CHAPTER 1

General Incomplete Block Design

1.1 INTRODUCTION AND EXAMPLES

One of the basic principles in experimental design is that of reduction of experimental error. We have seen (see Chapters I.9 and I.10) that this can be achieved quite often through the device of blocking. This leads to designs such as randomized complete block designs (Section I.9.2) or Latin square type designs (Chapter I.10). A further reduction can sometimes be achieved by using blocks that contain fewer experimental units than there are treatments.

The problem we shall be discussing then in this and the following chapters is the comparison of a number of treatments using blocks the size of which is less than the number of treatments. Designs of this type are called *incomplete block designs* (see Section I.9.8). They can arise in various ways of which we shall give a few examples.

In the case of field plot experiments, the size of the plot is usually, though by no means always, fairly well determined by experimental and agronomic techniques, and the experimenter usually aims toward a block size of less than 12 plots. If this arbitrary rule is accepted, and we wish to compare 100 varieties or crosses of inbred lines, which is not an uncommon situation in agronomy, we will not be able to accommodate all the varieties in one block. Instead, we might use, for example 10 blocks of 10 plots with different arrangements for each replicate (see Chapter 18).

Quite often a block and consequently its size are determined entirely on biological or physical grounds, as, for example, a litter of mice, a pair of twins, an individual, or a car. In the case of a litter of mice it is reasonable to assume that animals from the same litter are more alike than animals from different litters. The litter size is, of course, restricted and so is, therefore, the block size. Moreover, if one were to use female mice only for a certain investigation, the block size would be even more restricted, say to four or five animals. Hence,

Design and Analysis of Experiments. Volume 2: Advanced Experimental Design
By Klaus Hinkelmann and Oscar Kempthorne
ISBN 0-471-55177-5

comparing more than this number of treatments would require some type of incomplete block design.

Suppose we wish to compare seven treatments, $T_1, T_2, T_3, T_4, T_5, T_6, T_7$, say, using female mice, and suppose we have several litters with four females. We then could use the following incomplete block design, which, as will be explained later, is a balanced incomplete block design:

	Animal			
Litter	**1**	**2**	**3**	**4**
1	T_1	T_4	T_7	T_6
2	T_3	T_6	T_5	T_7
3	T_7	T_1	T_2	T_5
4	T_1	T_2	T_3	T_6
5	T_2	T_7	T_3	T_4
6	T_5	T_3	T_4	T_1
7	T_2	T_4	T_5	T_6

Notice that with this arrangement every treatment is replicated four times, and every pair of treatments occurs together twice in the same block; for example, T_1 and T_2 occur together in blocks 3 and 4.

Many sociological and psychological studies have been done on twins because they are "alike" in many respects. If they constitute a block, then the block size is obviously two. A number of incomplete block designs are available for this type of situation, for example, Kempthorne (1953) and Zoellner and Kempthorne (1954).

Blocks of size two arise also in some medical studies, when a patient is considered to be a block and his eyes or ears or legs are the experimental units.

With regard to a car being a block, this may occur if we wish to compare brands of tires, using the wheels as the experimental units. In this case one may also wish to take the effect of position of the wheels into account. This then leads to an incomplete design with two-way elimination of heterogeneity (see Chapters 6 and I.10).

These few examples should give the reader some idea why and how the need for incomplete block designs arises quite naturally in different types of research. For a given situation it will then be necessary to select the appropriate design from the catalogue of available designs. We shall discuss these different types of designs in more detail in the following chapters along with the appropriate analysis.

Before doing so, however, it seems appropriate to trace the early history and development of incomplete block designs. This development has been a remarkable achievement, and the reader will undoubtedly realize throughout the next chapters that the concept of incomplete block designs is fundamental to the understanding of experimental design as it is known today.

The origins of incomplete block designs go back to Yates (1936a) who introduced the concept of balanced incomplete block designs and their analysis utilizing both intra- and interblock information (Yates, 1940a). Other incomplete block designs were also proposed by Yates (1936b, 1937a, 1940b), who referred to these designs as quasi-factorial or lattice designs. Further contributions in the early history of incomplete block designs were made by Bose (1939, 1942) and Fisher (1940) concerning the structure and construction of balanced incomplete block designs. The notion of balanced incomplete block design was generalized to that of partially balanced incomplete block designs by Bose and Nair (1939), which encompass some of the lattice designs introduced earlier by Yates. Further extensions of the balanced incomplete block designs and lattice designs were made by Youden (1940) and Harshbarger (1947), respectively, by introducing balanced incomplete block designs for eliminating heterogeneity in two directions (generalizing the concept of the Latin square design) and rectangular lattices some of which are more general designs than partially balanced incomplete block designs. After this there has been a very rapid development in this area of experimental design, and we shall comment on many results more specifically in the following chapters.

1.2 GENERAL REMARKS ON THE ANALYSIS OF INCOMPLETE BLOCK DESIGNS

The analysis of incomplete block designs is different from the analysis of complete block designs in that comparisons among treatment effects and comparisons among block effects are no longer orthogonal to each other (see Section I.7.3). This is referred to usually by simply saying that treatments and blocks are not orthogonal. This nonorthogonality leads to an analysis analogous to that of the two-way classification with unequal subclass numbers. However, this is only partly true and applies only to the analysis that has come to be known as the *intrablock analysis*.

The name of the analysis is derived from the fact that contrasts in the treatment effects are estimated as linear combinations of comparisons of observations in the same block. In this way the block effects are eliminated and the estimates are functions of treatment effects and error (intrablock error) only. Coupled with the theory of least squares and the Gauss–Markov theorem (see I.4.16.2), this procedure will give rise to the best linear unbiased intrablock estimators for treatment comparisons. Historically, this has been the method first used for analyzing incomplete block designs (Yates, 1936a). We shall derive the intrablock analysis in Section 1.3.

Based upon considerations of efficiency, Yates (1939) argued that the intrablock analysis ignores part of the information about treatment comparisons, namely that information contained in the comparison of block totals. This analysis has been called *recovery of interblock information* or *interblock analysis*.

Yates (1939, 1940a) showed for certain types of lattice designs and for the balanced incomplete block design how these two types of analyses can be combined to yield more efficient estimators of treatment comparisons. Nair (1944) extended these results to partially balanced incomplete block designs, and Rao (1947a) gave the analysis for any incomplete block design showing the similarity between the intrablock analysis and the combined intra- and interblock analysis.

The intrablock analysis, as it is usually presented, is best understood by assuming that the block effects in the underlying linear model are fixed effects. But for the recovery of interblock information the block effects are then considered to be random effects. This leads sometimes to confusion with regard to the assumptions in the combined analysis, although it should be clear from the previous remark that then the block effects have to be considered random effects for *both* the intra- and interblock analysis. To emphasize it again, we can talk about intrablock analysis under the assumption of either fixed or random block effects. In the first case ordinary least squares (OLS) will lead to best linear unbiased estimators for treatment contrasts. This will, at least theoretically, not be true in the second case, which is the reason for considering the interblock information in the first place and using the Aitken equation (see I.4.16.2), which is also referred to as *generalized* (*weighted*) *least squares*.

We shall now derive the intrablock analysis (Section 1.3), the interblock analysis (Section 1.7), and the combined analysis (Section 1.8) for the general incomplete block design. Special cases will then be considered in the following chapters.

1.3 THE INTRABLOCK ANALYSIS

1.3.1 Notation and Model

Suppose we have t treatments replicated $r_1, r_2, \ldots, r_t$ times, respectively, and b blocks with $k_1, k_2, \ldots, k_b$ units, respectively. We then have

$$\sum_{i=1}^{t} r_i = \sum_{j=1}^{b} k_j = n$$

where n is the total number of observations.

Following the derivation of a linear model for observations from a randomized complete block design (RCBD), using the assumption of additivity in the broad sense (see Sections I.9.2.2 and I.9.2.6), an appropriate linear model for observations from an incomplete block design is

$$y_{ij\ell} = \mu + \tau_i + \beta_j + e_{ij\ell} \tag{1.1}$$

($i = 1, 2, \ldots, t;\ \ j = 1, 2, \ldots, b;\ \ \ell = 0, 1, \ldots, n_{ij}$), where τ_i is the effect of the ith treatment, β_j the effect of the jth block, and $e_{ij\ell}$ the error associated with the observation $y_{ij\ell}$. As usual, the $e_{ij\ell}$ contain both experimental and observational (sampling) error, that is, using notation established in Volume 1,

$$e_{ij\ell} = \epsilon_{ij\ell} + \eta_{ij\ell}$$

with $\epsilon_{ij\ell}$ representing experimental error and $\eta_{ij\ell}$ representing observational error. Also, based on previous derivations (see I.6.3.4), we can treat the $e_{ij\ell}$ as i.i.d. random variables with mean zero and variance $\sigma_e^2 = \sigma_\epsilon^2 + \sigma_\eta^2$. Note that because n_{ij}, the elements of the incidence matrix $\boldsymbol{N}$, may be zero, not all treatments occur in each block which is, of course, the definition of an incomplete block design.

Model (1.1) can also be written in matrix notation as

$$\boldsymbol{y} = \mu\boldsymbol{\Im} + \boldsymbol{X}_\tau\boldsymbol{\tau} + \boldsymbol{X}_\beta\boldsymbol{\beta} + \boldsymbol{e} \tag{1.2}$$

where $\boldsymbol{\Im}$ is a column vector consisting of n unity elements, $\boldsymbol{X}_\beta$ is the observation-block incidence matrix

$$\boldsymbol{X}_\beta = \begin{bmatrix} \boldsymbol{\Im}_{k_1} & & & \\ & \boldsymbol{\Im}_{k_2} & & \\ & & \ddots & \\ & & & \boldsymbol{\Im}_{k_b} \end{bmatrix}$$

with $\boldsymbol{\Im}_{k_j}$ denoting a column vector of k_j unity elements $(j = 1, 2, \ldots, b)$ and

$$\boldsymbol{X}_\tau = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_t)$$

is the observation-treatment incidence matrix, where $\boldsymbol{x}_i$ is a column vector with r_i unity elements and $(n - r_i)$ zero elements such that $\boldsymbol{x}_i'\boldsymbol{x}_i = r_i$ and $\boldsymbol{x}_i'\boldsymbol{x}_{i'} = 0$ for $i \neq i'(i, i' = 1, 2, \ldots, t)$.

1.3.2 Normal and Reduced Normal Equations

The normal equations (NE) for μ, τ_i, and β_j are then

$$\begin{aligned} n\widehat{\mu} + \sum_{i=1}^{t} r_i\widehat{\tau}_i + \sum_{j=1}^{b} k_j\widehat{\beta}_j &= G \\ r_i\widehat{\mu} + r_i\widehat{\tau}_i + \sum_{j=1}^{b} n_{ij}\widehat{\beta}_j &= T_i \qquad (i = 1, 2, \ldots, t) \\ k_j\widehat{\mu} + \sum_{i=1}^{t} n_{ij}\widehat{\tau}_i + k_j\widehat{\beta}_j &= B_j \qquad (j = 1, 2, \ldots, b) \end{aligned} \tag{1.3}$$

where

$$T_i = \sum_{j\ell} y_{ij\ell} = i\text{th treatment total}$$

$$B_j = \sum_{i\ell} y_{ij\ell} = j\text{th block total}$$

$$G = \sum_i T_i = \sum_j B_j = \text{overall total}$$

Equations (1.3) can be written in matrix notation as

$$\begin{pmatrix} \mathfrak{I}'_n \mathfrak{I}_n & \mathfrak{I}'_n \boldsymbol{X}_\tau & \mathfrak{I}'_n \boldsymbol{X}_\beta \\ \boldsymbol{X}'_\tau \mathfrak{I}_n & \boldsymbol{X}'_\tau \boldsymbol{X}_\tau & \boldsymbol{X}'_\tau \boldsymbol{X}_\beta \\ \boldsymbol{X}'_\beta \mathfrak{I}_n & \boldsymbol{X}'_\beta \boldsymbol{X}_\tau & \boldsymbol{X}'_\beta \boldsymbol{X}_\beta \end{pmatrix} \begin{pmatrix} \widehat{\mu} \\ \widehat{\boldsymbol{\tau}} \\ \widehat{\boldsymbol{\beta}} \end{pmatrix} = \begin{pmatrix} \mathfrak{I}'_n \boldsymbol{y} \\ \boldsymbol{X}'_\tau \boldsymbol{y} \\ \boldsymbol{X}'_\beta \boldsymbol{y} \end{pmatrix} \tag{1.4}$$

which, using the properties of $\mathfrak{I}, \mathbf{X}_\tau, \mathbf{X}_\beta$, can be written as

$$\begin{bmatrix} \mathfrak{I}'_n \mathfrak{I}_n & \mathfrak{I}'_\tau \boldsymbol{R} & \mathfrak{I}'_b \boldsymbol{K} \\ \boldsymbol{R} \mathfrak{I}_t & \boldsymbol{R} & \boldsymbol{N} \\ \boldsymbol{K} \mathfrak{I}_b & \boldsymbol{N}' & \boldsymbol{K} \end{bmatrix} \cdot \begin{bmatrix} \widehat{\mu} \\ \widehat{\boldsymbol{\tau}} \\ \widehat{\boldsymbol{\beta}} \end{bmatrix} = \begin{bmatrix} G \\ \boldsymbol{T} \\ \boldsymbol{B} \end{bmatrix} \tag{1.5}$$

where

$$\begin{aligned} \boldsymbol{R} &= \operatorname{diag}(r_i) && t \times t \\ \boldsymbol{K} &= \operatorname{diag}(k_j) && b \times b \\ \boldsymbol{N} &= (n_{ij}) && t \times b \quad \text{(the incidence matrix)} \\ \boldsymbol{T}' &= (T_1, T_2, \ldots, T_t) \\ \boldsymbol{B}' &= (B_1, B_2, \ldots, B_b) \\ \boldsymbol{\tau}' &= (\tau_1, \tau_2, \ldots, \tau_t) \\ \boldsymbol{\beta}' &= (\beta_1, \beta_2, \ldots, \beta_b) \end{aligned}$$

and the $\mathfrak{I}$'s are column vectors of unity elements with dimensions indicated by the subscripts. From the third set of equations in (1.5) we obtain

$$\widehat{\mu} \mathfrak{I}_b + \widehat{\boldsymbol{\beta}} = \boldsymbol{K}^{-1}(\boldsymbol{B} - \boldsymbol{N}' \widehat{\boldsymbol{\tau}}) \tag{1.6}$$

Substituting (1.6) into the second set of (1.5), which can also be expressed as $\boldsymbol{N}\boldsymbol{\Im}_b\widehat{\mu} + \boldsymbol{N}\widehat{\boldsymbol{\beta}} + \boldsymbol{R}\widehat{\boldsymbol{\tau}} = \boldsymbol{T}$ (since $\boldsymbol{N}\boldsymbol{\Im}_b = \boldsymbol{R}\boldsymbol{\Im}_t$), leads to the *reduced normal equations* (RNE) (see Section I.4.7.1) for $\boldsymbol{\tau}$

$$(\boldsymbol{R} - \boldsymbol{N}\boldsymbol{K}^{-1}\boldsymbol{N}')\widehat{\boldsymbol{\tau}} = \boldsymbol{T} - \boldsymbol{N}\boldsymbol{K}^{-1}\boldsymbol{B} \tag{1.7}$$

Standard notation for (1.7) is

$$\boldsymbol{C}\widehat{\boldsymbol{\tau}} = \boldsymbol{Q} \tag{1.8}$$

where

$$\boldsymbol{C} = \boldsymbol{R} - \boldsymbol{N}\boldsymbol{K}^{-1}\boldsymbol{N}' \tag{1.9}$$

and

$$\boldsymbol{Q} = \boldsymbol{T} - \boldsymbol{N}\boldsymbol{K}^{-1}\boldsymbol{B} \tag{1.10}$$

the (i, i') element of $\boldsymbol{C}$ being

$$c_{ii'} = \delta_{ii'} r_i - \sum_{j=1}^{b} \frac{n_{ij} n_{i'j}}{k_j}$$

with $\delta_{ii'} = 1$ for $i = i'$ and = 0 otherwise, and the ith element of $\boldsymbol{Q}$ being

$$Q_i = T_i - \sum_{j=1}^{b} \frac{n_{ij} B_j}{k_j}$$

And Q_i is called the ith *adjusted treatment total*, the adjustment being due to the fact that the treatments do not occur the same number of times in the blocks.

1.3.3 The *C* Matrix and Estimable Functions

We note that the matrix $\boldsymbol{C}$ of (1.9) is determined entirely by the specific design, that is, by the incidence matrix $\boldsymbol{N}$. It is, therefore, referred to as *the* $\boldsymbol{C}$ matrix (sometimes also as the *information matrix*) of that design. The $\boldsymbol{C}$ matrix is symmetric, and the elements in any row or any column of $\boldsymbol{C}$ add to zero, that is, $\boldsymbol{C}\boldsymbol{\Im} = \boldsymbol{0}$, which implies that $r(\boldsymbol{C}) = \text{rank}(\boldsymbol{C}) \leq t - 1$. Therefore, $\boldsymbol{C}$ does not have an inverse and hence (1.8) cannot be solved uniquely. Instead we write a solution to (1.8) as

$$\widehat{\boldsymbol{\tau}} = \boldsymbol{C}^{-}\boldsymbol{Q} \tag{1.11}$$

where $\boldsymbol{C}^{-}$ is a generalized inverse for $\boldsymbol{C}$ (see Section 1.3.4).

If we write $\boldsymbol{C} = (\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_t)$, where $\boldsymbol{c}_i$ is the ith column of $\boldsymbol{C}$, then the set of linear functions

$$\{\boldsymbol{c}_i'\boldsymbol{\tau}, i = 1, 2, \ldots, t\}$$

which span the totality of estimable functions of the treatment effects, has dimensionality $r(\boldsymbol{C})$. Let $\boldsymbol{c}'\boldsymbol{\tau}$ be an estimable function and $\boldsymbol{c}'\widehat{\boldsymbol{\tau}}$ its estimator, with $\widehat{\boldsymbol{\tau}}$ from (1.11). Then

$$\begin{aligned} E(\boldsymbol{c}'\widehat{\boldsymbol{\tau}}) &= E\left(\boldsymbol{c}'\boldsymbol{C}^-\boldsymbol{Q}\right) \\ &= \boldsymbol{c}'\boldsymbol{C}^- E(\boldsymbol{Q}) \\ &= \boldsymbol{c}'\boldsymbol{C}^-\boldsymbol{C}\boldsymbol{\tau} \end{aligned}$$

For $\boldsymbol{c}'\widehat{\boldsymbol{\tau}}$ to be an unbiased estimator for $\boldsymbol{c}'\boldsymbol{\tau}$ for any $\boldsymbol{\tau}$, we then must have

$$\boldsymbol{c}'\boldsymbol{C}^-\boldsymbol{C} = \boldsymbol{c}' \tag{1.12}$$

Since $\boldsymbol{C}\mathfrak{I} = 0$, it follows from (1.12) that $\boldsymbol{c}'\mathfrak{I} = 0$. Hence, only treatment contrasts are estimable. If $r(\boldsymbol{C}) = t - 1$, then all treatment contrasts are estimable. In particular, all differences $\tau_i - \tau_{i'} (i \neq i')$ are estimable, there being $t - 1$ linearly independent estimable functions of this type. Then the design is called a *connected design* (see also Section I.4.13.3).

1.3.4 Solving the Reduced Normal Equations

In what follows we shall assume that the design is connected; that is, $r(\boldsymbol{C}) = t - 1$. This means that $\boldsymbol{C}$ has $t - 1$ nonzero (positive) eigenvalues and one zero eigenvalue. From

$$\boldsymbol{C}\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \boldsymbol{0} = 0\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

it follows then that $(1, 1, \ldots, 1)'$ is an eigenvector corresponding to the zero eigenvalue. If we denote the nonzero eigenvalues of $\boldsymbol{C}$ by $d_1, d_2, \ldots, d_{t-1}$ and the corresponding eigenvectors by $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_{t-1}$ with $\boldsymbol{\xi}_i'\boldsymbol{\xi}_i = 1 \quad (i = 1, 2, \ldots, t-1)$ and $\boldsymbol{\xi}_i'\boldsymbol{\xi}_{i'} = 0 (i \neq i')$, then we can write $\boldsymbol{C}$ in its spectral decomposition as

$$\boldsymbol{C} = \sum_{i=1}^{t-1} d_i \boldsymbol{\xi}_i \boldsymbol{\xi}_i' \tag{1.13}$$

or with $d_t = 0$ and $\boldsymbol{\xi}_t' = 1/\sqrt{t}(1, 1, \ldots, 1)$, alternatively as

$$\boldsymbol{C} = \sum_{i=1}^{t} d_i \boldsymbol{\xi}_i \boldsymbol{\xi}_i' \tag{1.14}$$

We note that $\boldsymbol{\xi}_t' \boldsymbol{\xi}_t = 1$ and $\boldsymbol{\xi}_i' \boldsymbol{\xi}_t = 0$ for $i = 1, 2, \ldots, t-1$.

We now return to (1.8) and consider a solution to these equations of the form given by (1.11). Although there are many methods of finding generalized inverses, we shall consider here one particular method, which is most useful in connection with incomplete block designs, especially balanced and partially balanced incomplete block designs (see following chapters). This method is based on the following theorem, which is essentially due to Shah (1959).

Theorem 1.1 Let $\boldsymbol{C}$ be a $t \times t$ matrix as given by (1.9) with $r(\boldsymbol{C}) = t - 1$. Then $\widetilde{\boldsymbol{C}} = \boldsymbol{C} + a\mathfrak{I}\mathfrak{I}'$, where $a \neq 0$ is a real number, admits an inverse $\widetilde{\boldsymbol{C}}^{-1}$, and $\widetilde{\boldsymbol{C}}^{-1}$ is a generalized inverse for $\boldsymbol{C}$.

Proof

(a) We can rewrite $\widetilde{C}$ as

$$\widetilde{\boldsymbol{C}} = \boldsymbol{C} + a\mathfrak{I}\mathfrak{I}' = \boldsymbol{C} + a \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} (1, 1, \ldots, 1) = \boldsymbol{C} + at\, \boldsymbol{\xi}_t \boldsymbol{\xi}_t'$$

and because of (1.13)

$$\widetilde{\boldsymbol{C}} = \sum_{i=1}^{t-1} d_i \boldsymbol{\xi}_i \boldsymbol{\xi}_i' + at\, \boldsymbol{\xi}_t \boldsymbol{\xi}_t' \tag{1.15}$$

Clearly, $\widetilde{\boldsymbol{C}}$ has nonzero roots $d_1, d_2, \ldots, d_{t-1}, d_t = at$ and hence is nonsingular. Then

$$\widetilde{\boldsymbol{C}}^{-1} = \sum_{i=1}^{t-1} \frac{1}{d_i} \boldsymbol{\xi}_i \boldsymbol{\xi}_i' + \frac{1}{at} \boldsymbol{\xi}_t \boldsymbol{\xi}_t' \tag{1.16}$$

(b) To show that $\widetilde{\boldsymbol{C}}^{-1} = \boldsymbol{C}^{-}$ we consider $\boldsymbol{C}\widetilde{\boldsymbol{C}}^{-1}\boldsymbol{C}$. From (1.13), (1.15), and (1.16) we have

$$\widetilde{\boldsymbol{C}}\widetilde{\boldsymbol{C}}^{-1} = \boldsymbol{I} = \sum_{i=1}^{t} \boldsymbol{\xi}_i \boldsymbol{\xi}_i'$$

and

$$C\widetilde{C}^{-1} = \sum_{i=1}^{t-1} \boldsymbol{\xi}_i \boldsymbol{\xi}_i' = I - \boldsymbol{\xi}_t \boldsymbol{\xi}_t' = I - \frac{1}{t}\mathfrak{I}\mathfrak{I}' \tag{1.17}$$

$$C\widetilde{C}^{-1}C = C$$

which implies

$$\widetilde{C}^{-1} = C^{-}$$

We remark here already that determining C^{-} for the designs in the following chapters will be based on (1.17) rather than on (1.14).

Substituting $\widetilde{C}^{-1}$ into (1.13) then yields a solution of the RNE (1.8); that is,

$$\widehat{\boldsymbol{\tau}} = \widetilde{C}^{-1} Q \tag{1.18}$$

We note that because of (1.8) and (1.16)

$$\begin{aligned} E(\widehat{\boldsymbol{\tau}}) &= E\left(\widetilde{C}^{-1} Q\right) \\ &= \widetilde{C}^{-1} E(Q) \\ &= \widetilde{C}^{-1} E(C\widehat{\boldsymbol{\tau}}) \\ &= \widetilde{C}^{-1} C\boldsymbol{\tau} \\ &= \left(I - \frac{1}{t}\mathfrak{I}\mathfrak{I}'\right)\boldsymbol{\tau} \\ &= \begin{bmatrix} \tau_1 - \overline{\tau} \\ \tau_2 - \overline{\tau} \\ \vdots \\ \tau_t - \overline{\tau} \end{bmatrix} \end{aligned}$$

with $\overline{\tau} = 1/t \sum_i \tau_i$; that is, $E(\widehat{\boldsymbol{\tau}})$ is the same as if we had obtained a generalized inverse of C by imposing the condition $\sum_i \widehat{\tau_i} = 0$. □

1.3.5 Estimable Functions of Treatment Effects

We know from the Gauss–Markov theorem (see Section I.4.16.2) that for any linear estimable function of the treatment effects, say $c'\boldsymbol{\tau}$,

$$E(c'\widehat{\boldsymbol{\tau}}) = c'\boldsymbol{\tau} \tag{1.19}$$

is independent of the solution to the NE (see Section I.4.4.4). We have further

$$\mathrm{var}(c'\widehat{\boldsymbol{\tau}}) = c'\,\widetilde{C}^{-1}\,c\,\sigma_e^2 \tag{1.20}$$

with a corresponding result (but same numerical value) for any other solution obtained, using an available software package (see Section 1.14). We shall elaborate on this point briefly.

Let us rewrite model (1.2) as

$$\begin{aligned}\boldsymbol{y} &= \mu\boldsymbol{\Im} + \boldsymbol{X}_\beta\boldsymbol{\beta} + \boldsymbol{X}_\tau\boldsymbol{\tau} + \boldsymbol{e} \\ &= (\boldsymbol{\Im}\ \boldsymbol{X}_\beta\ \boldsymbol{X}_\tau)\begin{pmatrix}\boldsymbol{\mu}\\ \boldsymbol{\beta}\\ \boldsymbol{\tau}\end{pmatrix} + \boldsymbol{e} \\ &\equiv \boldsymbol{X\Theta} + \boldsymbol{e} \qquad (1.21)\end{aligned}$$

with

$$\boldsymbol{X} = (\boldsymbol{\Im} : \boldsymbol{X}_\beta : \boldsymbol{X}_\tau) \qquad (1.22)$$

and

$$\boldsymbol{\Theta}' = (\mu, \boldsymbol{\beta}', \boldsymbol{\tau}')$$

The NE for model (1.21) are

$$\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\Theta}^* = \boldsymbol{X}'\boldsymbol{y} \qquad (1.23)$$

A solution to (1.23) is given by, say,

$$\boldsymbol{\Theta}^* = (\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{y}$$

for some $(\boldsymbol{X}'\boldsymbol{X})^-$. Now $(\boldsymbol{X}'\boldsymbol{X})^-$ is a $(1+b+t)\times(1+b+t)$ matrix that we can partition conformably, using the form of $\boldsymbol{X}$ as given in (1.22), as

$$(\boldsymbol{X}'\boldsymbol{X})^- = \begin{pmatrix}\boldsymbol{A}_{\mu\mu} & \boldsymbol{A}_{\mu\beta} & \boldsymbol{A}_{\mu\tau}\\ \boldsymbol{A}'_{\mu\beta} & \boldsymbol{A}_{\beta\beta} & \boldsymbol{A}_{\beta\tau}\\ \boldsymbol{A}'_{\mu\tau} & \boldsymbol{A}'_{\beta\tau} & \boldsymbol{A}_{\tau\tau}\end{pmatrix} \qquad (1.24)$$

Here, $\boldsymbol{A}_{\tau\tau}$ is a $t\times t$ matrix that serves as the variance–covariance matrix for obtaining

$$\mathrm{var}(\boldsymbol{c}'\boldsymbol{\tau}^*) = \boldsymbol{c}'\,\boldsymbol{A}_{\tau\tau}\,\boldsymbol{c}\,\sigma_e^2 \qquad (1.25)$$

For any estimable function $\boldsymbol{c}'\boldsymbol{\tau}$ we have $\boldsymbol{c}'\widehat{\boldsymbol{\tau}} = \boldsymbol{c}'\boldsymbol{\tau}^*$ and also the numerical values for (1.20) and (1.25) are the same. If we denote the (i, i') element of $\widehat{\boldsymbol{C}}^{-1}$ by $c^{ii'}$ and the corresponding element of $\boldsymbol{A}_{\tau\tau}$ in (1.24) by $a^{ii'}$, then we have, for example, for $\boldsymbol{c}'\boldsymbol{\tau} = \tau_i - \tau_{i'}$

$$\mathrm{var}\,(\widehat{\tau}_i - \widehat{\tau}_{i'}) = \left(c^{ii} - 2c^{ii'} + c^{i'i'}\right)\sigma_e^2 = \left(a^{ii} - 2a^{ii'} + a^{i'i'}\right)\sigma_e^2 \qquad (1.26)$$

For a numerical example and illustration of computational aspects, see Section 1.13.

1.3.6 Analyses of Variance

It follows from general principles (see Section I.4.7.1) that the two forms of analysis of variance are as given in Tables 1.1 and 1.2. We shall henceforth refer to the analysis of variance in Table 1.1 as the treatment-after-block ANOVA or T | B-ANOVA as it is associated with the ordered model

$$\boldsymbol{y} = \mu\boldsymbol{\Im} + \boldsymbol{X}_\beta\boldsymbol{\beta} + \boldsymbol{X}_\tau\boldsymbol{\tau} + \boldsymbol{e}$$

whereas the analysis of variance in Table 1.2 is associated with the ordered model

$$\boldsymbol{y} = \boldsymbol{\mu}\boldsymbol{\Im} + \boldsymbol{X}_\tau\boldsymbol{\tau} + \boldsymbol{X}_\beta\boldsymbol{\beta} + \boldsymbol{e}$$

and hence shall be referred to as the block-after-treatment ANOVA or B | T-ANOVA. To indicate precisely the sources of variation and the associated sums of squares, we use the notation developed in Section I.4.7.2 for the general case as it applies to the special case of the linear model for the incomplete block

Table 1.1 T|B-ANOVA for Incomplete Block Design

Source	d.f.[a]	SS	E(MS)
$\boldsymbol{X}_\beta \mid \boldsymbol{\Im}$	$b-1$	$\sum_{j=1}^{b} \frac{B_j^2}{k_j} - \frac{G^2}{n}$	
$\boldsymbol{X}_\tau \mid \boldsymbol{\Im}, \boldsymbol{X}_\beta$	$t-1$	$\sum_{i=1}^{t} \widehat{\tau}_i Q_i$	$\sigma_e^2 + \frac{\boldsymbol{\tau}'\boldsymbol{C}\boldsymbol{\tau}}{t-1}$
$\boldsymbol{I} \mid \boldsymbol{\Im}, \boldsymbol{X}_\beta, \boldsymbol{X}_\tau$	$n-b-t+1$	Difference	σ_e^2
Total	$n-1$	$\sum_{ij\ell} y_{ij\ell}^2 - \frac{G^2}{n}$	

[a] d.f. = degrees of freedom.

Table 1.2 B|T-ANOVA for Incomplete Block Design

Source	d.f.	SS
$\boldsymbol{X}_\tau \mid \boldsymbol{\Im}$	$t-1$	$\sum_{i=1}^{t} \frac{T_i^2}{r_i} - \frac{G^2}{n}$
$\boldsymbol{X}_\beta \mid \boldsymbol{\Im}, \boldsymbol{X}_\tau$	$b-1$	Difference
$\boldsymbol{I} \mid \boldsymbol{\Im}, \boldsymbol{X}_\beta, \boldsymbol{X}_\tau$	$n-b-t+1$	From Table 1.1
Total	$n-1$	$\sum_{ij\ell} y_{ij\ell}^2 - \frac{G^2}{n}$

design, thereby avoiding the commonly used but not always clearly understood terms *blocks ignoring treatments* for $(\boldsymbol{X}_\beta \mid \mathfrak{I})$, *treatments eliminating blocks* for $(\boldsymbol{X}_\tau \mid \mathfrak{I}, \boldsymbol{X}_\beta)$, and *blocks eliminating treatments* for $(\boldsymbol{X}_\beta \mid \mathfrak{I}, \boldsymbol{X}_\tau)$.

The T | B-ANOVA follows naturally from the development of the RNE for the treatment effects. It is the appropriate ANOVA for the intrablock analysis as it allows to test the hypothesis

$$H_0\colon \tau_1 = \tau_2 = \cdots = \tau_t$$

by means of the (approximate) F test (see I.9.2.5)

$$F = \frac{\mathrm{SS}(\boldsymbol{X}_\tau \mid \mathfrak{I}, \boldsymbol{X}_\beta)/(t-1)}{\mathrm{SS}(\boldsymbol{I} \mid \mathfrak{I}, \boldsymbol{X}_\beta, \boldsymbol{X}_\tau)/(n-b-t+1)} \tag{1.27}$$

Also MS(Error)= $\mathrm{SS}(\boldsymbol{I} \mid \mathfrak{I}, \boldsymbol{X}_\beta, \boldsymbol{X}_\tau)/(n-b-t+1)$ is an estimator for σ_e^2 to be used for estimating $\mathrm{var}(\boldsymbol{c}'\widehat{\boldsymbol{\tau}})$ of (1.20).

The usefulness of the B | T-ANOVA in Table 1.2 will become apparent when we discuss specific aspects of the combined intra- and interblock analysis in Section 1.10. At this point we just mention that $\mathrm{SS}(\boldsymbol{X}_\beta \mid, \mathfrak{I}, \boldsymbol{X}_\tau)$ could have been obtained from the RNE for block effects. Computationally, however, it is more convenient to use the fact that $\mathrm{SS}(\boldsymbol{I} \mid \mathfrak{I}, \boldsymbol{X}_\beta, \boldsymbol{X}_\tau) = \mathrm{SS}(\boldsymbol{I} \mid \mathfrak{I}, \boldsymbol{X}_\tau, \boldsymbol{X}_\beta)$ and then obtain $\mathrm{SS}(\boldsymbol{X}_\beta \mid \mathfrak{I}, \boldsymbol{X}_\tau)$ by subtraction.

Details of computational procedures using SAS PROC GLM and SAS PROC Mixed (SAS1999–2000) will be described in Section 1.14.

1.4 INCOMPLETE DESIGNS WITH VARIABLE BLOCK SIZE

In the previous section we discussed the intrablock analysis of the general incomplete block design; that is, a design with possibly variable block size and possibly variable number of replications. Although most designed experiments use blocks of equal size, k say, there exist, however, experimental situations where blocks of unequal size arise quite naturally. We shall distinguish between two reasons why this can happen and why caution may have to be exercised before the analysis as outlined in the previous section can be used:

1. As pointed out by Pearce (1964, p. 699):

> With much biological material there are natural units that can be used as blocks and they contain plots to a number not under the control of the experimenter. Thus, the number of animals in a litter or the number of blossoms in a truss probably vary only within close limits.

2. Although an experiment may have been set up using a proper design, that is, a design with equal block size, missing plots due to accidents during the course of investigation will leave one for purpose of analysis with a design of variable block size.

In both cases there are two alternatives to handle the situation. In case 1 one may wish to reduce all blocks to a constant size, thereby reducing the number of experimental units available. If experimental units are at a premium, this may not be the most desirable course of action. The other alternative is to use the natural blocks and then use the analysis as given in the previous section. Before doing so we mention that its validity will depend on one very important assumption, and that is the constancy of the variance σ_e^2 for all blocks. In general, the size of σ_e^2 will depend on the size of the blocks: The larger the blocks, the larger σ_e^2 will be since it is in part a measure of the variability of the experimental units within blocks (see I.9.2.4). In fact, this is the reason for reducing the block size since it may also reduce the experimental error. Experience shows that such a reduction in σ_e^2 is not appreciable for only modest reduction in block size. It is therefore quite reasonable to assume that σ_e^2 is constant for blocks of different size if the number of experimental units varies only slightly.

In case 2 one possibility is to estimate the missing values and then use the analysis for the proper design. Such a procedure, however, would only be approximate. The exact analysis then would require the analysis with variable block size as in case 1. Obviously, the assumption of constancy of experimental error is satisfied here if is was satisfied for the original proper design.

1.5 DISCONNECTED INCOMPLETE BLOCK DESIGNS

In deriving the intrablock analysis of an incomplete block design in Section 1.3.4 we have made the assumption that the $\boldsymbol{C}$ matrix of (1.9) has maximal rank $t-1$, that is, the corresponding design is a connected design. Although connectedness is a desirable property of a design and although most designs have this property, we shall encounter designs (see Chapter 8) that are constructed on purpose as disconnected designs. We shall therefore comment briefly on this class of designs.

Following Bose (1947a) a treatment and a block are said to be associated if the treatment is contained in that block. Two treatments are said to be connected if it is possible to pass from one to the other by means of a chain consisting alternately of treatments and blocks such that any two adjacent members of the chain are associated. If this holds true for any two treatments, then the design is said to be connected, otherwise it is said to be *disconnected* (see Section I.4.13.3 for a more formal definition and Srivastava and Anderson, 1970). Whether a design is connected or disconnected can be checked easily by applying the definition given above to the incidence matrix $\boldsymbol{N}$: If one can connect two nonzero elements of $\boldsymbol{N}$ by means of vertical and horizontal lines such that the vertices are at nonzero elements, then the two treatments are connected. In order to check whether a design is connected, it is sufficient to check whether a given treatment is connected to all the other $t-1$ treatments. If a design is disconnected, it follows then that (possibly after suitable relabeling of the treatments) the matrix $\boldsymbol{N}\boldsymbol{N}'$ and hence $\boldsymbol{C}$ consist of disjoint block diagonal matrices such that the treatments associated with one of these submatrices are connected with each other.

Suppose $\boldsymbol{C}$ has m submatrices, that is,

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{C}_1 & & & \\ & \boldsymbol{C}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{C}_m \end{bmatrix}$$

where $\boldsymbol{C}_\nu$ is $t_\nu \times t_\nu$ $\left(\sum_{\nu=1}^{m} t_\nu = t\right)$. It then follows that rank $(\boldsymbol{C}_\nu) = t_\nu - 1 (\nu = 1, 2, \ldots, m)$ and hence rank$(\boldsymbol{C}) = t - m$. The RNE is still of the form (1.8) with a solution given by (1.11), where in $\boldsymbol{C}^- = \widetilde{\boldsymbol{C}}^{-1}$ we now have, modifying Theorem 1.1,

$$\widetilde{\boldsymbol{C}} = \boldsymbol{C} + \begin{bmatrix} a_1 \mathfrak{I}\mathfrak{I}' & & & \\ & a_2 \mathfrak{I}\mathfrak{I}' & & \\ & & \ddots & \\ & & & a_m \mathfrak{I}\mathfrak{I}' \end{bmatrix}$$

Table 1.3 T|B-ANOVA for Disconnected Incomplete Block Design

Source	d.f.	SS
$\boldsymbol{X}_\beta \mid \mathfrak{I}$	$b - 1$	$\sum_j \frac{B_j^2}{k_j} - \frac{G^2}{n}$
$\boldsymbol{X}_\tau \mid \mathfrak{I}, \boldsymbol{X}_\beta$	$t - m$	$\sum_i \widehat{\tau}_i Q_i$
$\boldsymbol{I} \mid \mathfrak{I}, \boldsymbol{X}_\beta, \boldsymbol{X}_\tau$	$n - b - t + m$	Difference
Total	$n - 1$	$\sum_{ij\ell} y_{ij\ell}^2 - \frac{G^2}{n}$

Table 1.4 B|T-ANOVA for Disconnected Incomplete Block Design

Source	d.f.	SS
$\boldsymbol{X}_\tau \mid \mathfrak{I}$	$t - 1$	$\sum_i \frac{T_i^2}{r_i} - \frac{G^2}{n}$
$\boldsymbol{X}_\beta \mid \mathfrak{I}, \boldsymbol{X}_\tau$	$b - m$	Difference
$\boldsymbol{I} \mid \mathfrak{I}, \boldsymbol{X}_\tau, \boldsymbol{X}_\beta$	$n - t - b + m$	From Table 1.3
Total	$n - 1$	$\sum_{ij\ell} y_{ij\ell}^2 - \frac{G^2}{n}$

with $a_\nu(\nu = 1, 2, \ldots, m)$ arbitrary constants ($\neq 0$) and the $\mathfrak{I}\mathfrak{I}'$ matrices are of appropriate dimensions. Following the development in Section 1.3.6, this then leads to the ANOVA tables as given in Tables 1.3 and 1.4.

1.6 RANDOMIZATION ANALYSIS

So far we have derived the analysis of data from incomplete block designs using a Gauss–Markov linear model as specified in (1.1). We have justified the appropriate use of such an infinite population theory model in our earlier discussions of error control designs (see, e.g., Sections I.6.3 and I.9.2) as a substitute for a derived, that is, finite, population theory model that takes aspects of randomization into account. In this section we shall describe in mathematical terms the randomization procedure for an incomplete block design, derive an appropriate linear model, and apply it to the analysis of variance. This will show again, as we have argued in Section I.9.2 for the RCBD, that treatment effects and block effects cannot be considered symmetrically for purposes of statistical inference.

1.6.1 Derived Linear Model

Following Folks and Kempthorne (1960) we shall confine ourselves to proper (i.e., all $k_j = k$), equireplicate (i.e., all $r_i = r$) designs. The general situation is then as follows: We are given a set of b blocks, each of constant size k; a master plan specifies b sets of k treatments; these sets are assigned at random to the blocks; in each block the treatments are assigned at random to the experimental units (EU). This randomization procedure is described more formally by the following design random variables:

$$\alpha_j^u = \begin{cases} 1 & \text{if the } u\text{th set is assigned to the } j\text{th block} \\ 0 & \text{otherwise} \end{cases} \tag{1.28}$$

and

$$\delta_{j\ell}^{uv} = \begin{cases} 1 & \text{if the } uv \text{ treatment is assigned to the} \\ & \ell\text{th unit of the } j\text{th block} \\ 0 & \text{otherwise} \end{cases} \tag{1.29}$$

The uv treatment is one of the t treatments that, for a given design, has been assigned to the uth set.

Assuming additivity in the strict sense (see Section I.6.3), the conceptual response of the uv treatment assigned to the ℓth EU in the jth block can be written as

$$T_{j\ell uv} = U_{j\ell} + T_{uv} \tag{1.30}$$

where $U_{j\ell}$ is the contribution from the ℓth EU in the jth block and T_{uv} is the contribution from treatment uv. We then write further

$$\begin{aligned} T_{j\ell uv} &= \overline{U}_{..} + (\overline{U}_{j.} - \overline{U}_{..}) + (U_{j\ell} - \overline{U}_{j.}) + \overline{T}_{..} + (T_{uv} - \overline{T}_{..}) \\ &= \mu + b_j + \tau_{uv} + u_{j\ell} \end{aligned} \tag{1.31}$$

where

$\mu = \overline{U}_{..} + \overline{T}_{..}$ is the overall mean

$b_j = \overline{U}_{j.} - \overline{U}_{..}$ is the effect of the jth block $(j = 1, 2, \ldots, b)$

$\tau_{uv} = t_{uv} - \overline{T}_{..}$ is the effect of the uv treatment $(u = 1, 2, \ldots, b;\ v = 1, 2, \ldots, k)$

$u_{j\ell} = U_{j\ell} - \overline{U}_{j.}$ is the unit error $(\ell = 1, 2, \ldots, k)$

with $\sum_j b_j = 0 = \sum_{uv} \tau_{uv} = \sum_\ell u_{j\ell}$. We then express the observed response for the uv treatment, y_{uv}, as

$$\begin{aligned} y_{uv} &= \sum_j \sum_\ell \alpha_j^u \, \delta_{j\ell}^{uv} \, T_{j\ell uv} \\ &= \mu + \tau_{uv} + \sum_j \alpha_j^u b_j + \sum_j \sum_\ell \alpha_j^u \, \delta_{j\ell}^{uv} \, u_{j\ell} \\ &= \mu + \tau_{uv} + \beta_u + \omega_{uv} \end{aligned} \tag{1.32}$$

where

$$\beta_u = \sum_j \alpha_j^u \, b_j \tag{1.33}$$

is a random variable with

$$E(\beta_u) = 0 \qquad E(\beta_u^2) = \frac{1}{b} \sum_j b_j^2 \qquad E(\beta_u \beta_{u'}) = \frac{1}{b(b-1)} \sum_j b_j^2 \qquad (u \neq u')$$

Also,

$$\omega_{uv} = \sum_j \sum_\ell \alpha_j^u \, \delta_{j\ell}^{uv} \, u_{j\ell} \tag{1.34}$$

is a random variable with

$$E(\omega_{uv}) = 0 \qquad E(\omega_{uv}^2) = \frac{1}{bk} \sum_{j\ell} u_{j\ell}^2$$

$$E(\omega_{uv}\omega_{uv'}) = -\frac{1}{bk(k-1)} \sum_j \sum_\ell u_{j\ell}^2 \qquad (v \neq v')$$

$$E(\omega_{uv}\omega_{u'v'}) = 0 \qquad (u \neq u')$$

In deriving the properties of the random variables β_u and ω_{uv} we have used, of course, the familiar distributional properties of the design random variables α_j^u and $\delta_{j\ell}^{uv}$, such as

$$P(\alpha_j^u = 1) = \frac{1}{b}$$

$$P(\alpha_j^u = 1 \mid \alpha_{j'}^u = 1) = 0 \qquad (j \neq j')$$

$$P(\alpha_j^u = 1 \mid \alpha_{j'}^{u'} = 1) = \frac{1}{b(b-1)} \qquad (u \neq u', j \neq j')$$

$$P(\delta_{j\ell}^{uv} = 1) = \frac{1}{k}$$

$$P(\delta_{j\ell}^{uv} = 1) \mid (\delta_{j\ell'}^{uv} = 1) = 0 \qquad (\ell \neq \ell')$$

$$P(\delta_{j\ell}^{uv} = 1) \mid (\delta_{j\ell'}^{uv'} = 1) = \frac{1}{k(k-1)} \qquad (\ell \neq \ell', v \neq v')$$

$$P(\delta_{j\ell}^{uv} = 1) \mid (\delta_{j'\ell'}^{u'v'} = 1) = \frac{1}{k^2} \qquad (j \neq j', u \neq u')$$

and so on.

1.6.2 Randomization Analysis of ANOVA Tables

Using model (1.32) and its distributional properties as induced by the design random variables α_j^u and $\delta_{j\ell}^{uv}$, we shall now derive expected values of the sums of squares in the analyses of variance as given in Tables 1.1 and 1.2:

1. $$\begin{aligned} E(\text{SS Total}) &= E \sum_{uv} (y_{uv} - \overline{y}_{..})^2 \\ &= E \sum_{uv} (\tau_{uv} + \beta_u + \omega_{uv})^2 \\ &= \sum_{uv} \tau_{uv}^2 + k \sum_j b_j^2 + \sum_{j\ell} u_{j\ell}^2 \end{aligned}$$

2.
$$\begin{aligned} E[\mathrm{SS}(\boldsymbol{X}_\beta \mid \mathfrak{I})] &= E \sum_{uv} (\overline{y}_{u.} - \overline{y}_{..})^2 \\ &= k\, E \sum_{u} \left(\overline{\tau}_{u.} + \beta_u + \frac{1}{k} \sum_{v} \omega_{uv}\right)^2 \\ &= k \sum_{u} \overline{\tau}_{u.}^2 + k \sum_{j} b_j^2 \end{aligned}$$

3. $E[\mathrm{SS}\,(\boldsymbol{I} \mid \mathfrak{I}, \boldsymbol{X}_\beta, \boldsymbol{X}_\tau)] = \dfrac{n-t-b+1}{b(k-1)} \displaystyle\sum_{j\ell} u_{j\ell}^2$

since the incomplete block designs considered are unbiased.

4. $E[\mathrm{SS}(\boldsymbol{X}_\tau \mid \mathfrak{I}, \boldsymbol{X}_\beta)]$ can be obtained by subtraction.

5. To obtain $E[\mathrm{SS}(\boldsymbol{X}_\tau \mid \mathfrak{I})]$ let

$$\gamma_{uv}^w = \begin{cases} 1 & \text{if the } w\text{th treatment corresponds to the} \\ & uv \text{ index} \qquad (w = 1, 2, \ldots, t) \\ 0 & \text{otherwise} \end{cases}$$

$$\gamma_{u}^w = \begin{cases} 1 & \text{if the } w\text{th treatment occurs in the } u\text{th block} \\ 0 & \text{otherwise} \end{cases}$$

where

$$\sum_{v} \gamma_{uv}^w = \gamma_u^w$$

and

$$\sum_{u} \gamma_u^w = r$$

Then

$$\begin{aligned} E[\mathrm{SS}(\boldsymbol{X}_\tau \mid \mathfrak{I})] &= E\left[\frac{1}{r} \sum_{w} \left(\sum_{uv} \gamma_{uv}^w y_{uv} - \frac{1}{t} \sum_{w} \sum_{uv} \gamma_{uv}^w y_{uv}\right)^2\right] \\ &= E\left[\frac{1}{r} \sum_{w} \left(r\tau_w + \sum_{u} \gamma_u^w\, \beta_u + \sum_{uv} \gamma_{uv}^w\, \omega_{uv}\right)^2\right] \\ &= r \sum_{w} \tau_w^2 + \frac{1}{r} \sum_{w} E\left(\sum_{u} \gamma_u^w\, \beta_u\right)^2 + \frac{1}{r} \sum_{w} E\left(\sum_{uv} \gamma_{uv}^w\, \omega_{uv}\right)^2 \end{aligned}$$

$$= r \sum_w \tau_w^2 + \frac{1}{r} \sum_w E \left(\sum_u \gamma_u^w \beta_u^2 + \sum_{\substack{uu' \\ u \neq u'}} \gamma_u^w \gamma_{u'}^w \beta_u \beta_{u'} \right)$$

$$+ \frac{1}{r} \sum_w E \left(\sum_{uv} \gamma_{uv}^w \omega_{uv}^2 + \sum_u \sum_{\substack{vv' \\ u \neq v'}} \gamma_{uv}^w \gamma_{uv'}^w \omega_{uv} \omega_{uv'} \right.$$

$$\left. + \sum_{\substack{uu' \\ u \neq u'}} \sum_{vv'} \gamma_{uv}^w \gamma_{u'v'}^w \omega_{uv} \omega_{u'v'} \right)$$

Now

$$E\left[\sum_u \gamma_u^w \beta_u^2\right] = r \frac{1}{b} \sum_j b_j^2$$

$$E\left[\sum_{\substack{uu' \\ u \neq u'}} \gamma_u^w \gamma_{u'}^w \beta_u \beta_{u'}\right] = - \sum_{\substack{uu' \\ u \neq u'}} \gamma_u^w \gamma_{u'}^w \frac{1}{b(b-1)} \sum_j b_j^2$$

$$= - \sum_u \gamma_u^w (r - \gamma_u^w) \frac{1}{b(b-1)} \sum_j b_j^2$$

$$= - \frac{r(r-1)}{b(b-1)} \sum_j b_j^2$$

$$E\left[\sum_{uv} \gamma_{uv}^w \omega_{uv}^2\right] = r \frac{1}{bk} \sum_{j\ell} u_{u\ell}^2$$

$$E\left[\sum_u \sum_{\substack{vv' \\ v \neq v'}} \gamma_{uv}^w \gamma_{uv'}^w \omega_{uv}^w \omega_{uv'}\right] = 0 \quad \text{since} \quad \gamma_{uv}^w \gamma_{uv'}^w \quad = \quad 0$$

$$E\left[\sum_{\substack{uu' \\ u \neq u'}} \gamma_{uv}^w \gamma_{u'v'}^w \omega_{uv} \omega_{u'v'}\right] = 0$$

and hence

$$E[(\boldsymbol{X}_\tau \mid \mathfrak{I}] = r \sum_w \tau_w^2 + \frac{t(b-r)}{b(b-1)} \sum_j b_j^2 + \frac{t}{bk} \sum_{j\ell} u_{j\ell}^2$$

Thus, we have for the mean squares (MS) from Tables 1.1 and 1.2 the expected values under randomization theory as given in Tables 1.5 and 1.6, respectively.

Table 1.5 E(MS) for T|B-ANOVA

Source	E(MS)
$\boldsymbol{X}_\beta \mid \mathfrak{I}$	$\dfrac{k}{b-1}\sum_u \overline{\tau}_{u.}^2 + \dfrac{k}{b-1}\sum_j b_j^2$
$\boldsymbol{X}_\tau \mid \mathfrak{I}, \boldsymbol{X}_\beta$	$\dfrac{1}{b(k-1)}\sum_{j\ell} u_{j\ell}^2 + \left(\sum_{uv} \tau_{uv}^2 - k\sum_u \overline{\tau}_{u.}^2\right) \Big/ (t-1)$
$\boldsymbol{I} \mid \mathfrak{I}, \boldsymbol{X}_\beta, \boldsymbol{X}_\tau$	$\dfrac{1}{b(k-1)}\sum_{j\ell} u_{j\ell}^2$

Table 1.6 E(MS) for B|T-ANOVA

Source	E(MS)
$\boldsymbol{X}_\tau \mid \mathfrak{I}$	$\dfrac{t}{bk(t-1)}\sum_{j\ell} u_{j\ell}^2 + \dfrac{t(b-r)}{b(b-1)(t-1)}\sum_j b_j^2 + \dfrac{r}{t-1}\sum_w \tau_w^2$
$\boldsymbol{X}_\beta \mid \mathfrak{I}, \boldsymbol{X}_\tau$	$\dfrac{t-k}{b(b-1)k(k-1)}\sum_{j\ell} u_{j\ell}^2 + \dfrac{bk-t}{(b-1)^2}\sum_j b_j^2$
$\boldsymbol{I} \mid \mathfrak{I}, \boldsymbol{X}_\beta, \boldsymbol{X}_\tau$	$\dfrac{1}{b(k-1)}\sum_{j\ell} u_{j\ell}^2$

If we define

$$\frac{1}{b(k-1)}\sum_{j\ell} u_{j\ell}^2 = \sigma_u^2$$

and

$$\frac{1}{b-1}\sum_j b_j^2 = \sigma_\beta^2$$

we can then express the expected values for the three important mean squares in ANOVA Tables 1.5 and 1.6 as

$$E\left[\mathrm{MS}(\boldsymbol{X}_\tau \mid \mathfrak{I}, \boldsymbol{X}_\beta)\right] = \sigma_u^2 + \frac{\sum_{uv} \tau_{uv}^2 - k\sum_u \overline{\tau}_{u.}^2}{t-1} \tag{1.35}$$

$$E\left[\mathrm{MS}(\boldsymbol{X}_\beta \mid \mathfrak{I}, \boldsymbol{X}_\tau)\right] = \frac{t-k}{(b-1)k}\sigma_u^2 + \frac{n-t}{b-1}\sigma_\beta^2 \tag{1.36}$$

$$E\left[\mathrm{MS}(\boldsymbol{I} \mid \mathfrak{I}, \boldsymbol{X}_\beta, \boldsymbol{X}_\tau)\right] = \sigma_u^2 \tag{1.37}$$

We make the following observations:

1. The quadratic form in the τ_{uv} in (1.35) is just a different way of writing $\boldsymbol{\tau}'\mathbf{C}\boldsymbol{\tau}$ in Table 1.1. Both expressions indicate that the quadratic form depends on the particular design chosen, and both equal zero when all the treatment effects are the same.
2. It follows from (1.35) and (1.37) that, based on the equality of the $E(\text{MS})$ under $H_0: \ \tau_1 = \tau_2 = \cdots = \tau_t$, the ratio

$$\text{MS}(\boldsymbol{X}_\tau \mid \mathfrak{I}, \boldsymbol{X}_\beta)/\text{MS}(\boldsymbol{I} \mid \mathfrak{I}, \boldsymbol{X}_\beta, \boldsymbol{X}_\tau) \tag{1.38}$$

   provides a test criterion for testing the above hypothesis. In fact, Ogawa (1974) has shown that the asymptotic randomization distribution of (1.38) is an F distribution with $t-1$ and $n-t-b+1$ d.f. We interpret this again to mean that the usual F test is an approximation to the randomization test based on (1.38).
3. Considering (1.36) and (1.37), there does not exist an exact test for testing the equality of block effects. This is in agreement with our discussion in Section I.9.2 concerning the asymmetry of treatment and block effects.
4. For $k = t$ and $r = b$, that is, for the RCBD, the results of Tables 1.5 and 1.6 agree with those in Table 9.1 of Section I.9.2.
5. With treatment-unit additivity in the broad sense (see Section I.6.3.3) the expressions in (1.35), (1.36), and (1.37) are changed by adding $\sigma_\nu^2 + \sigma_\eta^2$ to the right-hand sides (recall that $\sigma_u^2 + \sigma_\nu^2 + \sigma_\eta^2 \equiv \sigma_\epsilon^2 + \sigma_\eta^2 \equiv \sigma_e^2$). Remarks (2) and (3) above remain unchanged.
6. For the recovery of interblock information (to be discussed in Section 1.7), we need to estimate σ_β^2 (or a function of σ_β^2). Clearly, under the assumption of additivity in the broad sense, this cannot be done considering that

$$E\left[\text{MS}\left(\boldsymbol{X}_\beta \mid \mathfrak{I}, \boldsymbol{X}_\tau\right)\right] = \sigma_\nu^2 + \sigma_\eta^2 + \frac{t-k}{(b-1)k}\,\sigma_u^2 + \frac{n-t}{b-1}\,\sigma_\beta^2$$

   It is for this reason only that we shall resort to the approximation

$$E\left[\text{MS}\left(\boldsymbol{X}_\beta \mid \mathfrak{I}, \boldsymbol{X}_\tau\right)\right] \approx \sigma_\nu^2 + \sigma_\eta^2 + \sigma_u^2 + \frac{n-t}{b-1}\,\sigma_\beta^2 = \sigma_e^2 + \frac{n-t}{b-1}\,\sigma_\beta^2 \tag{1.39}$$

   which is the expected value based on an infinite population theory model [see (1.49) and (1.50)].

For a different approach to randomization analysis, see Calinski and Kageyama (2000).

1.7 INTERBLOCK INFORMATION IN AN INCOMPLETE BLOCK DESIGN

1.7.1 Introduction and Rationale

As mentioned earlier, Yates (1939, 1940a) has argued that for incomplete block designs comparisons among block totals (or averages) contain some information about treatment comparisons, and he referred to this as *recovery of interblock information*. The basic idea is as follows.

Consider, for purposes of illustration, the following two blocks and their observations from some design:

$$\begin{array}{llll} \text{Block 1:} & y_{51}, & y_{31}, & y_{11} \\ \text{Block 2:} & y_{22}, & y_{42}, & y_{32} \end{array}$$

Let

$$B_1 = y_{51} + y_{31} + y_{11}$$

and

$$B_2 = y_{22} + y_{42} + y_{32}$$

represent the block totals. Using model (1.1) we can write

$$\begin{aligned} B_1 - B_2 = (\tau_5 + \tau_2 + \tau_1) - (\tau_2 + \tau_4 + \tau_3) \\ + 3\beta_1 - 3\beta_2 + (e_{51} + e_{31} + e_{11}) \\ - (e_{22} + e_{42} + e_{32}) \end{aligned}$$

Assuming now that the β_j are random effects with mean zero, we find

$$E(B_1 - B_2) = \tau_1 + \tau_5 - \tau_2 - \tau_4$$

It is in this sense that block comparisons contain information about treatment comparisons. We shall now formalize this procedure.

1.7.2 Interblock Normal Equations

Consider the model equation (1.1)

$$\boldsymbol{y} = \mu\mathfrak{I} + \boldsymbol{X}_\tau\boldsymbol{\tau} + \boldsymbol{X}_\beta\boldsymbol{\beta} + \boldsymbol{e}$$

where $\boldsymbol{\beta}$ is now assumed to be a random vector with $E(\boldsymbol{\beta}) = \mathbf{0}$ and $\text{var}(\boldsymbol{\beta}) = \sigma^2_\beta \boldsymbol{I}$. As pointed out above the interblock analysis is based on block totals rather than on individual observations, that is, we now consider

$$\boldsymbol{X}'_\beta \boldsymbol{y} = \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_b \end{pmatrix} \mu + \boldsymbol{N}'\boldsymbol{\tau} + \boldsymbol{K}\boldsymbol{\beta} + \boldsymbol{X}'_\beta \boldsymbol{e} \tag{1.40}$$

We then have

$$E(\boldsymbol{X}'_\beta \boldsymbol{y}) = \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_b \end{pmatrix} \mu + \boldsymbol{N}'\boldsymbol{\tau} \tag{1.41}$$

and the variance–covariance matrix under what we might call a double error structure with both $\boldsymbol{\beta}$ and $\boldsymbol{e}$ being random vectors

$$\begin{aligned} \operatorname{var}(\boldsymbol{X}'_\beta \boldsymbol{y}) &= \boldsymbol{K}^2 \sigma_\beta^2 + \boldsymbol{K}\sigma_e^2 \\ &= \boldsymbol{K}\left(\boldsymbol{I} + \frac{\sigma_\beta^2}{\sigma_e^2}\boldsymbol{K}\right)\sigma_e^2 \\ &= \boldsymbol{L}\sigma_e^2 \quad \text{(say)} \end{aligned}$$

and

$$\boldsymbol{L} = \operatorname{diag}\{\ell_j\} = \operatorname{diag}\left\{k_j\left(1 + \frac{\sigma_\beta^2}{\sigma_e^2}k_j\right)\right\} = \operatorname{diag}\left\{\left(k_j \frac{w}{w'_j}\right)\right\} = \operatorname{diag}\{k_j \rho_j\}$$

with

$$w = \frac{1}{\sigma_e^2} \tag{1.42}$$

and

$$w'_j = \frac{1}{\sigma_e^2 + k_j \sigma_\beta^2} \tag{1.43}$$

and

$$\rho_j = \frac{w}{w'_j} = \frac{\sigma_e^2 + k_j \sigma_\beta^2}{\sigma_e^2} \tag{1.44}$$

The quantities w and w'_j of (1.42) and (1.43) are referred to as *intrablock* and *interblock weights*, respectively, as w is the reciprocal of the intrablock variance,

σ_e^2, and w'_j is the reciprocal of the interblock variance, that is, $\mathrm{var}(B_j)$ on a per observation basis, or $\mathrm{var}(B_j/k_j) = \sigma_e^2 + k_j\sigma_\beta^2$. We then use as our "observation" vector

$$z = \boldsymbol{L}^{-1/2}\boldsymbol{X}'_\beta \boldsymbol{y} \tag{1.45}$$

which has

$$\mathrm{var}(z) = \boldsymbol{I}\sigma_e^2$$

and hence satisfies the Gauss–Markov conditions.

If we write (1.41) as

$$E(\boldsymbol{X}'_\beta \boldsymbol{y}) = (\boldsymbol{k}\ \boldsymbol{N}')\begin{pmatrix}\mu\\ \boldsymbol{\tau}\end{pmatrix}$$

with $\boldsymbol{k} = (k_1, k_2, \ldots, k_b)'$, we have from (1.45) that

$$E(z) = \boldsymbol{L}^{-1/2}(\boldsymbol{k}\ \boldsymbol{N}')\begin{pmatrix}\mu\\ \boldsymbol{\tau}\end{pmatrix}$$

The resulting NE, which we shall refer to as the *interblock* NE, is then given by

$$\begin{pmatrix}\boldsymbol{k}'\\ \boldsymbol{N}\end{pmatrix}\boldsymbol{L}^{-1}(\boldsymbol{k}\ \boldsymbol{N}')\begin{pmatrix}\mu^*\\ \boldsymbol{\tau}^*\end{pmatrix} = \begin{pmatrix}\boldsymbol{k}'\\ \boldsymbol{N}\end{pmatrix}\boldsymbol{L}^{-1}\boldsymbol{X}'_\beta \boldsymbol{y} \tag{1.46}$$

or explicitly as

$$\begin{bmatrix} \sum_j \dfrac{k_j^2}{\ell_j} & \sum_j n_{1j}\dfrac{k_j}{\ell_j} & \cdots & \sum_j n_{tj}\dfrac{k_j}{\ell_j} \\ \sum_j n_{1j}\dfrac{k_j}{\ell_j} & \sum_j \dfrac{n_{1j}^2}{\ell_j} & \cdots & \sum_j \dfrac{n_{1j}n_{tj}}{\ell_j} \\ \vdots & \vdots & & \vdots \\ \sum_j n_{tj}\dfrac{k_j}{\ell_j} & \sum_j \dfrac{n_{1j}n_{tj}}{\ell_j} & \cdots & \sum_j \dfrac{n_{tj}^2}{\ell_j} \end{bmatrix} \begin{bmatrix}\mu^*\\ \tau_1^*\\ \vdots\\ \tau_t^*\end{bmatrix} = \begin{bmatrix} \sum_j \dfrac{k_j}{\ell_j}B_j \\ \sum_j \dfrac{n_{1j}}{\ell_j}B_j \\ \vdots \\ \sum_j \dfrac{n_{tj}}{\ell_j}B_j \end{bmatrix} \tag{1.47}$$

It can be seen easily that the rank of the coefficient matrix in (1.47) is t. To solve the interblock NE, we take $\mu^* = 0$ and hence reduce the set to the following t

equations in $\tau_1^*, \tau_2^*, \ldots, \tau_t^*$ where we have used the fact that $\ell_j = k_j \rho_j$:

$$\begin{bmatrix} \sum_j n_{1j}^2 \frac{\rho_j^{-1}}{k_j} & \sum_j n_{1j} n_{2j} \frac{\rho_j^{-1}}{k_j} & \cdots & \sum_j n_{1j} n_{tj} \frac{\rho_j^{-1}}{k_j} \\ \sum n_{2j} n_{1j} \frac{\rho_j^{-1}}{k_j} & \sum n_{2j}^2 \frac{\rho_j^{-1}}{k_j} & \cdots & \sum n_{2j} n_{tj} \frac{\rho_j^{-1}}{k_j} \\ \vdots & \vdots & & \vdots \\ \sum n_{tj} n_{1j} \frac{\rho_j^{-1}}{k_j} & \sum n_{tj} n_{2j} \frac{\rho_j^{-1}}{k_j} & \cdots & \sum n_{tj}^2 \frac{\rho_j^{-1}}{k_j} \end{bmatrix} \begin{bmatrix} \tau_1^* \\ \tau_2^* \\ \vdots \\ \tau_t^* \end{bmatrix}$$

$$= \begin{bmatrix} \sum_j \frac{\rho_j^{-1}}{k_j} n_{1j} B_j \\ \sum_j \frac{\rho_j^{-1}}{k_j} n_{2j} B_j \\ \vdots \\ \sum_j \frac{\rho_j^{-1}}{k_j} n_{tj} B_j \end{bmatrix} \tag{1.48}$$

The solution to the equations (1.48) is referred to as the *interblock information* about the treatment effects, with

$$E(\tau_i^*) = \mu + \tau_i + \text{const.} \cdot \sum_{i'=1}^{t} (\mu + \tau_{i'})$$

Hence

$$E\left(\sum_i c_i \tau_i^*\right) = \sum_i c_i \tau_i \quad \text{for} \quad \sum c_i = 0$$

We note here that typically (see Kempthorne, 1952) the interblock analysis is derived not in terms of the "observations" z [as given in (1.45)] but rather in terms of the block totals $\boldsymbol{X}_\beta' \boldsymbol{y}$. The resulting NE are then simply obtained by using $\boldsymbol{L} = \boldsymbol{I}$ in (1.46) and subsequent equations. The reason why we prefer our description is the fact that then the intra- and interblock information can be combined additively to obtain the so-called combined analysis (see Section 1.8) rather than in the form of a weighted average (see Kempthorne, 1952).

1.7.3 Nonavailability of Interblock Information

We conclude this section with the following obvious remarks:

1. For the special case $k_j = t$ for all $j, r_i = b$ for all i and all $n_{ij} = 1$, we have, of course, the RCBD. Then the elements in the coefficient matrix of (1.48) are all identical, and so are the right-hand sides. Consequently, (1.48) reduces to a single equation

$$b \sum_i \tau_i^* = \sum_j B_j$$

and no contrasts among the τ_i are estimable. Expressed differently, any contrast among block totals estimates zero, that is, no interblock information is available.

2. For a design with $b < t$ (and such incomplete block designs exist as we shall see in Chapter 4; see also the example in Section 1.7.1), the rank of the coefficient matrix of (1.48), $\boldsymbol{N}\boldsymbol{L}^{-1}\boldsymbol{N}'$, is less than t. Hence not all $\mu + \tau_i$ are estimable using block totals, which means that interblock information is not available for all treatment contrasts.

1.8 COMBINED INTRA- AND INTERBLOCK ANALYSIS

1.8.1 Combining Intra- and Interblock Information

The two types of information about treatment effects that we derived in Sections 1.3.2 and 1.7.2 can be combined to yield the "best" information about estimable functions of treatment effects. All we need to do is to add the coefficient matrices from the intrablock RNE (1.8) and the interblock NE (1.48) and do the same for the corresponding right-hand sides. This will lead to a system of equations in $\tau_1^{**}, \tau_2^{**}, \ldots, \tau_t^{**}$, say, and the solution to these equations will lead to the combined intra- and interblock estimators for treatment contrasts.

In the following section we shall derive the equations mentioned above more directly using the method of generalized least squares, that is, by using the Aitken equations described in Section I.4.16.2

1.8.2 Linear Model

In order to exhibit the double error structure that characterizes the underlying assumptions for the combined analysis, we rewrite model (1.1) as

$$y_{j\ell} = \mu + \tau_{j\ell} + \beta_j + e_{j\ell} \tag{1.49}$$

where $j = 1, 2, \ldots, b$; $\ell = 1, 2, \ldots, k_j$; $\tau_{j\ell}$ denotes the effect of the treatment applied to the ℓth experimental unit in the jth block, the β_j are assumed to be i.i.d. random variables with $E(\beta_j) = 0, \text{var}(\beta_j) = \sigma_\beta^2$, and the $e_{j\ell}$ are i.i.d.

random variables with $E(e_{j\ell}) = 0$, $\text{var}(e_{j\ell}) = \sigma_e^2$. We then have

$$E(y_{j\ell}) = \mu + \tau_{j\ell} \tag{1.50}$$

and

$$\text{cov}\left(y_{j\ell}, y_{j'\ell'}\right) = \begin{cases} \sigma_\beta^2 + \sigma_e^2 & \text{for } j = j', \ell = \ell' \\ \sigma_\beta^2 & \text{for } j = j', \ell \neq \ell' \\ 0 & \text{otherwise} \end{cases} \tag{1.51}$$

To use matrix notation it is useful to arrange the observations according to blocks, that is, write the observation vector as

$$\boldsymbol{y} = \left(y_{11}, y_{12}, \ldots, y_{1k_1}, y_{21}, y_{22}, \ldots, y_{b1}, y_{b2}, \ldots, y_{bk_b}\right)'$$

Letting

$$\boldsymbol{X} = (\mathfrak{I} \quad \boldsymbol{X}_\tau)$$

we rewrite (1.50) as

$$E(\boldsymbol{y}) = \boldsymbol{X}\begin{pmatrix} \mu \\ \boldsymbol{\tau} \end{pmatrix} \tag{1.52}$$

and the variance–covariance (1.51) as

$$\text{var}(\boldsymbol{y}) = \begin{bmatrix} \boldsymbol{V}_1 & & & \\ & \boldsymbol{V}_2 & & \boldsymbol{0} \\ & & \ddots & \\ \boldsymbol{0} & & & \boldsymbol{V}_b \end{bmatrix} \sigma_e^2 \equiv \boldsymbol{V}\sigma_e^2 \tag{1.53}$$

where $\boldsymbol{V}_j$ is given by

$$\boldsymbol{V}_j = \boldsymbol{I}_{k_j} + \frac{\sigma_\beta^2}{\sigma_e^2}\, \mathfrak{I}_{k_j}\, \mathfrak{I}'_{k_j} \tag{1.54}$$

1.8.3 Normal Equations

Applying now the principles of least squares to the model (1.52) with covariance structure (1.53) yields the Aitken equations (see Section I.4.16):

$$(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})\begin{pmatrix} \widehat{\widehat{\mu}} \\ \widehat{\widehat{\tau}} \end{pmatrix} = \boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y} \tag{1.55}$$

where

$$\boldsymbol{V}^{-1} = \text{diag}\left(\boldsymbol{V}_1^{-1}, \boldsymbol{V}_2^{-1}, \ldots, \boldsymbol{V}_b^{-1}\right)$$

and

$$\boldsymbol{V}_j^{-1} = \boldsymbol{I}_{k_j} - \frac{\sigma_\beta^2}{\sigma_e^2 + k_j\sigma_\beta^2}\ \boldsymbol{\mathfrak{I}}_{k_j}\boldsymbol{\mathfrak{I}}'_{k_j} \tag{1.56}$$

With

$$\frac{1}{\sigma_e^2} = w \qquad \frac{1}{\sigma_e^2 + k_j\sigma_\beta^2} = w'_j$$

and

$$\frac{w'_j}{w} = \rho_j^{-1}$$

Eq. (1.56) can be written as

$$\boldsymbol{V}_j^{-1} = \boldsymbol{I}_{k_j} - \frac{1-\rho_j^{-1}}{k_j}\ \boldsymbol{\mathfrak{I}}_{k_j}\boldsymbol{\mathfrak{I}}'_{k_j}$$

and hence

$$\boldsymbol{V}^{-1} = \boldsymbol{I}_n - \operatorname{diag}\left(\frac{1-\rho_1^{-1}}{k_1}\ \boldsymbol{\mathfrak{I}}_{k_1}\boldsymbol{\mathfrak{I}}'_{k_1}, \frac{1-\rho_2^{-1}}{k_2}\ \boldsymbol{\mathfrak{I}}_{k_2}\boldsymbol{\mathfrak{I}}'_{k_2}, \ldots, \frac{1-\rho_b^{-1}}{k_b}\ \boldsymbol{\mathfrak{I}}_{k_b}\boldsymbol{\mathfrak{I}}'_{k_b}\right) \tag{1.57}$$

Further, if we let $\left(1-p_j^{-1}\right)/k_j = \delta_j (j = 1, 2, \ldots, b)$, we have

$$\boldsymbol{X}'\boldsymbol{V}^{-1} = \boldsymbol{X}' - \begin{bmatrix} \left(1-\rho_1^{-1}\right)\boldsymbol{\mathfrak{I}}'_{k_1} & \left(1-\rho_2^{-1}\right)\boldsymbol{\mathfrak{I}}'_{k_2} & \cdots & \left(1-\rho_b^{-1}\right)\boldsymbol{\mathfrak{I}}'_{k_b} & \\ \delta_1 n_{11}\boldsymbol{\mathfrak{I}}'_{k_1} & \delta_2 n_{12}\boldsymbol{\mathfrak{I}}'_{k_2} & \cdots & \delta_b n_{1b}\boldsymbol{\mathfrak{I}}'_{k_b} & \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \delta_1 n_{t1}\boldsymbol{\mathfrak{I}}'_{k_1} & \delta_2 n_{t2}\boldsymbol{\mathfrak{I}}'_{k_2} & \cdots & \delta_b n_{tb}\boldsymbol{\mathfrak{I}}'_{k_b} & \end{bmatrix}$$

and

$$\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X} = \begin{bmatrix} \sum k_j\rho_j^{-1} & \sum n_{1j}\rho_j^{-1} & \sum n_{2j}\rho_j^{-1} & \cdots & \sum n_{tj}\rho_j^{-1} \\ \sum n_{1j}\rho_j^{-1} & r_1 - \sum \delta_j n_{1j}^2 & -\sum \delta_j n_{1j}n_{2j} & \cdots & -\sum \delta_j n_{1j}n_{tj} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \sum n_{tj}\rho_j^{-1} & -\sum \delta_j n_{1j}n_{tj} & -\sum \delta_j n_{2j}n_{tj} & \cdots & r_t - \sum \delta_j n_{tj}^2 \end{bmatrix} \tag{1.58}$$

$$X'\mathbf{V}^{-1}\boldsymbol{y} = \begin{bmatrix} \sum \rho_j^{-1} B_j \\ T_1 - \sum \delta_j n_{1j} B_j \\ \vdots \\ T_t - \sum \delta_j N_{tj} B_j \end{bmatrix} \tag{1.59}$$

By inspection one can verify that in the coefficient matrix (1.58) the elements in rows 2 to $t+1$ add up to the elements in row 1, which shows that (1.55) is not of full rank; in fact, it is of rank t. The easiest way to solve these equations then is to impose the condition $\widehat{\widehat{\mu}} = 0$. This means that we eliminate the first row and first column from (1.58) and the first element in (1.59) and solve the resulting system of t equations in the t unknowns $\widehat{\widehat{\tau}}_1, \widehat{\widehat{\tau}}_2, \ldots, \widehat{\widehat{\tau}}_t$. If we define $\boldsymbol{S} = \text{diag}(\rho_j)$, then this system of equations resulting from (1.58) and (1.59) can be written as

$$\left[\boldsymbol{R} - \boldsymbol{N}\boldsymbol{K}^{-1}\left(\boldsymbol{I} - \boldsymbol{S}^{-1}\right)\right]\widehat{\widehat{\boldsymbol{\tau}}} = \boldsymbol{T} - \boldsymbol{N}\boldsymbol{K}^{-1}\left(\boldsymbol{I} - \boldsymbol{S}^{-1}\right)\boldsymbol{B} \tag{1.60}$$

which we write for short as

$$\boldsymbol{A}\widehat{\widehat{\boldsymbol{\tau}}} = \boldsymbol{P} \tag{1.61}$$

with $\boldsymbol{A}$ and $\boldsymbol{P}$ as described above. The solution then is

$$\widehat{\widehat{\boldsymbol{\tau}}} = \boldsymbol{A}^{-1}\boldsymbol{P} \tag{1.62}$$

or, in terms of a generalized inverse for the original set of NE (1.55)

$$\begin{bmatrix} \widehat{\widehat{\boldsymbol{\mu}}} \\ \widehat{\widehat{\boldsymbol{\tau}}} \end{bmatrix} = \begin{bmatrix} 0 & \mathbf{0}' \\ \mathbf{0} & \boldsymbol{A}^{-1} \end{bmatrix} \boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y} \tag{1.63}$$

with

$$E\left(\widehat{\widehat{\tau}}_i\right) = \mu + \tau_i \qquad (i = 1, 2, \ldots, t)$$

If we denote the (i, i') element of $\boldsymbol{A}^{-1}$ by $a^{ii'}$, then

$$\text{var}\left(\widehat{\widehat{\tau_i - \tau_{i'}}}\right) = \text{var}\left(\widehat{\widehat{\tau}}_i - \widehat{\widehat{\tau}}_{i'}\right) = \left(a^{ii} + a^{i'i'} - 2a^{ii'}\right)\sigma_e^2 \tag{1.64}$$

More generally, the treatment contrast $\mathbf{c}'\boldsymbol{\tau}$ is estimated by $\mathbf{c}'\widehat{\widehat{\boldsymbol{\tau}}}$ with variance

$$\text{var}\left(\boldsymbol{c}'\widehat{\widehat{\boldsymbol{\tau}}}\right) = \boldsymbol{c}'\boldsymbol{A}^{-1}\boldsymbol{c}\sigma_e^2 \tag{1.65}$$

Expression (1.65) looks deceptively simple, but the reader should keep in mind that the elements of $\boldsymbol{A}^{-1}$ depend on σ_β^2 and σ_e^2. We shall return to estimating (1.65) in Section 1.10.

Finally, we note that the equations (1.60) show a striking similarity to the intrablock NE (1.7), except that the system (1.60) is of full rank and the elements of its coefficient matrix depend on the unknown parameters σ_β^2 and σ_e^2.

1.8.4 Some Special Cases

As a special case of the above derivations we mention explicitly the equireplicate, proper design, that is, the design with all $r_i = r$ and all $k_j = k$. We then define

$$w' = \frac{1}{\sigma_e^2 + k\sigma_\beta^2} \tag{1.66}$$

and

$$\rho = \frac{w}{w'} \tag{1.67}$$

and write (1.60) as

$$\left[r\boldsymbol{I} - \frac{1}{k}\left(1 - \rho^{-1}\right)\boldsymbol{N}\boldsymbol{N}'\right]\widehat{\widehat{\boldsymbol{\tau}}} = \boldsymbol{T} - \frac{1}{k}\left(1 - \rho^{-1}\right)\boldsymbol{N}\boldsymbol{B} \tag{1.68}$$

We shall comment briefly on the set of equations (1.68) for two special cases:

1. If $\rho^{-1} = 0$, that is, $\sigma_\beta^2 = \infty$, then (1.68) reduces to the NE (1.7) for the intrablock analysis. This means, of course, that in this case no interblock information is available. This is entirely plausible and suggests further that for “large” σ_β^2 the interblock information is very weak and perhaps not worthwhile considering.
2. If $\rho^{-1} = 1$, that is, $\sigma_\beta^2 = 0$, the solution to (1.68) is the same as that obtained for the completely randomized design (CRD) with the restriction $\hat{\mu} = 0$. This, of course, is a formal statement and should not imply that in this case the observations should be analyzed as if a CRD had been used.

1.9 RELATIONSHIPS AMONG INTRABLOCK, INTERBLOCK, AND COMBINED ESTIMATION

On several occasions we have pointed out that there exist certain relationships among the different types of analysis for incomplete block designs. It is worthwhile to exposit this in a little detail.

1.9.1 General Case

For the full model

$$\boldsymbol{y} = \boldsymbol{\Im}_n \mu + \boldsymbol{X}_\tau \boldsymbol{\tau} + \boldsymbol{X}_\beta \boldsymbol{\beta} + \boldsymbol{e}$$

with the double error structure we have

$$E(\boldsymbol{y}) = \boldsymbol{\Im}\mu + \boldsymbol{X}_\tau \boldsymbol{\tau}$$

and

$$\begin{aligned} \text{var}(\boldsymbol{y}) &= \boldsymbol{X}_\beta \boldsymbol{X}'_\beta \sigma^2_\beta + \boldsymbol{I}\sigma^2_e \\ &= (\boldsymbol{I} + \gamma \boldsymbol{X}_\beta \boldsymbol{X}'_\beta)\sigma^2_e \\ &\equiv \boldsymbol{V}\sigma^2_e \end{aligned}$$

as in (1.53) with $\gamma = \sigma^2_\beta/\sigma^2_e$. Then, as explained in Section 1.8, the estimators of estimable functions are obtained from the Aitken equations by minimizing

$$(\boldsymbol{y} - \boldsymbol{\Im}\mu - \boldsymbol{X}_\tau \boldsymbol{\tau})' \boldsymbol{V}^{-1} (\boldsymbol{y} - \boldsymbol{\Im}\mu - \boldsymbol{X}_\tau \boldsymbol{\tau}) \tag{1.69}$$

with respect to μ and $\boldsymbol{\tau}$. To simplify the algebra, let $\boldsymbol{\Theta} = \boldsymbol{\Im}\mu + \boldsymbol{X}_\tau \boldsymbol{\tau}$ and then write $\boldsymbol{\psi} = \boldsymbol{y} - \boldsymbol{\Theta}$. Expression (1.69) can now be written simply as $\boldsymbol{\psi}' \boldsymbol{V}^{-1} \boldsymbol{\psi}$. We then write

$$\boldsymbol{\psi} = \boldsymbol{P}\boldsymbol{\psi} + (\boldsymbol{I} - \boldsymbol{P})\boldsymbol{\psi}$$

where, again for brevity, we write

$$\begin{aligned} \boldsymbol{P} = \boldsymbol{P}_{\boldsymbol{x}_\beta} &= \boldsymbol{X}_\beta (\boldsymbol{X}'_\beta \boldsymbol{X}_\beta)^{-1} \boldsymbol{X}'_\beta \\ &= \boldsymbol{X}_\beta \boldsymbol{K}^{-1} \boldsymbol{X}'_\beta \end{aligned} \tag{1.70}$$

Then

$$\boldsymbol{\psi}' \boldsymbol{V}^{-1} \boldsymbol{\psi} = [\boldsymbol{\psi}' \boldsymbol{P} + \boldsymbol{\psi}' (\boldsymbol{I} - \boldsymbol{P})] \boldsymbol{V}^{-1} [\boldsymbol{P}\boldsymbol{\psi} + (\boldsymbol{I} - \boldsymbol{P})\boldsymbol{\psi}] \tag{1.71}$$

In order to expand the right-hand side of (1.71) we make use of the following results. From

$$(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{V} = (\boldsymbol{I} - \boldsymbol{P})(\boldsymbol{I} + \gamma \boldsymbol{X}_\beta \boldsymbol{X}'_\beta) = (\boldsymbol{I} - \boldsymbol{P})$$

[using (1.70)] it follows that

$$(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{V}^{-1} = (\boldsymbol{I} - \boldsymbol{P})$$

and hence

$$(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{V}^{-1}\boldsymbol{P} = (\boldsymbol{I} - \boldsymbol{P})\boldsymbol{P} = \boldsymbol{0}$$

Thus, (1.71) reduces to

$$\boldsymbol{\psi}'\boldsymbol{V}^{-1}\boldsymbol{\psi} = \boldsymbol{\psi}'\boldsymbol{P}\boldsymbol{V}^{-1}\boldsymbol{P}\boldsymbol{\psi} + \boldsymbol{\psi}'(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{\psi} \tag{1.72}$$

To handle the term $\boldsymbol{\psi}'\boldsymbol{P}\boldsymbol{V}^{-1}\boldsymbol{P}\boldsymbol{\psi}$ we note that $\boldsymbol{V}^{-1}$ in (1.57) can be rewritten as

$$\boldsymbol{V}^{-1} = \boldsymbol{I} - \gamma\boldsymbol{X}_\beta(\boldsymbol{I} + \gamma\boldsymbol{X}'_\beta\boldsymbol{X}_\beta)^{-1}\boldsymbol{X}'_\beta$$

so that

$$\begin{aligned}
\boldsymbol{P}\boldsymbol{V}^{-1}\boldsymbol{P} &= \boldsymbol{P} - \gamma\boldsymbol{X}_\beta(\boldsymbol{I} + \gamma\boldsymbol{X}'_\beta\boldsymbol{X}_\beta)^{-1}\boldsymbol{X}'_\beta \\
&= \boldsymbol{X}_\beta\left[\boldsymbol{K}^{-1} - \gamma(\boldsymbol{I} + \gamma\boldsymbol{K})^{-1}\right]\boldsymbol{X}'_\beta \\
&= \boldsymbol{X}_\beta \ \mathrm{diag}\left[\frac{1}{k_j(1 + \gamma k_j)}\right]\boldsymbol{X}'_\beta \\
&= \boldsymbol{X}_\beta(\boldsymbol{K} + \gamma\boldsymbol{K}^2)^{-1}\boldsymbol{X}'_\beta
\end{aligned}$$

Hence

$$\boldsymbol{\psi}'\boldsymbol{P}\boldsymbol{V}^{-1}\boldsymbol{P}\boldsymbol{\psi} = \left(\boldsymbol{X}'_\beta\boldsymbol{\psi}\right)'\left(\boldsymbol{K} + \gamma\boldsymbol{K}^2\right)^{-1}\left(\boldsymbol{X}'_\beta\boldsymbol{\psi}\right) \tag{1.73}$$

Then we note that $\boldsymbol{X}'_\beta\boldsymbol{\psi}$ is the vector of block totals of the vector $\boldsymbol{\psi}$. Since $\boldsymbol{\psi} = \boldsymbol{y} - \boldsymbol{\Theta}$, we have

$$\begin{aligned}
\mathrm{var}\left(\boldsymbol{X}'_\beta\boldsymbol{\psi}\right) &= \mathrm{var}\left(\boldsymbol{X}'_\beta\boldsymbol{y}\right) \\
&= \left(\boldsymbol{X}'_\beta\boldsymbol{X}_\beta\right)^2\sigma^2_\beta + \boldsymbol{X}'_\beta\boldsymbol{X}_\beta\sigma^2_e \\
&= \sigma^2_e\left(\boldsymbol{K} + \gamma\boldsymbol{K}^2\right)
\end{aligned}$$

Hence $\boldsymbol{\psi}'\boldsymbol{P}\boldsymbol{V}^{-1}\boldsymbol{P}\boldsymbol{\psi}$ is equal to

$$Q_2 \equiv [\boldsymbol{B} - E(\boldsymbol{B})]'\left(\boldsymbol{K} + \gamma\boldsymbol{K}^2\right)^{-1}[\boldsymbol{B} - E(\boldsymbol{B})] \tag{1.74}$$

with $\boldsymbol{B} = \boldsymbol{X}'_\beta\boldsymbol{y}$ being the vector of block totals.

The second expression in (1.72) is actually the quadratic form that needs to be minimized to obtain the RNE for $\mathbf{\Theta}$ (see Section I.4.7.1), and this is equivalent to minimizing

$$Q_1 \equiv (\boldsymbol{y} - \boldsymbol{X}_\tau \boldsymbol{\tau})'(\boldsymbol{I} - \boldsymbol{P})(\mathbf{y} - \boldsymbol{X}_\tau \boldsymbol{\tau}) \tag{1.75}$$

We thus summarize: To obtain the intrablock estimator, we minimize Q_1; to obtain the interblock estimator, we minimize Q_2; and to obtain the combined estimator, we minimize $Q_1 + Q_2$.

1.9.2 Case of Proper, Equireplicate Designs

We now consider the case with $\boldsymbol{R} = r\boldsymbol{I}$ and $\boldsymbol{K} = k\boldsymbol{I}$. The intrablock NE are [see (1.7)]

$$\left(r\boldsymbol{I} - \frac{1}{k}\boldsymbol{N}\boldsymbol{N}'\right)\hat{\boldsymbol{\tau}} = \boldsymbol{Q}$$

For the interblock observational equations

$$\boldsymbol{B} = k\mathfrak{I}\mu + \boldsymbol{N}'\boldsymbol{\tau} + \text{error}$$

or, absorbing μ into $\boldsymbol{\tau}$, that is, replacing $\mathfrak{I}\mu + \boldsymbol{\tau}$ by $\boldsymbol{\tau}$,

$$\boldsymbol{B} = \boldsymbol{N}'\boldsymbol{\tau} + \text{error}$$

we have the interblock NE [see (1.47)]

$$\boldsymbol{N}\boldsymbol{N}'\boldsymbol{\tau}^* = \boldsymbol{N}\boldsymbol{B}$$

As we have pointed out earlier (see Section 1.8) and as is obvious from (1.72), the combined NE are

$$\left[\left(r\boldsymbol{I} - \frac{1}{k}\boldsymbol{N}\boldsymbol{N}'\right) + \frac{1}{k(1+\gamma k)}\,\boldsymbol{N}\boldsymbol{N}'\right]\hat{\hat{\boldsymbol{\tau}}} = \boldsymbol{Q} + \frac{1}{k(1+\gamma k)}\,\boldsymbol{N}\boldsymbol{B} \tag{1.76}$$

The form of (1.76) shows again two things:

1. The intrablock and interblock NE are related.
2. The matrix $\boldsymbol{N}\boldsymbol{N}'$ determines the nature of the estimators, both intrablock and interblock.

The properties of $\boldsymbol{NN}'$ can be exploited to make some statements about estimable functions of treatment effects. Being real symmetric, $\boldsymbol{NN}'$ is orthogonally diagonalizable. We know that

$$\left(r\boldsymbol{I} - \frac{1}{k}\boldsymbol{NN}'\right) \mathfrak{I} = \boldsymbol{0}$$

or

$$\boldsymbol{NN}'\mathfrak{I} = rk\mathfrak{I}$$

so that one root of $\boldsymbol{NN}'$ is $rk \equiv \delta_t$, say, with associated eigenvector $(1/\sqrt{t})\ \mathfrak{I} \equiv \boldsymbol{\xi}_t$, say. Suppose $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_{t-1}$ complete the full set of orthonormal eigenvectors with associated eigenvalues $\delta_1, \delta_2, \ldots, \delta_{t-1}$. Then with

$$\boldsymbol{O} = (\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_t)$$

and

$$\boldsymbol{NN}'\boldsymbol{\xi}_i = \delta_i \boldsymbol{\xi}_i$$

the NE are equivalent to

$$\boldsymbol{O}'\left(r\boldsymbol{I} - \frac{1}{k}\boldsymbol{NN}'\right)\boldsymbol{OO}'\,\widehat{\boldsymbol{\tau}} = \boldsymbol{O}'\boldsymbol{Q} \tag{1.77}$$

and

$$\boldsymbol{O}'\boldsymbol{NN}'\boldsymbol{OO}'\boldsymbol{\tau}^* = \boldsymbol{O}'\boldsymbol{NB} \tag{1.78}$$

respectively. If we write

$$\boldsymbol{\nu} = \begin{bmatrix} \nu_1 \\ \nu_2 \\ \vdots \\ \nu_t \end{bmatrix} = \boldsymbol{O}'\boldsymbol{\tau}$$

then Eq. (1.77) and (1.78) reduce to

$$\left(r - \frac{\delta_i}{k}\right)\widehat{\nu}_i = a_i$$

and

$$\delta_i \nu_i^* = b_i$$

respectively, where

$$\boldsymbol{a} = (a_1, a_2, \ldots, a_t)' = \boldsymbol{O}'\boldsymbol{Q}$$

and

$$\boldsymbol{b} = (b_1, b_2, \ldots, b_t)' = \boldsymbol{O}'\boldsymbol{N}\boldsymbol{B}$$

We see then that we have both intrablock and interblock estimators of ν_i if δ_i is not equal to 0 or to rk. For the component ν_t, only the interblock estimator exists. If other roots are equal to rk, then the intrablock estimators for the corresponding treatment parameters do not exist. Similarly, if other roots are equal to zero, then the interblock estimators for the corresponding treatment parameters do not exist. The treatment parameters $\nu_1, \nu_2, \ldots, \nu_{t-1}$ are necessarily treatment contrasts. If the design is connected, then no $\delta_i (i = 1, 2, \ldots, t-1)$ will equal rk.

The combined NE (1.76) are now transformed to

$$\left[\left(r - \frac{\delta_i}{k}\right) + \frac{1}{k(1+\gamma k)}\,\delta_i\right]\widehat{\widehat{\nu}}_i = a_i + \frac{b_i}{k(1+\gamma k)} \qquad (i = 1, 2, \ldots, t) \quad (1.79)$$

We know that

$$\text{var}(\boldsymbol{Q}) = \left(r\boldsymbol{I} - \frac{1}{k}\boldsymbol{N}\boldsymbol{N}'\right)\sigma_e^2$$

and

$$\begin{aligned}\text{var}(\boldsymbol{N}\boldsymbol{B}) &= \boldsymbol{N}\ \text{var}(\boldsymbol{B})\boldsymbol{N}' \\ &= \boldsymbol{N}\ \frac{k}{\sigma_e^2}(1+\gamma k)\boldsymbol{I}\boldsymbol{N}'\end{aligned}$$

Hence

$$\text{var}(\boldsymbol{a}) = \sigma_e^2 \ \ \text{diag}\left(r - \frac{\delta_i}{k}\right)$$

and

$$\text{var}(\boldsymbol{b}) = \sigma_e^2\ k(1+\gamma k)\text{diag}(\delta_i)$$

So we see from (1.79) that combined estimation of the parameter vector $\boldsymbol{\nu}$ consists of combining intrablock and interblock estimators of components of $\boldsymbol{\nu}$, weighting inversely as their variances.

1.10 ESTIMATION OF WEIGHTS FOR THE COMBINED ANALYSIS

The estimator for the treatment effects as given by (1.61) depends on the weights w and w'_j as can be seen from (1.58) and (1.59). If these weights were known, or alternatively as is apparent from (1.68), if the ratios of the interblock variance and the intrablock variance,

$$\rho_j = \frac{w}{w'_j} = \frac{\sigma_e^2 + k_j \sigma_\beta^2}{\sigma_e^2}$$

were known, then the solution (1.63) to the NE (1.55) would lead to best linear unbiased estimators for estimable functions of treatment effects. Usually, however, these parameters are not known and have to be estimated from the data. If the estimators are used instead of the unknown parameters, then the solutions to the normal equations (1.55) lose some of their good properties. It is for this reason that the properties of the combined estimator have to be examined critically, in particular with regard to their dependence on the type of estimator for the ρ_j's, and with regard to the question of how the combined estimator compares with the intrablock estimator. Before we discuss these questions in some more detail, we shall outline the "classical" procedure for estimating the ρ_j. Since this method was proposed first by Yates (1940a) we shall refer to it as the *Yates procedure* or to the estimators as the *Yates estimators*.

1.10.1 Yates Procedure

One way to estimate w and w'_j is to first estimate σ_e^2 and σ_β^2 and then use these estimators to estimate w and w'_j. If the estimators are denoted by $\widehat{\sigma}_e^2$ and $\widehat{\sigma}_\beta^2$, respectively, then we estimate w and w'_j as

$$\widehat{w} = \frac{1}{\widehat{\sigma}_e^2} \qquad \widehat{w}'_j = \frac{1}{\widehat{\sigma}_e^2 + k_j \widehat{\sigma}_\beta^2} \qquad (j = 1, 2, \ldots, b) \tag{1.80}$$

Obviously, from Table 1.1

$$\widehat{\sigma}_e^2 = \mathrm{MS}(\boldsymbol{I}|\mathfrak{I}, \boldsymbol{X}_\beta, \boldsymbol{X}_{\boldsymbol{\tau}}) \tag{1.81}$$

To estimate σ_β^2 we turn to Table 1.2. Under model (1.50) with covariance structure (1.51) we find [see also (1.39)]

$$E[\mathrm{SS}(\boldsymbol{X}_\beta|\mathfrak{I}, \boldsymbol{X}_\tau)] = (b-1)\sigma_e^2 + \left(n - \sum_{ij} \frac{1}{r_i} n_{ij}^2\right) \sigma_\beta^2 \tag{1.82}$$

Hence it follows from (1.81) and (1.82) that

$$\widehat{\sigma}_\beta^2 = \frac{b-1}{n - \sum \frac{1}{r_i} n_{ij}^2} \left[\mathrm{SS}(\boldsymbol{X}_\beta|\mathfrak{I}, \boldsymbol{X}_\tau) - \mathrm{SS}(\boldsymbol{I}|\mathfrak{I}, \boldsymbol{X}_\beta, \boldsymbol{X}_\tau)\right] \tag{1.83}$$

The estimators (1.81) and (1.83) are then substituted into (1.80) to obtain $\widehat{w}$ and $\widehat{w}_j (j = 1, 2, \ldots, b)$. If in (1.83) $\widehat{\sigma}_\beta^2 \leq 0$ for a given data set, we take

$$\widehat{w}'_j = \widehat{w} = \frac{1}{\mathrm{MS}(\boldsymbol{I}|\mathfrak{I}, \boldsymbol{X}_\beta, \boldsymbol{X}_\tau)}$$

In either case $\widehat{w}$ and $\widehat{w}'_j$ are substituted into (1.58) and (1.59) and hence into the solution (1.61). Also, $\mathrm{var}\left(\widehat{\widehat{\tau}}_i - \widehat{\widehat{\tau}}_{i'}\right)$ is estimated by substituting $\widehat{w}$, $\widehat{w}'_j$ and $\widehat{\sigma}_e^2$ into (1.62) and (1.64).

For alternative estimation procedures see Section 1.11, and for a numerical example see Section 1.14.3.

1.10.2 Properties of Combined Estimators

As we have already pointed out, the fact that the unknown parameters in (1.61) are replaced by their estimators will have an effect on the properties of the estimators for treatment effects. The two properties we are concerned about here are unbiasedness and minimum variance.

Let $\boldsymbol{c}'\boldsymbol{\tau}$ be an estimable function of the treatment effects, let $t = \boldsymbol{c}'\widehat{\boldsymbol{\tau}}$ be its intrablock estimator, $t(\rho) = \boldsymbol{c}'\widehat{\widehat{\boldsymbol{\tau}}}$ its combined (Aitken) estimator with ρ known (for the present discussion we shall confine ourselves to proper designs), and $t(\widehat{\rho}) = \boldsymbol{c}'\widetilde{\boldsymbol{\tau}}$ the combined estimator when in (1.68) ρ is replaced by $\widehat{\rho} = \widehat{w}/\widehat{w}'$.

Roy and Shah (1962) have shown that for the general incomplete block design, although the Yates procedure leads to a biased estimator for ρ, the estimators for treatment contrasts obtained by the method just described are unbiased, that is,

$$E\,[t(\widehat{\rho})] = \boldsymbol{c}'\boldsymbol{\tau}$$

With regard to $\mathrm{var}[t(\widehat{\rho})]$, it is clear that due to sampling fluctuations of $\widehat{\rho}$ we have

$$\mathrm{var}[t(\rho)] < \mathrm{var}[t(\widehat{\rho})]$$

that is, the combined estimators no longer have minimum variance. The crucial question in this context, however, is: When is $\mathrm{var}[t(\widehat{\rho})] < \mathrm{var}(t)$? In other words: When is the combined estimator more efficient than the intrablock estimator?

The answer to this question depends on several things such as (1) the true value of ρ, (2) the type of estimator for ρ, and (3) the number of treatments and blocks. It is therefore not surprising that so far no complete answer has been given.

The most general result for the Yates estimator (and a somewhat larger class of estimators) is that of Shah (1964) based upon some general results by Roy and Shah (1962) [see also Bhattacharya (1998) for proper designs]. It is shown there that the combined estimator for any treatment contrast in any (proper) incomplete block design has variance smaller than that of the corresponding intrablock estimator if ρ does not exceed 2, or, equivalently, if $\sigma_\beta^2 \leq \sigma_e^2/k$. This is a condition that, if blocking is effective at all, one would not in general expect to be satisfied. The problem therefore remains to find methods of constructing estimators for ρ such that the combined estimator for treatment contrasts is uniformly better than the corresponding intrablock estimator, in the sense of having smaller variances for all values of ρ. For certain incomplete block designs this goal has been achieved. We shall mention these results in the following chapters.

The only general advice that we give at this point in conjunction with the use of the Yates estimator is the somewhat imprecise advice to use the intrablock rather than the combined estimator if the number of treatments is "small." The reason for this is that in such a situation the degrees of freedom for $\mathrm{MS}(\boldsymbol{I}|\boldsymbol{\Im}, \boldsymbol{X}_\beta, \boldsymbol{X}_\tau)$ and $\mathrm{MS}(\boldsymbol{X}_\beta|\boldsymbol{\Im}, \boldsymbol{X}_\tau)$ are likely to be small also, which would imply that σ_e^2 and σ_β^2, and hence ρ, cannot be estimated very precisely.

1.11 MAXIMUM-LIKELIHOOD TYPE ESTIMATION

In this section we discuss alternatives to the Yates procedure (see Section 1.10) of estimating the variance components σ_β^2 and σ_e^2 for the combined analysis. These estimators are maximum-likelihood type estimators. This necessitates the assumption of normality, which is not in agreement with our underlying philosophy of finite population randomization analysis. The reason for discussing them, however, is the fact that they can easily be implemented in existing software, in particular SAS PROC MIXED (SAS, 1999–2000) (see Section 1.14).

1.11.1 Maximum-Likelihood Estimation

It is convenient to rewrite model (1.49) with its covariance structure (1.51) in matrix notation as follows:

$$\begin{aligned} \boldsymbol{y} &= \mu\boldsymbol{\Im} + \boldsymbol{X}_\tau\boldsymbol{\tau} + \boldsymbol{X}_\beta\boldsymbol{\beta} + \boldsymbol{e} \\ &= \boldsymbol{X}\boldsymbol{\alpha} + \boldsymbol{U}\boldsymbol{\beta} + \boldsymbol{e} \end{aligned} \tag{1.84}$$

where $\boldsymbol{X\alpha}$ represents the fixed part, with $\boldsymbol{X} = (\boldsymbol{\Im}\ \boldsymbol{X}_\tau)$, $\boldsymbol{\alpha}' = (\mu, \boldsymbol{\tau}')$, and $\boldsymbol{U\beta} + \boldsymbol{e}$ represents the random part. Thus

$$E(\boldsymbol{y}) = \boldsymbol{X\alpha}$$

and

$$\begin{aligned} \mathrm{var}(\boldsymbol{y}) &= \boldsymbol{U}\boldsymbol{U}'\sigma_\beta^2 + \boldsymbol{I}_n\sigma_e^2 \\ &= \boldsymbol{V}\sigma_e^2 \end{aligned} \tag{1.85}$$

with $\boldsymbol{V} = \gamma \boldsymbol{U}\boldsymbol{U}' + \boldsymbol{I}_n$ and $\gamma = \sigma_\beta^2/\sigma_e^2$ [see also (1.53), (1.54)]. We then assume that

$$\boldsymbol{y} \sim N_n(\boldsymbol{X\alpha}, \boldsymbol{V}\sigma_e^2) \tag{1.86}$$

that is, $\boldsymbol{y}$ follows as a multivariate normal distribution (see I.4.17.1).

The logarithm of the likelihood function for $\mathbf{y}$ of (1.86) is then given by

$$\lambda = -\tfrac{1}{2}n \log \pi - \tfrac{1}{2}n\sigma_e^2 - \tfrac{1}{2} \log|\boldsymbol{V}| - \tfrac{1}{2}(\boldsymbol{y} - \boldsymbol{X\alpha})'\boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{X\alpha})/\sigma_e^2 \tag{1.87}$$

Hartley and Rao (1967) show that the maximum-likelihood (ML) estimator of $\boldsymbol{\alpha}$, γ, and σ_e^2 are obtained by solving the following equations:

$$\frac{1}{\sigma_e^2}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y} - \boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X\alpha}) = \boldsymbol{0} \tag{1.88}$$

$$-\frac{1}{2}\,\mathrm{tr}(\boldsymbol{V}^{-1}\boldsymbol{U}\boldsymbol{U}') + \frac{1}{2\sigma_e^2}(\boldsymbol{y} - \boldsymbol{X\alpha})'\boldsymbol{V}^{-1}\boldsymbol{U}\boldsymbol{U}'\boldsymbol{V}^{-1}(By - \boldsymbol{X\alpha}) = 0$$

$$-\frac{n}{2\sigma_e^2} + \frac{1}{2\sigma_e^4}(\boldsymbol{y} - \boldsymbol{X\alpha})'\boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{X\alpha}) = 0$$

where $\mathrm{tr}(\boldsymbol{A})$ represents the trace of the matrix $\boldsymbol{A}$.

The basic feature of this method is that the fixed effects and the variance components associated with the random effects are estimated simultaneously in an iterative procedure. We shall not go into the details of the numerical implementation (see, e.g., Hemmerle and Hartley, 1973), but refer to the example given in Section 1.14.4 using SAS.

1.11.2 Restricted Maximum-Likelihood Estimation

Specifically for the estimation of weights for the purpose of recovery of interblock information, Patterson and Thompson (1971) introduced a modified maximum-likelihood procedure. The basic idea is to obtain estimators for the variance components that are free of the fixed effects in the sense that the likelihood does not contain the fixed effect. Operationally this is accomplished by dividing the likelihood function (1.87) into two parts, one being based on treatment contrasts and the other being based on error contrasts, that is, contrasts with expected value zero. Maximizing this second part will lead to estimates of functions of σ_β^2 and σ_e^2. Because of the procedure employed, these estimates are by some referred to as residual maximum-likelihood estimates (REML), by others as restricted maximum likelihood estimates (REML). The latter name is derived from the fact that maximizing the part of the likelihood free of the fixed effects can be thought of as maximizing the likelihood over a restricted parameter set, an idea first proposed by Thompson (1962) for random effects models and generalized for the general linear mixed model by Corbeil and Searle (1976), based on the work by Patterson and Thompson (1971). We shall give below a brief outline of the basic idea of REML estimation following Corbeil and Searle (1976).

Consider model (1.84) and assume that the observations are ordered by treatments, where the ith treatment is replicated r_i times ($i = 1, 2, \ldots, t$). Then the matrix $\boldsymbol{X}$ can simply be written as

$$\boldsymbol{X} = \begin{bmatrix} \mathfrak{I}_{r_1} & & & \\ & \mathfrak{I}_{r_2} & & 0 \\ & & \ddots & \\ 0 & & & \mathfrak{I}_{r_t} \end{bmatrix} \equiv \sum_{i=1}^{t}{}^{+} \mathfrak{I}_{r_i} \tag{1.89}$$

To separate the log-likelihood function (1.87) into the two parts mentioned above, we employ the transformation (as proposed by Patterson and Thompson, 1971)

$$\boldsymbol{y}'[\boldsymbol{S} \vdots \boldsymbol{V}^{-1}\boldsymbol{X}] \tag{1.90}$$

where

$$\boldsymbol{S} = \boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$$

$$= \sum_{i=1}^{t}{}^{+} \left(\boldsymbol{I}_{r_i} - \frac{1}{r_i} \mathfrak{I}_{r_i} \mathfrak{I}'_{r_i} \right) \tag{1.91}$$

is symmetric and idempotent. Furthermore, $\boldsymbol{SX} = \boldsymbol{0}$, and hence $\boldsymbol{Sy}$ is distributed $N(\boldsymbol{0}, \boldsymbol{SVS}\sigma_e^2)$ independently of $\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y}$.

It follows from (1.91) that $\boldsymbol{S}$ is singular. Hence, instead of $\boldsymbol{S}$ we shall use in (1.90) a matrix, $\boldsymbol{T}$ say, which is derived from $\boldsymbol{S}$ by deleting its r_1th, $(r_1 + r_2)$th, $(r_1 + r_2 + r_3)$th, ..., $(r_1 + r_2 + \cdots + r_t)$th rows, thereby reducing an $n \times n$ matrix to an $(n - t) \times n$ matrix (with $n - t$ representing the number of linearly independent error contrasts). More explicitly, we can write $\boldsymbol{T}$ as

$$\boldsymbol{T} = \sum_{i=1}^{t}{}^{+} \left[\boldsymbol{I}_{r_i-1} \ \vdots \ \boldsymbol{0}_{r_i-1} - \frac{1}{r_i} \mathfrak{I}_{r_i-1} \mathfrak{I}'_{r_i} \right]$$

$$= \sum_{i=1}^{t}{}^{+} \left[\boldsymbol{I}_{r_i-1} - \frac{1}{r_i} \mathfrak{I}_{r_i-1} \mathfrak{I}'_{r_i-1} \ \vdots \ -\frac{1}{r_i} \mathfrak{I}_{r_i-1} \right] \tag{1.92}$$

It follows from (1.89) and (1.92) that

$$\boldsymbol{TX} = \boldsymbol{0} \tag{1.93}$$

Considering now the transformation

$$z = \begin{pmatrix} \boldsymbol{T} \\ \boldsymbol{X}'\boldsymbol{V}^{-1} \end{pmatrix} \boldsymbol{y}$$

it follows from (1.86) and (1.93) that

$$z \sim N\left[\begin{pmatrix} \mathbf{0} \\ \boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}\boldsymbol{\alpha} \end{pmatrix},\ \begin{pmatrix} \boldsymbol{T}\boldsymbol{V}\boldsymbol{T}'\sigma_e^2 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}\sigma_e^2 \end{pmatrix}\right] \tag{1.94}$$

Clearly, the likelihood function of z consists of two parts, one for $\boldsymbol{T}\boldsymbol{y}$, which is free of fixed effects, and one for $\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y}$ pertaining to the fixed effects. In particular, the log likelihood of $\boldsymbol{T}\boldsymbol{y}$ then is

$$\begin{aligned} \lambda_1 = \tfrac{1}{2}(n-t)\log 2\pi - \tfrac{1}{2}(n-t)\log \sigma_e^2 \\ -\tfrac{1}{2}\log|\boldsymbol{T}\boldsymbol{V}\boldsymbol{T}'| - \tfrac{1}{2}\boldsymbol{y}'\boldsymbol{T}'(\boldsymbol{T}\boldsymbol{V}\boldsymbol{T}')^{-1}\boldsymbol{y}/\sigma_e^2 \end{aligned} \tag{1.95}$$

The REML estimators for $\gamma = \sigma_\beta^2/\sigma_e^2$ and σ_e^2 are obtained by solving the equations

$$\frac{\partial \lambda_1}{\partial \gamma} = 0 \tag{1.96}$$

$$\frac{\partial \lambda_1}{\partial \sigma_e^2} = 0 \tag{1.97}$$

The resulting equations have no analytic solutions and have to be solved iteratively. We denote the solutions, that is, estimates by $\widetilde{\gamma}$ and $\widetilde{\sigma}_e^2$, respectively.

The fixed effects, represented by $\boldsymbol{\alpha}$ in (1.84), can be estimated by considering the log likelihood of $\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y}$, which is given by

$$\begin{aligned} \lambda_2 = -\tfrac{1}{2}t\log 2\pi - \tfrac{1}{2}t\ \log \sigma_e^2 \\ -\tfrac{1}{2}\log|\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}| \\ -\tfrac{1}{2}\ (\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\alpha})'\,\boldsymbol{V}^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1}(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\alpha})/\sigma_e^2 \end{aligned} \tag{1.98}$$

Solving

$$\frac{\partial \lambda_2}{\partial \boldsymbol{\alpha}} = 0$$

leads to the estimator

$$\widehat{\boldsymbol{\alpha}} = (\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y} \tag{1.99}$$

This, of course, assumes that $\boldsymbol{V}$ is known. Since it is not, we substitute $\widetilde{\gamma}$ from (1.96) and (1.97) for γ in (1.99) and obtain the estimate

$$\widetilde{\boldsymbol{\alpha}} = (\boldsymbol{X}'\widetilde{\boldsymbol{V}}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\widetilde{\boldsymbol{V}}^{-1}\boldsymbol{y} \tag{1.100}$$

where $\widetilde{\boldsymbol{V}}$ denotes $\boldsymbol{V}$ with γ replaced by $\widetilde{\gamma}$.

An approximate estimate of the variance of $\widetilde{\alpha}$ is given by

$$\widetilde{\text{var}}(\widetilde{\boldsymbol{\alpha}}) \cong (\boldsymbol{X}'\widetilde{\boldsymbol{V}}^{-1}\boldsymbol{X})^{-1}\widetilde{\sigma}_e^2 \tag{1.101}$$

For a numerical example using REML see Section 1.14.4.

1.12 EFFICIENCY FACTOR OF AN INCOMPLETE BLOCK DESIGN

We have seen in Sections I.9.3 and I.10.2.9, for example, how we can compare different error control designs with each other by using the notion of relative efficiency. In this case, we compare two error control designs after we have performed the experiment using a particular error control design. For example, after we have used an RCBD we might ask: How would we have done with a corresponding CRD? In other cases, however, we may want to compare error control designs before we begin an experiment. In particular, we may want to compare an incomplete block design (IBD) with either a CRD or an RCBD, or we may want to compare competing IBDs with each other. For this purpose we shall use a quantity that is referred to as the *efficiency factor* of the IBD. It compares, apart from the residual variance, σ_e^2, the average variance of simple treatment comparisons for the two competing designs.

1.12.1 Average Variance for Treatment Comparisons for an IBD

Let us now consider

$$\underset{i \neq i'}{\text{av.}}\ \text{var}(\widehat{\tau}_i - \widehat{\tau}_{i'}) \tag{1.102}$$

for a connected IBD. We know, of course, that (1.102) is a function of $\boldsymbol{C}^-$, a generalized inverse of the $\boldsymbol{C}$ matrix. Suppose now that all the block sizes are equal to k. Then we have

$$\boldsymbol{C} = \boldsymbol{R} - \frac{1}{k}\boldsymbol{N}\boldsymbol{N}'$$

and we know that $\boldsymbol{C}$ has one zero root, $d_t = 0$ say, with associated normalized eigenvector $\boldsymbol{\xi}_t = (1/\sqrt{t})\mathfrak{I}$. Let the other roots be $d_1, d_2, \ldots, d_{t-1}$ with associated orthonormal eigenvectors $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_{t-1}$. Then

$$\boldsymbol{\xi}_i'\boldsymbol{C} = d_i\boldsymbol{\xi}_i' \qquad (i = 1, 2, \ldots, t-1)$$

and from

$$\boldsymbol{\xi}_i'\boldsymbol{C}\boldsymbol{\tau} = d_i\boldsymbol{\xi}_i'\boldsymbol{\tau}$$

it follows that

$$\widehat{\boldsymbol{\xi}_i'\boldsymbol{\tau}} = \frac{1}{d_i}\boldsymbol{\xi}_i'\boldsymbol{Q}$$

and

$$\operatorname{var}(\boldsymbol{\xi}_i'\widehat{\boldsymbol{\tau}}) = \frac{1}{d_i^2}\boldsymbol{\xi}_i'\boldsymbol{C}\boldsymbol{\xi}_i\sigma_e^2 = \frac{1}{d_i}\sigma_e^2 \tag{1.103}$$

Using the fact that $\boldsymbol{\xi}_t = (1/\sqrt{t})\mathfrak{I}$, that $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_{t-1}$ are mutually perpendicular and perpendicular to $\boldsymbol{\xi}_1$, and that

$$\sum_{i=1}^{t} \boldsymbol{\xi}_i\ \boldsymbol{\xi}_i' = \mathbf{I}$$

we have with $\boldsymbol{z}' = (z_1, z_2, \ldots, z_t)$

$$\begin{aligned}\sum_{i=1}^{t-1} (\boldsymbol{\xi}_i'\boldsymbol{z})^2 &= \boldsymbol{z}' \left(\sum_{i=1}^{t-1} \boldsymbol{\xi}_i\ \boldsymbol{\xi}_i'\right)\boldsymbol{z} \\ &= \boldsymbol{z}'(\boldsymbol{I} - \boldsymbol{\xi}_t\ \boldsymbol{\xi}_t')\boldsymbol{z} \\ &= \sum_{i=1}^{t} z_i^2 - \frac{1}{t}\left(\sum_{i=1}^{t} z_i\right)^2 \\ &= \sum_{i=1}^{t} (z_i - \overline{z})^2 \end{aligned} \tag{1.104}$$

It is also easy to verify that

$$\frac{1}{t(t-1)} \sum_{i \neq i'} (z_i - z_{i'})^2 = \frac{2}{t-1} \sum_{i=1}^{t} (z_i - \overline{z})^2 \tag{1.105}$$

Taking $z_i = \widehat{\tau}_i - \tau_i$, substituting into (1.105) using (1.104) and then taking expectations and using (1.103), yields for (1.102)

$$\underset{i \neq i'}{\text{av.}} \operatorname{var}(\widehat{\tau}_i - \widehat{\tau}_{i'}) = \frac{2}{t-1} \sum_{i=1}^{t-1} \frac{1}{d_i} \sigma_e^2 \tag{1.106}$$

1.12.2 Definition of Efficiency Factor

It is natural in attempting to evaluate the efficiency of an IBD to compare it with a CRD since this is always a possible competing design. For a CRD with r_i replications for treatment i, the average variance of treatment differences is

$$\underset{i \neq i'}{\text{av.}} \operatorname{var}(\widehat{\tau}_i - \widehat{\tau}_{i'}) = \underset{i \neq i'}{\text{av.}} \left(\frac{1}{r_i} + \frac{1}{r_i'}\right) \sigma^2_{e(\text{CRD})} = \frac{2}{\overline{r}_h} \sigma^2_{e(\text{CRD})} \tag{1.107}$$

where $\overline{r}_h$ is the harmonic mean of the r_i, that is,

$$\frac{1}{\overline{r}_h} = \frac{1}{t} \sum_i \frac{1}{r_i}$$

We shall digress here for a moment and show that the best CRD is the one with all $r_i = r$, and that is the design with which we shall compare the IBD. For this and later derivations we need the "old" result that the harmonic mean of a set of positive numbers is not greater than the arithmetic mean. It seems useful to give an elementary proof of this.

Let the set of numbers be $\{x_i, i = 1, 2, \ldots, m\}$. Consider the quadratic

$$q(\beta) = \sum_{i=1}^{m} \left(\sqrt{x_i} - \beta \frac{1}{\sqrt{x_i}}\right)^2$$

Clearly $q(\beta) \geq 0$ for all β. The minimizing value of β is obtained by using least squares which gives the NE

$$\sum_i \frac{1}{x_i} \widetilde{\beta} = m$$

The minimum sum of squares is

$$\sum_i x_i - \widetilde{\beta} m$$

Hence

$$\sum_i x_i - \frac{m^2}{\sum_i \frac{1}{x_i}} \geq 0$$

or

$$\left(\frac{1}{m}\sum_i x_i\right)\left(\frac{1}{m}\sum_i \frac{1}{x_i}\right) \geq 1$$

or

$$\frac{\overline{x}}{\overline{x}_h} \geq 1$$

with equality if and only if $x_i = x$ for all i.

This result implies that the best CRD will have $r_i = r$ and $r = n/t$ where n is the total numbers of EUs. This can happen, of course, only if n/t is an integer. If n/t is not an integer so that $n = pt + q$ $(0 < q < t)$, then the best CRD will have q treatments replicated $p + 1$ times.

Consider now the case of an IBD with b blocks of size k and r_i replications for the ith treatment. Then the total number of EUs is $n = bk = \sum r_i$. Suppose also that $n = rt$, so that an equireplicate CRD is possible. The average variance for such a design is $2\sigma^2_{e(\text{CRD})}/r$, whereas the average variance for the IBD is $2\sigma^2_{e(\text{IBD})}/c$ where, as shown in (1.106), c is the harmonic mean of the positive eigenvalues of $\left(\boldsymbol{R} - (1/k)\ \boldsymbol{NN}'\right)$ (see Kempthorne, 1956). It is natural to write $c = rE$, so that with $\sigma^2_{e(\text{CRD})} = \sigma^2_{e(\text{IBD})}$ we have

$$\frac{\text{av. var}(\widehat{\tau}_i - \widehat{\tau}_{i'})_{\text{CRD}}}{\text{av. var}(\widehat{\tau}_i - \widehat{\tau}_{i'})_{\text{IBD}}} = \frac{2/r}{2/rE} = E \tag{1.108}$$

The quantity E thus defined is called the *efficiency factor* of the IBD. It is clearly a numerical property of the treatment-block configuration only and hence a characteristic of a given IBD.

We add the following remarks:

1. The same definition of E in (1.108) could have been obtained by using the average variance for an RCBD with $b = r$ blocks instead of the average variance for an equireplicate CRD assuming that $\sigma^2_{e(\text{RCBD})} = \sigma^2_{e(\text{IBD})}$.
2. Although E is a useful quantity to compare designs, it does not, of course, give the full story. It compares average variances only under the assumption of equality of residual variances, whereas we typically expect $\sigma^2_{e(\text{IBD})} < \sigma^2_{e(\text{CRD})}$ and $\sigma^2_{e(\text{IBD})} < \sigma^2_{e(\text{RCBD})}$.
3. The efficiency factor pertains only to the intrablock analysis and ignores the interblock information.
4. Each IBD will have associated with it an efficiency factor E. In order to compare two competing IBDs with the same n and with efficiency factors E_1 and E_2, respectively, we would typically choose the one with the higher E value.

1.12.3 Upper Bound for the Efficiency Factor

Using again the fact that the harmonic mean of positive numbers is not greater than the arithmetic mean, we have

$$\begin{aligned}(t-1)c \leq \sum_{i=1}^{t-1} d_i &= \text{trace}\left(\boldsymbol{R} - \frac{1}{k}\boldsymbol{N}\boldsymbol{N}'\right)\\ &= \sum_{i=1}^{t} r_i - \frac{1}{k}\sum_{ij} n_{ij}^2\\ &= n - \frac{1}{k}\sum_{ij} n_{ij}^2 \qquad (1.109)\end{aligned}$$

The largest value of the right-hand side of (1.109) is obtained for the smallest value of $\sum_{ij} n_{ij}^2$. Since n_{ij} is one of the numbers $0, 1, 2, \ldots, k$, the minimum value of $\sum n_{ij}^2$ will be achieved when n of the n_{ij}'s are 1 and the remaining are zero. Since then $n_{ij}^2 = n_{ij}$ and $\sum_j n_{ij} = r_i$, it follows from (1.109) that

$$(t-1)c \leq \sum_i r_i - \frac{1}{k}\sum_i r_i = \frac{k-1}{k}t\overline{r}$$

or, since $c = rE$,

$$E \leq \frac{(k-1)t/(t-1)k}{\overline{r}/r}$$

But since $t\overline{r} = n = tr$, we have finally

$$E \leq \frac{(k-1)t}{(t-1)k} \qquad (1.110)$$

Since for an IBD $k < t$, we can write further

$$E \leq \frac{(k-1)t}{(t-1)k} < 1 \qquad (1.111)$$

We shall see later (see Chapter 2, also Section I.9.8.2) that the upper bound given in (1.110) will be achieved for the balanced incomplete block design.

Sharper upper bounds for certain classes of IBDs are given by Jacroux (1984), Jarrett (1983), Paterson (1983), and Tjur (1990); see also John and Williams (1995).

1.13 OPTIMAL DESIGNS

We have argued in the previous section that in order to compare two designs, d_1 and d_2 say, we may consider their efficiency factors E_1 and E_2, respectively, and choose the design with the higher efficiency factor. In particular, if the efficiency factor of one of those designs achieves the upper bound for that class of designs, we would consider that design to be optimal in some sense. Such considerations have led to the development of the notion of *optimal designs* and to various criteria for optimality. We shall describe briefly some of these criteria.

1.13.1 Information Function

Initial contributions to the formal discussion of optimal designs were made by Wald (1943) and Ehrenfeld (1953). Extending their results, Kiefer (1958) provided a systematic account of different optimality criteria. These can be discussed either in terms of maximizing a suitable function of the information matrix or minimizing a corresponding function of the dispersion matrix of a maximal set of orthonormal treatment contrast estimates.

In the context of our discussion the information matrix is given by $\boldsymbol{C}$ in (1.9). Let $\boldsymbol{P}'\boldsymbol{\tau}$ represent a set of $t-1$ orthonormal contrasts of the treatment effects. Using intrablock information from a connected design d, the estimator for $\boldsymbol{P}'\boldsymbol{\tau}$ is given by $\boldsymbol{P}'\widehat{\boldsymbol{\tau}}$ with $\widehat{\tau}$ from (1.18). The dispersion matrix for $\boldsymbol{P}'\widehat{\boldsymbol{\tau}}$ is then given by

$$\boldsymbol{V}_d\, \sigma_e^2 = \boldsymbol{P}'\boldsymbol{C}_d^- \boldsymbol{P}\sigma_e^2 = (\boldsymbol{P}'\boldsymbol{C}_d\boldsymbol{P})^{-1}\sigma_e^2 \tag{1.112}$$

[see (1.20)], where $\boldsymbol{C}_d^-$ and hence $\boldsymbol{C}_d$ refer to the specific design d used. The information matrix for the design d is then defined as

$$\boldsymbol{C}_d^* = (\boldsymbol{P}'\boldsymbol{C}_d^-\boldsymbol{P})^{-1} \tag{1.113}$$

which shows, of course, the connection between $\boldsymbol{C}_d^*$ and $\boldsymbol{C}_d$.

An *information function* or *optimality criterion* is then a real-valued function ϕ that has the following properties (see Pukelsheim, 1993):

1. Function ϕ is a monotonic function; that is, an information matrix $\boldsymbol{C}^*$ is at least as good as another information matrix $\boldsymbol{D}^*$ if $\phi(\boldsymbol{C}^*) \geq \phi(\boldsymbol{D}^*)$;
2. Function ϕ is a concave function, that is, $\phi[(1-\alpha)\,\boldsymbol{C}^* + \alpha\boldsymbol{D}^*] = (1-\alpha)\phi(\boldsymbol{C}^*) + \alpha\,\phi(\boldsymbol{D}^*)$ for $\alpha\ \in (0,\,1)$;
3. Function ϕ is positively homogeneous, that is, $\phi(\delta\boldsymbol{C}^*) = \delta\,\phi(\boldsymbol{C}^*)$.

Condition (2) says that information cannot be increased by interpolation. And condition (3) says that even if we define the information matrix to be directly proportional to the number of observations, n, and inversely proportional to σ_e^2, that is, the information matrix is of the form $(n/\sigma_e^2)\boldsymbol{C}^*$, we need to consider only $\boldsymbol{C}^*$.

Let $\mathcal{D}$ be the set of competing designs. The problem of finding an optimal design d in $\mathcal{D}$ can then be reduced to finding a design that maximizes $\phi(\boldsymbol{C}_d^*)$ over d in $\mathcal{D}$ (see Cheng, 1996). Such a design is called ϕ-optimal.

As indicated above, an alternative, and historically original, approach to finding an optimal design is to consider minimization of some convex and nonincreasing function Φ of dispersion matrices, as indicated by the relationship between (1.112) and (1.113). Accordingly, we shall then talk about a Φ-optimal design.

1.13.2 Optimality Criteria

Several optimality criteria, that is, several functions ϕ or Φ have been considered for studying optimal designs. These criteria can be expressed conveniently in terms of the eigenvalues of $\boldsymbol{C}_d^*$ or, equivalently, the nonzero eigenvalues of $\boldsymbol{C}_d$, say $\mu_{d1} \geq \mu_{d2} \geq \cdots \geq \mu_{d,t-1}$.

The most commonly used optimality criteria are D-, A-, and E-optimality, which maximize the following information functions:

1. *D-optimality*: Determinant criterion or

$$\phi(\boldsymbol{C}_d^*) = \prod_{i=1}^{t-1} \mu_{di}$$

2. *A-optimality*: Average variance criterion or

$$\phi(\boldsymbol{C}_d^*) = \left(\frac{1}{t-1} \sum_{i=1}^{t-1} \mu_{di}^{-1} \right)^{-1}$$

3. *E-optimality*: Smallest eigenvalue criterion or

$$\phi(\boldsymbol{C}_d^*) = \mu_{d,t-1}$$

In terms of the corresponding Φ function, these optimality criteria can be expressed as minimizing

1. $$\Phi(\boldsymbol{V}_d) = \det \boldsymbol{V}_d = \prod_{i=1}^{t-1} \mu_{di}^{-1} \tag{1.114}$$

2. $$\Phi(\boldsymbol{V}_d) = \operatorname{tr} \boldsymbol{V}_d = \sum_{i=1}^{t-1} \mu_{di}^{-1} \tag{1.115}$$

3. $$\Phi(\boldsymbol{V}_d) = \text{maximum eigenvalue of } \boldsymbol{V}_d = \mu_{d,t-1}^{-1} \tag{1.116}$$

The statistical meaning of these criteria is that minimizing (1.114) minimizes the generalized variance of $\boldsymbol{P}'\widehat{\boldsymbol{\tau}}$, (1.115) minimizes the average variance of the set $\boldsymbol{P}'\widehat{\boldsymbol{\tau}}$, and (1.116) minimizes the maximum variance of a single normalized contrast.

1.13.3 Optimal Symmetric Designs

There exist special classes of designs for which all the nonzero eigenvalues of the information matrix $\boldsymbol{C}_d$ are equal. Such designs are called *symmetric designs*. Examples of symmetric designs are balanced incomplete block designs (Chapter 2), Latin square designs (Section I.10.2), Youden squares (Section I.10.5), and so forth. The information matrix of a symmetric design is of the form $a\boldsymbol{I} + b\,\mathfrak{I}\mathfrak{I}'$, which is referred to as a *completely symmetric matrix*.

In general, if a design is Φ_1 optimal, it may not be Φ_2 optimal for two different optimality criteria Φ_1 and Φ_2. However, for symmetric designs Kiefer (1958) showed that they are A-, D-, and E-optimal. This led to the definition of *universal optimality* (Kiefer, 1975a) or *Kiefer optimality* (Pukelsheim, 1993).

Definition 1.1 (Kiefer, 1975a) Let $\mathcal{B}_{t,0}$ be the set of all $t \times t$ nonnegative definite matrices with zero row and column sums, and let Φ be a real-valued function on $\mathcal{B}_{t,0}$ such that

- **(a)** Φ is convex,
- **(b)** $\Phi(\delta\boldsymbol{C})$ is nonincreasing in the scalar $\delta \geq 0$, and
- **(c)** Φ is invariant under each simultaneous permutation of rows and columns.

A design d^* is called *universally optimal* in $\mathcal{D}$ if d^* minimizes $\Phi(\boldsymbol{C}_d)$ for every Φ satisfying conditions (a), (b), and (c). □

To help identify universally optimal designs we have the following fundamental theorem.

Theorem 1.2 (Kiefer, 1975a) Suppose a class $\mathcal{C} = \{\boldsymbol{C}_d, d \in \mathcal{D}\}$ of matrices in $\mathcal{B}_{t,0}$ contains a $\boldsymbol{C}_{d^*}$ for which

- **(a)** $\boldsymbol{C}_{d^*}$ is completely symmetric, and
- **(b)** $\operatorname{tr} \boldsymbol{C}_{d^*} = \max_{d \in \mathcal{D}} \operatorname{tr} \boldsymbol{C}_d$,

Then d^* is universally optimal in $\mathcal{D}$.

1.13.4 Optimality and Research

We have just discussed the notion of design optimality and some of the available optimality criteria. Other criteria have been introduced in other contexts, in particular, in the area of regression or response surface designs (see, e.g., Atkinson and Donev, 1992). And thus optimality has become a powerful concept, but we need to remember that, although it has statistical meaning, it is a mathematical concept. It has a definite and restricted connotation and it may be difficult to apply it in the larger context of designing a "very good" experiment for a researcher who has a scientific or technological problem.

An immediate difficulty is that there is no simple classification of real problems. There are discovery problems, for example, finding the point in factor space at which yield is maximum. There are exploration problems, for example, to obtain a "good" representation of the nature of the dependence of a noisy dependent variable on a given set of independent variables. There is the mathematical problem in that same context that the dependence is known to be of a definite functional form with some specified but unknown parameters, which are to be determined from observations at some locations in the factor space. A common problem in technology and some scientific areas is what is called screening of factors. It is useful and important to think about this overall picture because there is a tendency to interpret the term "optimality of design" in a very limited context, a context that is very valuable, but misleading, in the sphere of total human investigation.

The moral of the situation is multifold: (1) Researchers have to make a choice about problems and often work on unrealistic ones as the closest workable approximation to real live problems and should not be criticized for so doing; (2) almost any optimality problem is to some extent artificial and limited because criteria of value of designs must be introduced, and in almost any investigative situation it is difficult to map the possible designs valuewise into the real line; and (3) a solution to a mathematically formulated problem may have limited value, so to promote one design that is optimal only with respect to a particular criterion of value, C_1, and to declare another design to be of poor value because it is not optimal may be unfair because that design may be better with respect to a different criterion of value, C_2 say. And for one researcher C_1 may be irrelevant, whereas C_2 may be more appropriate.

Considerations of optimality involve, of course, comparison of designs. But how does one do this when error reduction needs to be taken into account? For example, how does one compare the randomized complete block and the Latin square design? Or how does one compare designs when different aspects of statistical inference are involved? This was at the basis of heated discussion between Neyman, who was interested in hypothesis testing, and Fisher and Yates, who were interested in precision of estimation (see Neyman, Iwaszkiewicz, and Kolodziejczyk, 1935).

Informal optimality considerations early on gave probably rise to the heuristic (or perhaps mathematical) idea of symmetry and balancedness, and we shall encounter these characteristics throughout much of the book. Even though these properties do not always guarantee optimality in many cases they lead to near optimality. And from a practical point of view that may be good enough. On the other hand, if a balanced design is an optimal design, but we cannot use that design because of practical constraints and need to use instead a near-balanced design, then we have a way to evaluate the efficiency of the design we are going to use. For example, we know that a balanced incomplete block design (Chapter 2) is optimal. However, we cannot use the design and need to use a partially balanced incomplete block design (Chapter 4). We may then choose, if possible, a design with efficiency "close" to that of the optimal design.

Thus, the insistence on an optimal design may be frustrating for the user because practical reasons may dictate otherwise and because an experimenter rarely has one criterion of value. Rather, he has many criteria of value and is in a mathematical programming situation in which he wishes the design to have "reasonable" efficiencies with respect to criteria $C_1, C_2, \ldots, C_K$. The dilemma is often that the design that is optimal with respect to C_1 is completely nonoptimal with respect to C_2.

In summary, mathematical ideals and requirements of empirical science do not always meet, but it is worth trying to find common ground. In the end, practical considerations may dictate compromises in many instances of experimental research.

1.14 COMPUTATIONAL PROCEDURES

In this section we shall discuss some computational aspects of performing the intrablock analysis and the combined intra- and interblock analysis, mainly in the context of SAS procedures (SAS, 2000). For the intrablock analysis (see Section 1.3) we shall use SAS PROC GLM, and for the combined analysis (see Section 1.8) we shall use SAS PROC MIXED.

1.14.1 Intrablock Analysis Using SAS PROC GLM

Consider the following data set IBD (Table 1.7) with $t = 4$ treatments in $b = 5$ blocks of size $k = 2$, such that treatments 1 and 4 are replicated 3 times, and treatments 2 and 3 are replicated 2 times. An example might be 5 pairs of identical twins representing the blocks, each twin being an experimental unit to whom different drugs are assigned according to the given plan.

The SAS PROC GLM input statements for the intrablock analysis and the results are given in Table 1.8. We shall comment briefly on some aspects of the SAS output (see Table 1.8):

1. The coefficient matrix as well as the right-hand side (RHS) of (1.5) are given under the heading "The $\boldsymbol{X}'\boldsymbol{X}$ Matrix."
2. A generalized inverse for the coefficient matrix $\boldsymbol{X}'\boldsymbol{X}$ is obtained by first eliminating the rows and columns for β_5 and τ_4 from $\boldsymbol{X}'\boldsymbol{X}$ as a consequence of imposing the conditions $\beta_5^* = 0$ and $\tau_4^* = 0$ (we shall denote the SAS solutions to the NE by $\boldsymbol{\beta}^*$ and $\boldsymbol{\tau}^*$). The reduced matrix is of full rank and thus can be inverted. The inverted matrix is restored to the original dimension by inserting zeros in the rows and columns corresponding to β_5^* and τ_4^*. This matrix, together with $\widehat{\sigma}_e^2$ =MS(Error) from the ANOVA table, can be used to find the standard errors for the estimators of estimable functions for treatment effects.

Table 1.7 Data for Incomplete Block Design (t = 4, b = 5, k = 2, r1 = 3, r2 = r3 = 2, r4 = 3)

```
options pageno=1 nodate;
data IBD;
input TRT BLOCK Y @@;
datalines;
1 1 10  2 1 12
3 2 23  4 2 28
1 3 13  3 3 27
2 4 14  4 4 20
1 5 15  4 5 32
;
run;

proc print data=IBD;
title1 'TABLE 1.7';
title2 'DATA FOR INCOMPLETE BLOCK DESIGN';
title3 '(t=4, b=5, k=2, r1=3, r2=r3=2, r4=3)';
run;
```

Obs	TRT	BLOCK	Y
1	1	1	10
2	2	1	12
3	3	2	23
4	4	2	28
5	1	3	13
6	3	3	27
7	2	4	14
8	4	4	20
9	1	5	15
10	4	5	32

3. The general form of an estimable function for treatment effects is given by

$$L7\tau_1 + L8\tau_2 + L9\tau_3 - (L7 + L8 + L9)\tau_4$$

for any values of $L7$, $L8$, and $L9$; that is, only contrasts are estimable.

4. The general form of estimable functions can also be used to identify the solutions to the NE by putting sequentially (and in order) each $Li = 1$ and the remaining $Lj = 0, (j \neq i)$. For example, the "Estimate" of "Intercept" is actually the estimate of $\mu + \beta_5 + \tau_4$; that is, $L1 = 1, Lj = 0, (j \neq 1)$. Expressed in terms of the SAS solutions we thus have

$$\mu^* = \widehat{\mu + \beta_5 +}\, \tau_4 = 31.125$$

Another example, putting $L7 = 1, Lj = 0 (j \neq 7)$, yields

$$\tau_1^* = \widehat{\tau_1 - \tau_4} = -15.25$$

Table 1.8 Intrablock Analysis with Post-hoc Comparisons

```
proc glm data=IBD;
class BLOCK TRT;
model Y = BLOCK TRT/XPX inverse solution e;

1smeans TRT/stderr e;

estimate 'TRT1 - (TRT2+TRT3)/2' TRT 1 -.5 -.5 0;
estimate 'TRT1 - TRT4' TRT 1 0 0 -1;
estimate 'TRT2 - TRT3' TRT 0 1 -1 0;
title1 'TABLE 1.8';
title2 'INTRA-BLOCK ANALYSIS';
title3 'WITH POST-HOC COMPARISONS';
run;
```

The GLM Procedure

Class Level Information

Class	Levels	Values
BLOCK	5	1 2 3 4 5
TRT	4	1 2 3 4

Number of observations 10

The X'X Matrix

	Intercept	BLOCK 1	BLOCK 2	BLOCK 3	BLOCK 4	BLOCK 5
Intercept	10	2	2	2	2	2
BLOCK 1	2	2	0	0	0	0
BLOCK 2	2	0	2	0	0	0
BLOCK 3	2	0	0	2	0	0
BLOCK 4	2	0	0	0	2	0
BLOCK 5	2	0	0	0	0	2
TRT 1	3	1	0	1	0	1
TRT 2	2	1	0	0	1	0
TRT 3	2	0	1	1	0	0
TRT 4	3	0	1	0	1	1
Y	194	22	51	40	34	47

Table 1.8 (*Continued*)

The X'X Matrix

	TRT 1	TRT 2	TRT 3	TRT 4	Y
Intercept	3	2	2	3	194
BLOCK 1	1	1	0	0	22
BLOCK 2	0	0	1	1	51
BLOCK 3	1	0	1	0	40
BLOCK 4	0	1	0	1	34
BLOCK 5	1	0	0	1	47
TRT 1	3	0	0	0	38
TRT 2	0	2	0	0	26
TRT 3	0	0	2	0	50
TRT 4	0	0	0	3	80
Y	38	26	50	80	4300

X'X Generalized Inverse (g2)

	Intercept	BLOCK 1	BLOCK 2	BLOCK 3	BLOCK 4	BLOCK 5
Intercept	0.75	-0.375	-0.625	-0.375	-0.625	0
BLOCK 1	-0.375	1.3125	0.4375	0.5625	0.6875	0
BLOCK 2	-0.625	0.4375	1.3125	0.6875	0.5625	0
BLOCK 3	-0.375	0.5625	0.6875	1.3125	0.4375	0
BLOCK 4	-0.625	0.6875	0.5625	0.4375	1.3125	0
BLOCK 5	0	0	0	0	0	0
TRT 1	-0.5	-0.25	0.25	-0.25	0.25	0
TRT 2	-0.25	-0.625	0.125	-0.125	-0.375	0
TRT 3	-0.25	-0.125	-0.375	-0.625	0.125	0
TRT 4	0	0	0	0	0	0
Y	31.125	-7.6875	-4.0625	-1.9375	-9.3125	0

X'X Generalized Inverse (g2)

	TRT 1	TRT 2	TRT 3	TRT 4	Y
Intercept	-0.5	-0.25	-0.25	0	31.125
BLOCK 1	-0.25	-0.625	-0.125	0	-7.6875
BLOCK 2	0.25	0.125	-0.375	0	-4.0625
BLOCK 3	-0.25	-0.125	-0.625	0	-1.9375
BLOCK 4	0.25	-0.375	0.125	0	-9.3125
BLOCK 5	0	0	0	0	0
TRT 1	1	0.5	0.5	0	-15.25
TRT 2	0.5	1.25	0.25	0	-9.625
TRT 3	0.5	0.25	1.25	0	-3.125
TRT 4	0	0	0	0	0
Y	-15.25	-9.625	-3.125	0	18.1875

Table 1.8 (*Continued*)

General Form of Estimable Functions

Effect		Coefficients
Intercept		L1
BLOCK	1	L2
BLOCK	2	L3
BLOCK	3	L4
BLOCK	4	L5
BLOCK	5	L1-L2-L3-L4-L5
TRT	1	L7
TRT	2	L8
TRT	3	L9
TRT	4	L1-L7-L8-L9

The GLM Procedure

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	518.2125000	74.0303571	8.14	0.1137
Error	2	18.1875000	9.0937500		
Corrected Total	9	536.4000000			

R-Square	Coeff Var	Root MSE	Y Mean
0.966093	15.54425	3.015585	19.40000

Source	DF	Type I SS	Mean Square	F Value	Pr > F
BLOCK	4	261.4000000	65.3500000	7.19	0.1259
TRT	3	256.8125000	85.6041667	9.41	0.0975

Source	DF	Type III SS	Mean Square	F Value	Pr > F
BLOCK	4	79.1458333	19.7864583	2.18	0.3388
TRT	3	256.8125000	85.6041667	9.41	0.0975

Table 1.8 (***Continued***)

Parameter		Estimate	Standard Error	t Value	Pr > |t|
Intercept		31.12500000 B	2.61157280	11.92	0.0070
BLOCK	1	-7.68750000 B	3.45478608	-2.23	0.1560
BLOCK	2	-4.06250000 B	3.45478608	-1.18	0.3607
BLOCK	3	-1.93750000 B	3.45478608	-0.56	0.6314
BLOCK	4	-9.31250000 B	3.45478608	-2.70	0.1145
BLOCK	5	0.00000000 B	.	.	.
TRT	1	-15.25000000 B	3.01558452	-5.06	0.0369
TRT	2	-9.62500000 B	3.37152599	-2.85	0.1039
TRT	3	-3.12500000 B	3.37152599	-0.93	0.4518
TRT	4	0.00000000 B	.	.	.

NOTE: The X'X matrix has been found to be singular, and a generalized inverse was used to solve the normal equations. Terms whose estimates are followed by the letter 'B' are not uniquely estimable.

Least Squares Means

Coefficients for TRT Least Square Means

TRT Level

Effect		1	2	3	4
Intercept		1	1	1	1
BLOCK	1	0.2	0.2	0.2	0.2
BLOCK	2	0.2	0.2	0.2	0.2
BLOCK	3	0.2	0.2	0.2	0.2
BLOCK	4	0.2	0.2	0.2	0.2
BLOCK	5	0.2	0.2	0.2	0.2
TRT	1	1	0	0	0
TRT	2	0	1	0	0
TRT	3	0	0	1	0
TRT	4	0	0	0	1

TRT	Y LSMEAN	Standard Error	Pr > |t|
1	11.2750000	1.9774510	0.0294
2	16.9000000	2.6632921	0.0239
3	23.4000000	2.6632921	0.0127
4	26.5250000	1.9774510	0.0055

Dependent Variable: Y

Parameter	Estimate	Standard Error	t Value	Pr > |t|
TRT1 - (TRT2+TRT3)/2	-8.8750000	2.61157280	-3.40	0.0768
TRT1 - TRT4	-15.2500000	3.01558452	-5.06	0.0369
TRT2 - TRT3	-6.5000000	4.26468053	-1.52	0.2670

5. The top part of the ANOVA table provides the partition

$$\text{SS(MODEL)} + \text{SS(ERROR)} = \text{SS(TOTAL)}$$

and from that produces

$$\text{MS(ERROR)} = \widehat{\sigma}_e^2 = 9.094$$

6. The lower part of the ANOVA table provides type I SS (sequential SS for ordered model; see I.4.7.2) and type III SS (partial SS). From the latter we obtain the P value (.0975) for the test of

$$H_0: \quad \tau_1 = \tau_2 = \cdots = \tau_t$$

versus

$$H_1: \quad \text{not all } \tau_i \text{ are the same}$$

We stress that the P value for blocks should be ignored (see I.9.2).

7. The e-option of LSMEANS gives the coefficients for the solution vector to compute the treatment least squares means, for example,

$$\begin{aligned}\text{LSMEAN(TRT}\quad 1) &= \mu^* + .2\sum \beta_i^* + \tau_1^* \\ &= 31.125 + .2(-7.6875 - 4.0625 - 1.9375 \\ &\quad - 9.3125 + 0) - 15.25 \\ &= 11.275\end{aligned}$$

The standard error is computed by making use of the G inverse (see item 2) and MS(ERROR).

8. The t tests are performed for the prespecified contrasts among the least-squares means; for example,

$$\begin{aligned}\text{TRT2-TRT3} &= \tau_2^* - \tau_3^* = 9.625 + 3.125 = -\ 6.5 \\ se(\tau_2^* - \tau_3^*) &= [(1.25 + 1.25 - 2 \times .25) \times 9.094]^{1/2} = 4.265 \\ t &= \frac{-\ 6.5}{4.265} = -1.52\end{aligned}$$

1.14.2 Intrablock Analysis Using the Absorb Option in SAS PROC GLM

A computational method as described in Section 1.3 using the RNE can be implemented in SAS PROC GLM by using the ABSORB OPTION. This is illustrated in Table 1.9.

Table 1.9 Intra Block Analysis Using Reduced Normal Equations With Post-hoc Comparisons

```
proc glm data=IBD;
class BLOCK TRT;
absorb BLOCK;
model Y = TRT/XPX inverse solution e;

estimate 'TRT1 - (TRT2+TRT3)/2' TRT 1 -.5 -.5 0;
estimate 'TRT1 - TRT4' TRT 1 0 0 -1;
estimate 'TRT2 - TRT3' TRT 0 1 -1 0;
title1 'TABLE 1.9';
title2 'INTRA-BLOCK ANALYSIS USING REDUCED NORMAL EQUATIONS';
title3 'WITH POST-HOC COMPARISONS';
run;
```

The GLM Procedure

Class Level Information

Class	Levels	Values
BLOCK	5	1 2 3 4 5
TRT	4	1 2 3 4

Number of observations 10

The X'X Matrix

	TRT 1	TRT 2	TRT 3	TRT 4	Y
TRT 1	1.5	-0.5	-0.5	-0.5	-16.5
TRT 2	-0.5	1	0	-0.5	-2
TRT 3	-0.5	0	1	-0.5	4.5
TRT 4	-0.5	-0.5	-0.5	1.5	14
Y	-16.5	-2	4.5	14	275

The GLM Procedure

X'X Generalized Inverse (g2)

	TRT 1	TRT 2	TRT 3	TRT 4	Y
TRT 1	1	0.5	0.5	0	-15.25
TRT 2	0.5	1.25	0.25	0	-9.625
TRT 3	0.5	0.25	1.25	0	-3.125
TRT 4	0	0	0	0	0
Y	-15.25	-9.625	-3.125	0	18.1875

General Form of Estimable Functions
Given that the coefficients for all absorbed effects are zero

Table 1.9 (*Continued*)

Effect		Coefficients
TRT	1	L1
TRT	2	L2
TRT	3	L3
TRT	4	-L1-L2-L3

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	518.2125000	74.0303571	8.14	0.1137
Error	2	18.1875000	9.0937500		
Corrected Total	9	536.4000000			

R-Square	Coeff Var	Root MSE	Y Mean
0.966093	15.54425	3.015585	19.40000

Source	DF	Type I SS	Mean Square	F Value	Pr > F
BLOCK	4	261.4000000	65.3500000	7.19	0.1259
TRT	3	256.8125000	85.6041667	9.41	0.0975

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TRT	3	256.8125000	85.6041667	9.41	0.0975

Parameter	Estimate	Standard Error	t Value	Pr > |t|
TRT1 - (TRT2+TRT3)/2	-8.8750000	2.61157280	-3.40	0.0768
TRT1 - TRT4	-15.2500000	3.01558452	-5.06	0.0369
TRT2 - TRT3	-6.5000000	4.26468053	-1.52	0.2670

Parameter		Estimate	Standard Error	t Value	Pr > |t|
TRT	1	-15.25000000 B	3.01558452	-5.06	0.0369
TRT	2	-9.62500000 B	3.37152599	-2.85	0.1039
TRT	3	-3.12500000 B	3.37152599	-0.93	0.4518
TRT	4	0.00000000 B	.	.	.

NOTE: The X'X matrix has been found to be singular, and a generalized inverse was used to solve the normal equations. Terms whose estimates are followed by the letter 'B' are not uniquely estimable.

We make the following comments about the SAS output:

1. The $\boldsymbol{X}'\boldsymbol{X}$ matrix is now the $\boldsymbol{C}$-matrix of (1.9).
2. The $\boldsymbol{X}'\boldsymbol{X}$ generalized inverse is obtained by the SAS convention of setting $\tau_4^* = 0$. This g inverse is therefore different from $\widetilde{\boldsymbol{C}}^{-1}$ of (1.16).
3. The ANOVA table provides the same information as in Table 1.8, except that it does not give solutions for the intercept and blocks. Hence, this analysis cannot be used to obtain treatment least-squares means and their standard error.

1.14.3 Combined Intra- and Interblock Analysis Using the Yates Procedure

Using SAS PROC VARCOMP illustrates the estimation of σ_β^2 according to the method described in Section 1.10.1. The result is presented in Table 1.10. The option type I produces Table 1.2 with

$$\mathrm{E[MS(BLOCK)]} = \mathrm{E}\left[\mathrm{MS}(\boldsymbol{X}_{\boldsymbol{\beta}} | \mathfrak{I}, \boldsymbol{X}_{\boldsymbol{\tau}})\right]$$

as given in (1.82). This yields $\widehat{\sigma}_e^2 = 9.09$ (as in Table 1.8) and $\widehat{\sigma}_\beta^2 = 7.13$.

Substituting $\widehat{\boldsymbol{\rho}} = \dfrac{9.09 + 2 \times 7.13}{9.09} = 2.57$ into (1.60) we obtain

$$\widetilde{\boldsymbol{A}} = \begin{pmatrix} 2.0814 & -0.3062 & -0.3062 & -0.3062 \\ -0.3062 & 1.3877 & 0.0000 & -0.3062 \\ -0.3062 & 0.0000 & 1.3877 & -0.3062 \\ -0.3062 & -0.3062 & -0.3062 & 2.0814 \end{pmatrix}$$

and

$$\widetilde{\boldsymbol{A}}^{-1} = \begin{pmatrix} 0.541654 & 0.146619 & 0.146619 & 0.122823 \\ 0.146619 & 0.785320 & 0.064704 & 0.146619 \\ 0.146619 & 0.064704 & 0.785320 & 0.146619 \\ 0.122823 & 0.146619 & 0.146619 & 0.541654 \end{pmatrix}$$

and

$$\widetilde{\boldsymbol{P}} = \begin{pmatrix} 4.6242 \\ 8.8528 \\ 22.1358 \\ 39.5816 \end{pmatrix}$$

Table 1.10 Estimation of Block Variance Component Using the Yates Procedure

```
proc varcomp data=IBD method=type1;
class BLOCK TRT;
model Y = TRT BLOCK/fixed=1;
title1 'TABLE 1.10';
title2 'ESTIMATION OF BLOCK VARIANCE COMPONENT';
title3 'USING THE YATES PROCEDURE';
run;
```

```
                 Variance Components Estimation Procedure

                        Class Level Information

                     Class     Levels     Values

                     BLOCK          5     1 2 3 4 5

                     TRT            4     1 2 3 4

                       Number of observations 10

                        Dependent Variable: Y

                      Type 1 Analysis of Variance

                            Sum of
Source             DF      Squares     Mean Square     Expected Mean Square

TRT                 3   439.066667      146.355556     Var(Error) + 0.6667 Var(BLOCK) + Q(TRT)
BLOCK               4    79.145833       19.786458     Var(Error) + 1.5 Var(BLOCK)
Error               2    18.187500        9.093750     Var(Error)
Corrected Total     9   536.400000               .     .

                           Type 1 Estimates

                     Variance Component     Estimate

                     Var(BLOCK)              7.12847
                     Var(Error)              9.09375
```

We then obtain

$$\widetilde{\boldsymbol{\tau}} = \widetilde{\boldsymbol{A}}^{-1}\widetilde{\boldsymbol{P}} = \begin{pmatrix} 11.9097 \\ 14.8659 \\ 24.4379 \\ 26.5510 \end{pmatrix}$$

We note that the elements of $\widetilde{\boldsymbol{\tau}}$ are actually the treatment least-squares means. Their estimated variances and the estimated variances for the treatment contrasts are obtained from $\widetilde{\boldsymbol{A}}^{-1} \times 9.09$ (see Tables 1.13 and 1.14).

1.14.4 Combined Intra- and Interblock Analysis Using SAS PROC MIXED

We illustrate here the numerical implementation of the ML and REML procedures as described in Sections 1.11.1 and 1.11.2, respectively, using SAS PROC MIXED. The results of the ML estimation are given in Table 1.11.

It takes four interations to obtain the solutions, yielding $\widehat{\sigma}_\beta^2 = 7.45$ and $\widehat{\sigma}_e^2 = 4.14$ and hence $\widehat{\gamma} = 1.80$ (Notice that these are quite different from the estimates obtained by the Yates procedure (Section 1.14.3) and the REML procedure as given below).

Since SAS uses a different parametrization than the one used in (1.84) it obtains "estimates" of μ and $\tau_i (i = 1, 2, 3, 4)$ separately. The type 3 coefficients indicate that the solutions for $\mu, \tau_i (i = 1, 2, 3, 4)$ are actually estimates of $\mu + \tau_4, \tau_1 - \tau_4, \tau_2 - \tau_4, \tau_3 - \tau_4$, respectively. From these solutions the least-squares means are then obtained as

$$\begin{aligned}
\text{LSMEAN(TRT 1)} = \widehat{\mu} + \widehat{\tau}_1 &= 11.65 = \widehat{\alpha}_1 \\
\text{LSMEAN(TRT 2)} = \widehat{\mu} + \widehat{\tau}_2 &= 15.63 = \widehat{\alpha}_2 \\
\text{LSMEAN(TRT 3)} = \widehat{\mu} + \widehat{\tau}_3 &= 24.09 = \widehat{\alpha}_3 \\
\text{LSMEAN(TRT 4)} = \widehat{\mu} &= 26.53 = \widehat{\alpha}_4
\end{aligned}$$

where the $\widehat{\alpha}_i$ denote the solutions to (1.87).

The REML procedure is illustrated in Table 1.12. It takes three iterations to obtain the estimates $\widetilde{\sigma}_\beta^2 = 6.35$ and $\widetilde{\sigma}_e^2 = 10.17$, and hence $\widetilde{\gamma} = 0.62$. We note that the REML and ML least-squares means are numerically quite similar even though $\widehat{\gamma}$ and $\widehat{\sigma}_e^2$ are substantially different from $\widetilde{\gamma}$ and $\widetilde{\sigma}_e^2$, respectively.

1.14.5 Comparison of Estimation Procedures

For a small proper incomplete block design we have employed the above four methods of estimating treatment least-squares means and treatment comparisons:

M1: Intrablock analysis

M2: Combined analysis: Yates

Table 1.11 Combined Analysis Using Maximum Likelihood With Post-hoc Comparisons

```
proc mixed data=IBD method=ML;

class BLOCK TRT;
model Y = TRT/ solution E3 ddfm=Satterth;

random BLOCK;
lsmeans TRT;
estimate 'TRT1 - (TRT2+TRT3)/2' TRT 1 -.5 -.5 0;
estimate 'TRT1 - TRT4' TRT 1 0 0 -1;
estimate 'TRT2 - TRT3' TRT 0 1 -1 0;
title1 'TABLE 1.11';
title2 'COMBINED ANALYSIS USING MAXIMUM LIKELIHOOD';
title3 'WITH POST-HOC COMPARISONS';
run;
                         The Mixed Procedure

                          Model Information

          Data Set                      WORK.IBD
          Dependent Variable            Y
          Covariance Structure          Variance Components
          Estimation Method             ML
          Residual Variance Method      Profile
          Fixed Effects SE Method       Model-Based
          Degrees of Freedom Method     Satterthwaite


                       Class Level Information

                    Class     Levels     Values

                    BLOCK          5     1 2 3 4 5
                    TRT            4     1 2 3 4


                             Dimensions

                     Covariance Parameters      2
                     Columns in X               5
                     Columns in Z               5
                     Subjects                   1
                     Max Obs Per Subject       10
                     Observations Used         10
                     Observations Not Used      0
                     Total Observations        10
```

Table 1.11 (*Continued*)

Iteration History

Iteration	Evaluations	-2 Log Like	Criterion
0	1	51.13433487	
1	2	50.29813032	0.00396161
2	1	50.22541829	0.00030283
3	1	50.22033455	0.00000230
4	1	50.22029773	0.00000000

Convergence criteria met.

Covariance Parameter Estimates

Cov Parm	Estimate
BLOCK	7.4528
Residual	4.1426

Fit Statistics

-2 Log Likelihood	50.2
AIC (smaller is better)	62.2
AICC (smaller is better)	90.2
BIC (smaller is better)	59.9

Solution for Fixed Effects

Effect	TRT	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		26.5329	1.7767	8.87	14.93	<.0001
TRT	1	-14.8822	1.9330	4.75	-7.70	0.0007
TRT	2	-10.9029	2.1495	4.82	-5.07	0.0043
TRT	3	-2.4380	2.1495	4.82	-1.13	0.3099
TRT	4	0	.	.	.	.

Type 3 Coefficients for TRT

Effect	TRT	Row1	Row2	Row3
Intercept				
TRT	1	1		
TRT	2		1	
TRT	3			1
TRT	4	-1	-1	-1

Table 1.11 (*Continued*)

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
TRT	3	4.76	23.37	0.0028

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|
TRT1 - (TRT2+TRT3)/2	-8.2117	1.7085	4.3	-4.81	0.0072
TRT1 - TRT4	-14.8822	1.9330	4.75	-7.70	0.0007
TRT2 - TRT3	-8.4650	2.6087	5.86	-3.24	0.0182

Least Squares Means

Effect	TRT	Estimate	Standard Error	DF	t Value	Pr > |t|
TRT	1	11.6506	1.7767	8.87	6.56	0.0001
TRT	2	15.6299	2.0786	9.98	7.52	<.0001
TRT	3	24.0949	2.0786	9.98	11.59	<.0001
TRT	4	26.5329	1.7767	8.87	14.93	<.0001

M3: Combined analysis: ML

M4: Combined analysis: REML

In Tables 1.13 and 1.14 we present the estimates and their standard errors (exact or approximate) for these methods for purely numerical comparisons.

Based on the numerical results, we make the following observations, which should not necessarily be generalized:

1. For the least-squares means, M1 produces slightly smaller standard errors than M2, but the result is reversed for the contrast estimates.
2. The results for M2 and M4 are very similar, both with respect to estimates and standard errors.
3. In both tables M3 produces the smallest standard errors.

1.14.6 Testing of Hypotheses

To test the hypothesis

$$H_0: \tau_1 = \tau_2 = \cdots = \tau_t$$

Table 1.12 Combined Analysis Using Residual Maximum Likelihood With Post-hoc Comparisons

```
proc mixed data=IBD;

class BLOCK TRT;
model Y = TRT/ solution E3 ddfm=Satterth;
random BLOCK;
lsmeans TRT;
estimate 'TRT1 - (TRT2+TRT3)/2' TRT 1 -.5 -.5 0;
estimate 'TRT1 - TRT4' TRT 1 0 0 -1;
estimate 'TRT2 - TRT3' TRT 0 1 -1 0;
title1 'TABLE 1.12';
title2 'COMBINED ANALYSIS USING RESIDUAL MAXIMUM LIKELIHOOD';
title3 'WITH POST-HOC COMPARISONS';
run;
                         The Mixed Procedure

                          Model Information

          Data Set                      WORK.IBD
          Dependent Variable            Y
          Covariance Structure          Variance Components
          Estimation Method             REML
          Residual Variance Method      Profile
          Fixed Effects SE Method       Model-Based
          Degrees of Freedom Method     Satterthwaite


                       Class Level Information

                   Class     Levels     Values

                   BLOCK          5     1 2 3 4 5
                   TRT            4     1 2 3 4


                             Dimensions

                    Covariance Parameters      2
                    Columns in X               5
                    Columns in Z               5
                    Subjects                   1
                    Max Obs Per Subject       10
                    Observations Used         10
                    Observations Not Used      0
                    Total Observations        10
```

Table 1.12 (*Continued*)

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	37.32907360	
1	2	37.14557447	0.00022052
2	1	37.14253025	0.00000186
3	1	37.14250585	0.00000000

Convergence criteria met.

Covariance Parameter Estimates

Cov Parm	Estimate
BLOCK	6.3546
Residual	10.1681

Fit Statistics

-2 Res Log Likelihood	37.1
AIC (smaller is better)	41.1
AICC (smaller is better)	45.1
BIC (smaller is better)	40.4

Solution for Fixed Effects

Effect	TRT	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		26.5596	2.2615	5.93	11.74	<.0001
TRT	1	-14.5682	2.8843	2.68	-5.05	0.0196
TRT	2	-11.9152	3.2045	2.66	-3.72	0.0415
TRT	3	-2.0305	3.2045	2.66	-0.63	0.5766
TRT	4	0	.	.	.	.

Type 3 Coefficients for TRT

Effect	TRT	Row1	Row2	Row3
Intercept				
TRT	1	1		
TRT	2		1	
TRT	3			1
TRT	4	-1	-1	-1

Table 1.12 (*Continued*)

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
TRT	3	2.42	10.82	0.0615

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|
TRT1 - (TRT2+TRT3)/2	-7.5953	2.5979	2.2	-2.92	0.0891
TRT1 - TRT4	-14.5682	2.8843	2.68	-5.05	0.0196
TRT2 - TRT3	-9.8847	3.7522	3.82	-2.63	0.0607

Least Squares Means

Effect	TRT	Estimate	Standard Error	DF	t Value	Pr > |t|
TRT	1	11.9914	2.2615	5.93	5.30	0.0019
TRT	2	14.6444	2.7365	5.52	5.35	0.0023
TRT	3	24.5291	2.7365	5.52	8.96	0.0002
TRT	4	26.5596	2.2615	5.93	11.74	<.0001

versus

$$H_1\text{: not all } \tau_i \text{ are equal}$$

we consider a set of $t - 1$ linearly independent contrasts, say $\boldsymbol{C\tau}$, and test equivalently

$$H_0\text{: } \boldsymbol{C\tau} = \mathbf{0}$$

Table 1.13 Comparison of Least-Squares Means

TRT	M1		M2		M3		M4	
i	LSM(TRT$_i$)	SE	LSM(TRT$_i$)	SE	LSM(TRT$_i$)	SE	LSM(TRT$_i$)	SE
1	11.28	1.98	11.91	2.22	11.65	1.78	11.99	2.26
2	16.90	2.66	14.87	2.67	15.63	2.08	14.64	2.74
3	23.40	2.66	24.44	2.67	24.09	2.08	24.53	2.74
4	26.53	1.98	26.55	2.22	26.53	1.78	26.56	2.26

Table 1.14 Comparison of Contrast Estimates

	M1		M2		M3		M4	
CONTRAST[a]	$\widehat{C}$	SE	$\widetilde{C}$	SE	$\widetilde{C}$	SE	$\widetilde{C}$	SE
C1	−8.88	2.61	−7.75	2.47	−8.21	1.71	−7.59	2.60
C2	−15.25	3.02	−14.64	2.76	−14.88	1.93	−14.57	2.88
C3	−6.50	4.26	−9.57	3.62	−8.47	2.61	−9.88	3.75

[a] C1 = TRT1 − (TRT2 + TRT3)/2
C2 = TRT1 − TRT4
C3 = TRT2 − TRT3

Table 1.15 Comparison of Testing $H_0: \tau_1 = \tau_2 = \tau_3 = \tau_4$

Method	F Ratio	Denominator d.f.	P Value	Source
M1	9.41	2	.0975	Table 1.8
M2	11.73	2	.0796	See below
M3	23.37	4.76	.0028	Table 1.11
M4	10.82	2.42	.0615	Table 1.12

versus

$$H_1: \boldsymbol{C\tau} \neq \boldsymbol{0}$$

We then compute the test statistic

$$F = (\boldsymbol{C}\widetilde{\boldsymbol{\tau}})'[\boldsymbol{C}\widetilde{\boldsymbol{A}}^{-1}\boldsymbol{C}]^{-1}\boldsymbol{C}\widetilde{\boldsymbol{\tau}}/[(t-1)\ \text{MS}(E)]$$

which follows approximately an F distribution with $t - 1$ and $n - t - b + 1$ d.f.

For the data set in Table 1.7, using

$$\boldsymbol{C} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix}$$

with the Yates procedure we obtain $F = 11.73$ with 3 and 2 d.f.

A comparison of the four methods of analysis [as described in (Section 1.13.5)] concerning the test of treatment effects is given in Table 1.15

It is interesting to note that the results for M1, M2, and M4 are in close agreement, whereas the result for M3 is quite different. This appears due to the fact that the estimate of σ_e^2 for M3, namely 4.1426, is quite different from the corresponding estimates using M2, namely, 9.09375, and M4, namely 10.1681.