r(k-1)$.

In resolved designs in general, a treatment contrast $\mathbf{v}$ is said to be *confounded with blocks* in a given large block if v_ω takes a constant value throughout each block in that large block. On the other hand, a treatment contrast $\mathbf{v}$ is said to be *orthogonal to blocks* in a given large block if the sum of the values in $\mathbf{v}$ is zero in every block in that large block. Lattice designs have the property that each treatment contrast that is confounded with blocks in any large block is orthogonal to blocks in every other large block.

If there exist $k+1$ mutually orthogonal $k\times k$ Latin squares then we obtain a lattice design with $k+1$ large blocks. Now $r(k-1) = (k+1)(k-1) = k^2-1 = t-1$, and so the design is balanced with $\lambda = 1$. If k is a prime number or a power of a prime number, the methods in Section 9.2 give $k+1$ mutually orthogonal $k\times k$ Latin squares.

In this case, $k+1$ extra treatments can be added to give a non-resolved balanced incomplete-block design with k^2+k+1 treatments, block size $k+1$ and $\lambda = 1$. Label the new treatments j_1, …, j_{k+1}. Insert treatment j_i into every block in the i-th large block. Then add one further block containing $\{j_1, j_2, \ldots, j_{k+1}\}$.

Example 11.3 (Balanced design) For a balanced incomplete-block design for 13 treatments in 13 blocks of size 4, start with the version of Example 11.2 which has four large blocks. Insert treatment 10 into every block in the first large block; insert treatment 11 into every block in the second large block; insert treatment 12 into every block in the third large block; insert treatment 13 into every block in the fourth large block. Finally, add one new block containing all the new treatments. This gives the following blocks.

$$
\begin{array}{llll}
\{1,2,3,10\}, & \{4,5,6,10\}, & \{7,8,9,10\}, & \\
\{1,4,7,11\}, & \{2,5,8,11\}, & \{3,6,9,11\}, & \\
\{1,6,8,12\}, & \{2,4,9,12\}, & \{3,5,7,12\}, & \\
\{1,5,9,13\}, & \{2,6,7,13\}, & \{3,4,8,13\}, & \{10,11,12,13\}.
\end{array}
$$

11.4 Randomization

To randomize a resolved incomplete-block design, proceed as follows.

(i) Randomize large blocks, that is, randomize the numbering of the large blocks.

(ii) Within each large block independently, randomize blocks; that is, randomly order the blocks within each large block (this will affect which treatments occur in which real block).

(iii) Within each block independently, randomize plots; that is, randomly order the treatments which have been assigned to that block.

Do *not* randomize the treatment labels. As we shall see in Section 11.7, if the design is not balanced then different treatment contrasts have different variances. A good designer chooses the treament labels in such a way that the most important contrasts have lower variance. Randomizing treatment labels would undo the effect of careful matching of real treatments to treatment labels.

A common mistake is to randomize treatment labels independently within each large block. This can have the effect of making the concurrences very unequal, and so making the design much less efficient than it need be, even on average, not just for the important contrasts.

Randomization of unresolved incomplete-block designs is similar, except that the first step is omitted.

(i) Randomize blocks.

(ii) Within each block independently, randomize plots.

Example 11.4 (Example 1.9 continued: Detergents) Here ten housewives each do four washloads in an experiment to compare new detergents. Thus $b = 10$ and $k = 4$. If there are five new detergents then $t = 5$. The obvious balanced incomplete-block design for 5 treatments in blocks of size 4 is the unreduced design, which has 5 blocks. To obtain a balanced design in 10 blocks, simply take two copies of the unreduced design. This gives the systematic design in Table 11.1.

One of the many possible randomized plans for this experiment is shown in Table 11.4. Both blocks and plots have been randomized by the method given at the end of Section 2.2. The stream of random numbers (to three decimal places) used to randomize the blocks is in Table 11.2, while the stream of random digits used to randomize the plots within blocks is in Table 11.3.

Housewife 1

Washload	1	2	3	4
Detergent	1	2	3	4

Housewife 2

Washload	1	2	3	4
Detergent	1	2	3	5

Housewife 3

Washload	1	2	3	4
Detergent	1	2	4	5

Housewife 4

Washload	1	2	3	4
Detergent	1	3	4	5

Housewife 5

Washload	1	2	3	4
Detergent	2	3	4	5

Housewife 6

Washload	1	2	3	4
Detergent	1	2	3	4

Housewife 7

Washload	1	2	3	4
Detergent	1	2	3	5

Housewife 8

Washload	1	2	3	4
Detergent	1	2	4	5

Housewife 9

Washload	1	2	3	4
Detergent	1	3	4	5

Housewife 10

Washload	1	2	3	4
Detergent	2	3	4	5

Table 11.1: Systematic design in Example 11.4

0.653	0.686	0.911	0.402	0.130	0.486	0.678	0.995	0.707	0.808
4	6	9	2	1	3	5	10	7	8

Table 11.2: Stream of random numbers, used to randomize the blocks in Example 11.4

0	5	9	3	|	6	2	3	5	|	0	3	9	5	|	1	1	9	2	9	6	|	9	8	7	6	|	7	2	2	5	6	|	5	9	2	6	|	7	3	8	3	0
1	4	5	3	|	4	1	2	3	|	1	3	5	4	|	1	X	5	2	X	3	|	4	3	2	1	|	5	1	X	4	2	|	3	5	2	4	|	4	3	5	X	2

|	2	0	3	9	|	8	9	1	9	2
|	2	1	3	5	|	4	5	1	X	2

Table 11.3: Stream of random digits, used to randomize the plots within blocks in Example 11.4

Housewife 1

Washload	1	2	3	4
Detergent	1	4	5	3

Housewife 2

Washload	1	2	3	4
Detergent	4	1	2	3

Housewife 3

Washload	1	2	3	4
Detergent	1	3	5	4

Housewife 4

Washload	1	2	3	4
Detergent	1	5	2	3

Housewife 5

Washload	1	2	3	4
Detergent	4	3	2	1

Housewife 6

Washload	1	2	3	4
Detergent	5	1	4	2

Housewife 7

Washload	1	2	3	4
Detergent	3	5	2	4

Housewife 8

Washload	1	2	3	4
Detergent	4	3	5	2

Housewife 9

Washload	1	2	3	4
Detergent	2	1	3	5

Housewife 10

Washload	1	2	3	4
Detergent	4	5	1	2

Table 11.4: Randomized plan in Example 11.4

11.5 Analysis of balanced incomplete-block designs

Assume the fixed-effects model, so that

$$\mathbb{E}(Y_\omega) = \tau_i + \zeta_j \tag{11.2}$$

if plot ω is in block j and receives treatment i; also, $\mathrm{Cov}(\mathbf{Y}) = \xi\mathbf{I}$. As we showed in Section 4.5, we can replace every treatment parameter τ_i by $\tau_i + c$ for some fixed constant c, so long as we replace every block parameter ζ_j by $\zeta_j - c$. The algorithm that follows estimates $\tau_i + c$ where $c = -\sum_{i=1}^{t} \tau_i/t$.

Suppose that $\mathbf{y}$ is the vector of observations from an experiment carried out in an incomplete-block design. To analyse the data by hand, proceed as follows.

(i) Subtract the block mean from every observation in each block, to give a new vector $\mathbf{y}^{(B)}$.

(ii) Calculate treatment means from $\mathbf{y}^{(B)}$.

(iii) Multiply each treatment mean by

$$\frac{t-1}{t}\frac{k}{k-1}$$

to obtain the treatment estimates $\hat{\tau}_i$.

(iv) In $\mathbf{y}$, subtract $\hat{\tau}_i$ from every observation on treatment i, to give a new vector $\mathbf{y}^{(T)}$.

(v) Obtain $\hat{\zeta}_j$, the estimate of the parameter for block j, as the mean of the entries in block j in $\mathbf{y}^{(T)}$.

(vi) In $\mathbf{y}^{(T)}$, subtract $\hat{\zeta}_j$ from every entry in block j, to obtain the vector of residuals.

(vii) Estimate the plots stratum variance by

$$\frac{\text{sum of squares of residuals}}{N-t-b+1}.$$

Theorem 11.4 *If $\lambda_1, \ldots, \lambda_t$ are real numbers such that $\sum_{i=1}^t \lambda_i = 0$ then, under the fixed-effects model, the above procedure gives $\sum \lambda_i \hat{\tau}_i$ as the best linear unbiased estimator of $\sum \lambda_i \tau_i$. Moreover, the variance of this estimator is equal to*

$$\left(\frac{1}{r}\sum_{i=1}^t \lambda_i^2\right)\left(\frac{t-1}{t}\frac{k}{k-1}\right)\xi,$$

where ξ is the plots stratum variance.

Although we shall not prove the whole of Theorem 11.4, we shall show that it gives unbiased estimators of τ_i and ζ_j under the assumption that the parameters in Equation (11.2) have been chosen so that $\sum_{i=1}^t \tau_i = 0$. Suppose that $k = 3$ and that block 1 contains treatments 1, 2 and 3. Then the expected values of the responses in block 1 are $\tau_1+\zeta_1$, $\tau_2+\zeta_1$ and $\tau_3+\zeta_1$, whose mean is $(\tau_1+\tau_2+\tau_3)/3+\zeta_1$. After subtracting this mean, the expected values are $\tau_1-(\tau_1+\tau_2+\tau_3)/3$, $\tau_2-(\tau_1+\tau_2+\tau_3)/3$ and $\tau_3-(\tau_1+\tau_2+\tau_3)/3$. In general, the expected value on a plot ω with treatment i is $\tau_i - {\sum'}_{B(\omega)}\tau_j/k$, where the sum ${\sum'}_{B(\omega)}$ is taken over all treatments in the block $B(\omega)$. Since treatment i occurs r times altogether, and λ times with each other treatment, the mean of the values $\tau_i - {\sum'}_{B(\omega)}\tau_j/k$, over all plots which receive treatment i, is

$$\begin{aligned}
\tau_i - \frac{1}{rk}\left(r\tau_i + \lambda\sum_{j\neq i}\tau_j\right) &= \tau_i - \frac{1}{rk}\left((r-\lambda)\tau_i + \lambda\sum_{j=1}^t \tau_j\right)\\
&= \frac{(rk-r+\lambda)\tau_i}{rk}, \qquad \text{because } \textstyle\sum_{j=1}^t \tau_j = 0,\\
&= \frac{(r(k-1)(t-1)+r(k-1))\tau_i}{rk(t-1)}, \qquad \text{by Theorem 11.1,}\\
&= \frac{(k-1)t}{k(t-1)}\,\tau_i.
\end{aligned}$$

Thus the estimators $\hat{\tau}_i$ are unbiased.

Because $\mathbb{E}(\hat{\tau}_i) = \tau_i$ for all i, the act of subtracting $\hat{\tau}_i$ from $\mathbb{E}(Y_\omega)$ for all plots ω which receive treatment i leaves an expected value of ζ_j for every plot in block j.

Hence the mean of these values is also equal to ζ_j, and so the estimators $\hat{\zeta}_j$ are unbiased.

Note that the foregoing argument does not work for incomplete-block designs which are not balanced.

What about the random-effects model? Here the assumptions are that $\mathbb{E}(Y_\omega) = \tau_i$ if $T(\omega) = i$ and that

$$\mathrm{Cov}(\mathbf{Y}) = \xi_0 P_{V_U} + \xi_B (P_{V_B} - P_{V_U}) + \xi (I - P_{V_B}),$$

where P_{V_U} is the matrix of orthogonal projection onto V_U, which is $N^{-1}\mathbf{J}$, and P_{V_B} is the matrix of orthogonal projection onto V_B, which is equal to $k^{-1}\mathbf{J}_B$. Often we assume that ξ_B is so large relative to ξ that we will base all estimation on the data $(\mathbf{y} - P_{V_B}\mathbf{y})$ orthogonal to blocks; that is, we will begin with the first step of the preceding algorithm.

Theorem 11.5 *If $\lambda_1, \ldots, \lambda_t$ are real numbers such that $\sum_{i=1}^t \lambda_i = 0$ then, under the random-effects model, the preceding procedure gives $\sum \lambda_i \hat{\tau}_i$ as the best linear unbiased estimator of $\sum \lambda_i \tau_i$ which uses only the data $\mathbf{y} - P_{V_B}\mathbf{y}$ orthogonal to blocks. Moreover, the variance of this estimator is equal to*

$$\left(\frac{1}{r}\sum_{i=1}^t \lambda_i^2\right)\left(\frac{t-1}{t}\frac{k}{k-1}\right)\xi,$$

where ξ is the plots stratum variance.

11.6 Efficiency

Definition The *efficiency* for a treatment estimator $\sum \lambda_i \hat{\tau}_i$ in an incomplete-block design Δ relative to a complete-block design Γ with the same values of t and r is

$$\frac{\mathrm{Var}\left(\sum \lambda_i \hat{\tau}_i\right) \text{ in } \Gamma}{\mathrm{Var}\left(\sum \lambda_i \hat{\tau}_i\right) \text{ in } \Delta}.$$

Thus efficiency is large when variance is small.

Let σ^2 be the plots stratum variance in Γ. Then the variance of the estimator $\sum \lambda_i \hat{\tau}_i$ in Γ is equal to $\left[\left(\sum \lambda_i^2\right)/r\right]\sigma^2$. Suppose that the variance of the estimator $\sum \lambda_i \hat{\tau}_i$ in Δ is $v\xi$, where ξ is the plots stratum variance in Δ. Then the efficiency for this estimator is

$$\frac{\left(\sum \lambda_i^2/r\right)\sigma^2}{v\xi} = \frac{\sum \lambda_i^2}{rv} \times \frac{\sigma^2}{\xi}.$$

The right-hand factor, σ^2/ξ, depends only on the variability of the experimental material in the blocks of the two sizes. It is usually unknown, but a reasonable guess at its size may often be made in advance of the experiment. If the blocking is good then ξ should be less than σ^2, so this factor should be greater than 1. The left-hand factor, $\sum \lambda_i^2/(rv)$, depends only on properties of the design Δ; it is called the *efficiency factor* for the treatment contrast $\sum\lambda_i\tau_i$ in the design Δ. It can be shown that efficiency factors are always between 0 and 1.

Theorem 11.4 shows that if Δ is a balanced incomplete-block design then

$$v = \frac{\sum\lambda_i^2}{r} \times \frac{t-1}{t} \times \frac{k}{k-1}$$

and so the efficiency factor for every treatment contrast is equal to

$$\frac{t}{t-1}\frac{k-1}{k}.$$

Since $tk - t < tk - k$, this is less than 1.

Example 11.5 (Comparing block designs) Suppose that we want to do an experiment to compare seven treatments, and that we have enough treatment material for three replicates. We might have a choice between seven blocks of size three, with plots stratum variance ξ, and three blocks of size seven, with plots stratum variance σ^2. In the first case we can use the balanced incomplete-block design in Figure 11.1. Now

$$\frac{t-1}{t} \times \frac{k}{k-1} = \frac{6}{7} \times \frac{3}{2} = \frac{9}{7},$$

and so the variance of each estimator of a difference between two treatments is equal to

$$\frac{2}{r} \times \frac{9}{7} \times \xi = \frac{2}{3} \times \frac{9}{7} \times \xi = \frac{6}{7}\,\xi.$$

In the second case we can use a complete-block design, and the variance of each estimator of a difference between two treatments is equal to $2\sigma^2/r$, which is $2\sigma^2/3$. Thus the incomplete-block design is better if and only if $6\xi/7 < 2\sigma^2/3$; that is, if and only if $\xi < 7\sigma^2/9$.

Sometimes the notion of efficiency and the relative sizes of the plots stratum variances enables us to choose between block designs whose block sizes differ. At other times the block size is fixed and the concept of efficiency enables us to choose between two different designs for the same block size.

11.7 Analysis of lattice designs

Lattice designs also have a method of analysis by hand that is relatively straightforward. For simplicity, we consider only the fixed-effects model: the adaptation to the random-effects model is the same as in Section 11.5.

There are three types of contrast. Recall that, for an equi-replicate design, the constrast corresponding to the linear combination $\sum\lambda_i\tau_i$ is $\sum\lambda_i\mathbf{u}_i$, where $\mathbf{u}_i$ takes the value 1 on ω if $T(\omega) = i$ and the value 0 otherwise.

(i) If a contrast is confounded with blocks in one large block then it cannot be estimated from that large block, because there is no way of distinguishing the effect of the treatment contrast from the effect of the corresponding block contrast. However, it is then orthogonal to blocks in every other large block, so it can be estimated from all the other large blocks in the normal way. There are $r-1$ other large blocks, and so

$$\mathrm{Var}\left(\sum_i \lambda_i \hat{\tau}_i\right) = \frac{\sum \lambda_i^2}{r-1}\,\xi.$$

Thus the efficiency for this contrast is

$$\frac{\left(\sum \lambda_i^2/r\right)\sigma^2}{\left(\sum \lambda_i^2/(r-1)\right)\xi} = \frac{r-1}{r}\frac{\sigma^2}{\xi}$$

and the efficiency factor is $(r-1)/r$.

(ii) If a contrast is orthogonal to blocks in every large block then it can be estimated from every large block in the normal way. Then

$$\mathrm{Var}\left(\sum_i \lambda_i \hat{\tau}_i\right) = \frac{\sum \lambda_i^2}{r}\,\xi$$

and the efficiency factor is equal to 1.

(iii) Every other treatment contrast can be expressed as a sum of contrasts of the first two types. Moreover, these can be chosen to be orthogonal to each other, so their estimators are uncorrelated and the variance of their sum is the sum of their variances.

(iv) The preceding steps give unbiased estimators for the treatment parameters $\hat{\tau}_i$ (up to a constant c). From these, obtain estimators for the block parameters ζ_j and the plots stratum variance ξ just as in Section 11.5.

Example 11.2 revisited (Lattice design) Table 11.5 shows the version of this design with two large blocks. The contrast **v** corresponds to the linear combination $(\tau_1+\tau_2+\tau_3)-(\tau_4+\tau_5+\tau_6)$ while the contrast **w** corresponds to $(2\tau_1-\tau_2-\tau_3)-(2\tau_4-\tau_5-\tau_6)$.

The contrast **v** is confounded with blocks in the first large block. If we tried to estimate $(\tau_1+\tau_2+\tau_3)-(\tau_4+\tau_5+\tau_6)$ from the first large block we would use $Y_1+Y_2+Y_3-Y_4-Y_5-Y_6$, whose expectation is equal to $(\tau_1+\tau_2+\tau_3)-(\tau_4+\tau_5+\tau_6)+3\zeta_1-3\zeta_2$. There is no way to disentangle the τ parameters from the ζ parameters. However, this contrast is orthogonal to blocks in the second large block, so we can use $Y_{10}-Y_{11}+Y_{13}-Y_{14}+Y_{16}-Y_{17}$ to estimate $(\tau_1+\tau_2+\tau_3)-(\tau_4+\tau_5+\tau_6)$. The variance of this estimator is equal to

$$\frac{1^2+1^2+1^2+(-1)^2+(-1)^2+(-1)^2}{1}\,\xi = 6\xi.$$

Large block	block j	Plot ω	Treatment	$\mathbb{E}(Y_\omega)$	$\mathbf{v}$	$\mathbf{w}$	$\mathbf{x}$	$\mathbf{z}$
1	1	1	1	$\tau_1+\zeta_1$	1	2	1	1
1	1	2	2	$\tau_2+\zeta_1$	1	−1	−1	0
1	1	3	3	$\tau_3+\zeta_1$	1	−1	0	−1
1	2	4	4	$\tau_4+\zeta_2$	−1	−2	1	0
1	2	5	5	$\tau_5+\zeta_2$	−1	1	−1	−1
1	2	6	6	$\tau_6+\zeta_2$	−1	1	0	1
1	3	7	7	$\tau_7+\zeta_3$	0	0	1	−1
1	3	8	8	$\tau_8+\zeta_3$	0	0	−1	1
1	3	9	9	$\tau_9+\zeta_3$	0	0	0	0
2	4	10	1	$\tau_1+\zeta_4$	1	2	1	1
2	4	11	4	$\tau_4+\zeta_4$	−1	−2	1	0
2	4	12	7	$\tau_7+\zeta_4$	0	0	1	−1
2	5	13	2	$\tau_2+\zeta_5$	1	−1	−1	0
2	5	14	5	$\tau_5+\zeta_5$	−1	1	−1	−1
2	5	15	8	$\tau_8+\zeta_5$	0	0	−1	1
2	6	16	3	$\tau_3+\zeta_6$	1	−1	0	−1
2	6	17	6	$\tau_6+\zeta_6$	−1	1	0	1
2	6	18	9	$\tau_9+\zeta_6$	0	0	0	0

Table 11.5: Treatment contrasts in Example 11.2

On the other hand, the contrast **w** is orthogonal to blocks in both large blocks, so we use $(2Y_1-Y_2-Y_3-2Y_4+Y_5+Y_6+2Y_{10}-2Y_{11}-Y_{13}+Y_{14}-Y_{16}+Y_{17})/2$ to estimate $(2\tau_1-\tau_2-\tau_3)-(2\tau_4-\tau_5-\tau_6)$. The variance of this estimator is equal to

$$\frac{2^2+(-1)^2+(-1)^2+(-2)^2+1^2+1^2}{2}\,\xi = 6\xi.$$

The sum of these two estimators is used to estimate $3(\tau_1-\tau_4)$. The first estimator does not use any data from the first large block, so it is obviously independent of the part of the second estimator that uses the first large block. On the second large block, the two estimators are orthogonal to each other, so Theorem 2.5 shows that they are uncorrelated. Hence the variance of their sum is equal to $6\xi+6\xi$, and so $\mathrm{Var}(\hat\tau_1-\hat\tau_4)=12\xi/9=4\xi/3$. The efficiency for $\hat\tau_1-\hat\tau_4$ is

$$\frac{(2/2)\sigma^2}{(4/3)\xi}=\frac{3}{4}\frac{\sigma^2}{\xi},$$

while the efficiency factor is $3/4$.

To estimate the difference $\tau_1-\tau_5$ we use the contrasts **v**, **x** and **z** shown in Table 11.5. The contrast **x** is confounded with blocks in the second large block, so we use it in the first large block only to estimate $\tau_1-\tau_2+\tau_4-\tau_5+\tau_7-\tau_8$ with variance 6ξ. The contrast **z** is orthogonal to blocks in both large blocks, so we use it in both large blocks to estimate $\tau_1 - \tau_3 - \tau_5 + \tau_6 - \tau_7 + \tau_8$ with variance 3ξ. The sum of these three estimators gives an estimator for $3(\tau_1 - \tau_5)$ with variance $6\xi + 6\xi + 3\xi$, because the three contrasts are orthogonal to each other in each block. Hence $\mathrm{Var}(\hat{\tau}_1 - \hat{\tau}_5) = 15\xi/9 = 5\xi/3$. The efficiency for $\hat{\tau}_1 - \hat{\tau}_5$ is

$$\frac{(2/2)\sigma^2}{(5/3)\xi} = \frac{3}{5}\frac{\sigma^2}{\xi},$$

while the efficiency factor is $3/5$.

Theorem 11.6 *In a lattice design with replication r and block size k, the estimator $\hat{\tau}_i - \hat{\tau}_j$ has variance*

$$\begin{cases} \dfrac{2(k+1)}{kr}\,\xi & \text{if } \lambda_{ij} = 1 \\[2ex] \dfrac{2(kr-k+r)}{kr(r-1)}\,\xi & \text{if } \lambda_{ij} = 0, \end{cases}$$

where ξ is the plots stratum variance. In particular, in a balanced lattice design ($r = k+1$), every such estimator has variance

$$\frac{2(k+1)}{kr} = \frac{2r}{(r-1)r} = \frac{2}{r-1}.$$

Note that

$$\frac{k+1}{kr} = \frac{(k+1)(r-1)}{kr(r-1)} = \frac{kr-k+r-1}{kr(r-1)} < \frac{kr-k+r}{kr(r-1)},$$

so the variance is bigger if treatments i and j do not occur in a block together.

11.8 Optimality

Ideally we want to choose a design in which the variance of every estimator of the form $\hat{\tau}_i - \hat{\tau}_j$ is a small as possible.

Definition The *overall efficiency factor E* for an incomplete-block design for t treatments replicated r times is defined by

$$\frac{2}{t(t-1)}\sum_{i=1}^{t-1}\sum_{j=i+1}^{t}\mathrm{Var}(\hat{\tau}_i - \hat{\tau}_j) = \frac{2\xi}{rE},$$

where ξ is the plots stratum variance.

Definition An incomplete-block design is *optimal* if it has the largest value of E over all incomplete-block designs with the same values of t, r and k.

The next two results justify our concentration on balanced designs and on lattice designs.

Theorem 11.7 *For an incomplete-block design, the following conditions are equivalent.*

(i) The design is balanced.

(ii) All estimators of the form $\hat{\tau}_i - \hat{\tau}_j$ have the same variance.

(iii) All estimators of the form $\hat{\tau}_i - \hat{\tau}_j$ have the same efficiency fcator.

(iv) The overall efficiency factor satisfies

$$E = \frac{t}{t-1}\frac{k-1}{k}$$

where t is the number of treatments and k is the block size.

Moreover, if the design is not balanced then

$$E < \frac{t}{t-1}\frac{k-1}{k}.$$

Hence balanced incomplete-block designs are optimal.

Theorem 11.8 *Lattice designs are optimal.*

11.9 Supplemented balance

Balanced designs and lattice designs are suitable whether or not the treatments are structured. Incomplete-block designs especially suitable for factorial treatments are given in Chapter 12. The other important treatment structure has one or more control treatments.

Definition An incomplete-block design for a set of treatments which includes one or more control treatments has *supplemented balance* if every control treatment occurs once in each block and the remaining parts of the blocks form a balanced incomplete-block design for the other treatments.

Example 11.4 revisited (Detergents) Suppose that one standard detergent is included in the experiment to compare 5 new detergents. If each housewife does one washload with the standard detergent then she can do three with new detergents. The unreduced design for 5 treatments in blocks of size 3 has 10 blocks. Inserting the standard detergent X into each block gives the following design, which has supplemented balance.

$$\begin{array}{ccccc} \{1,2,3,X\}, & \{1,2,4,X\}, & \{1,2,5,X\}, & \{1,3,4,X\}, & \{1,3,5,X\}, \\ \{1,4,5,X\}, & \{2,3,4,X\}, & \{2,3,5,X\}, & \{2,4,5,X\}, & \{3,4,5,X\} \end{array}$$

Questions for Discussion

11.1 Show that $\{1,3,4,5,9\}$ is a difference set modulo 11.

11.2 For each of the following sets of values of t, b and k, *either* construct a balanced incomplete-block design for t treatments in b blocks of size k *or* prove that no such balanced incomplete-block design exists.

(a) $t = 5 \quad b = 10 \quad k = 3$

(b) $t = 7 \quad b = 14 \quad k = 3$

(c) $t = 10 \quad b = 15 \quad k = 3$

(d) $t = 11 \quad b = 11 \quad k = 5$

(e) $t = 21 \quad b = 14 \quad k = 6$

(f) $t = 25 \quad b = 30 \quad k = 5$

(g) $t = 25 \quad b = 40 \quad k = 5$

(h) $t = 31 \quad b = 31 \quad k = 6$

11.3 Construct a lattice design for 25 treatments in 20 blocks of size 5.

11.4 Suppose that there are 48 experimental units, arranged in 12 blocks of size 4. Construct a suitable design for

(a) 16 treatments;

(b) 9 new treatments and one control treatment.

11.5 Find the efficiency factor for treatment contrasts for the design in Example 11.3.

Suppose that the plots stratum variance for this design is $5/8$ of the plots stratum variance for an alternative design in 4 complete blocks of size 13. Which design is better?

11.6 A horticulture research institute wants to compare 9 methods of treating a certain variety of houseplant while it is being grown in a greenhouse in preparation for the Christmas market. One possibility is to ask twelve small growers to test three treatments each in separate chambers in their greenhouses. A second possibility is to ask three large commercial growers to test nine methods each, also in separate greenhouse chambers. If the plots stratum variance is the same in both cases, which design is more efficient?

Also compare the designs in terms of likely cost, difficulty and representativeness of the results.