Relations among partitions

R. A. Bailey

Abstract

Combinatorialists often consider a balanced incomplete-block design to consist of a set of points, a set of blocks, and an incidence relation between them which satisfies certain conditions. To a statistician, such a design is a set of experimental units with two partitions, one into blocks and the other into treatments; it is the relation between these two partitions which gives the design its properties. The most common binary relations between partitions that occur in statistics are refinement, orthogonality and balance. When there are more than two partitions, the binary relations may not suffice to give all the properties of the system. I shall survey work in this area, including designs such as double Youden rectangles.

1 Introduction

Many combinatorialists think of a balanced incomplete-block design (BIBD) as a set $\mathcal{P}$ of points together with a collection $\mathcal{B}$ of subsets of $\mathcal{P}$, called *blocks*, which satisfy various conditions. For example, see [52]. Some papers, such as [16, 65, 201], call a BIBD simply a *design*. Others think of it as the pair of sets $\mathcal{P}$ and $\mathcal{B}$ with a binary incidence relation between their elements. These views are both rather different from that of a statistician who is involved in designing experiments. The following examples introduce the statistical point of view, as well as serving as a basis for the combinatorial ideas in this paper.

Example 1.1 A horticultural enthusiast wants to compare three varieties of lettuce for people to grow in their own gardens. He enlists twelve people in his neighbourhood. Each of these prepares three patches in their vegetable garden, and grows one of the lettuce varieties on each patch, so that each gardener grows all three varieties.

Here the patches of land are experimental units. There may be some differences between the gardeners, so the three patches in a single garden form what is called a *block*. Each variety occurs just once in each block, and so the blocks are said to be *complete*. Complete-block designs were advocated by Fisher in [78], and are frequently used in practice.

Example 1.2 Now suppose that the number of lettuce varieties is increased to nine. It is not reasonable to expect an amateur gardener to

A	D	G	A	B	C	A	B	C	A	B	C
B	E	H	D	E	F	E	F	D	F	D	E
C	F	I	G	H	I	I	G	H	H	I	G

Figure 1: Balanced incomplete-block design in Example 1.2: columns represent blocks and letters represent varieties

grow nine different varieties, so each gardener still uses only three patches of ground, and thus can grow only three varieties. The blocks are now *incomplete*, in the terminology of Yates [227].

One possible layout is shown in Figure 1. This incomplete-block design has the property that each pair of distinct varieties concur in the same number of blocks (here, exactly one). Yates originally called incomplete-block designs with this property *symmetrical*, but the adjective had been changed to *balanced* within a few years [46, 80].

To a statistician, the partition of the set of experimental units into blocks is inherent and is known before the decision is taken about which variety to allocate to each unit. This allocation gives another partition of the set of experimental units, and it is the relation between these two partitions that is regarded as *balance*. It is not a symmetric relation, in general. In Example 1.2 the varieties are balanced with respect to the blocks, but the blocks are not balanced with respect to the varieties because some pairs of blocks have one variety in common while others have none. This relation is discussed in more detail in Section 5.

In fact, statisticians usually call these partitions *factors*, because the names of the parts are relevant. In Example 1.2 the names of the varieties are not interchangeable; we probably want to find out which one does best. Thus a factor is typically regarded as a function from the set of experimental units to a finite set: if B and L denote the factors for blocks and lettuce varieties respectively and ω is a vegetable patch then $B(\omega)$ is the block (garden) containing ω and $L(\omega)$ is the variety grown on ω. Furthermore, $|B(\omega)|$ is the size of the block containing ω, while $|L(\omega)|$ is the number of patches with the same variety as that grown on ω.

A response Y_ω, such as total yield of edible lettuce in kilograms, is measured on each patch ω. It is usually assumed that Y_ω is a random variable and that there are constants τ_i and β_j such that

$$Y_\omega = \tau_{L(\omega)} + \beta_{B(\omega)} + \varepsilon_\omega, \tag{1.1}$$

where the final terms ε_ω are independent random variables with zero mean and the same variance σ^2; often they are assumed to be normally dis-

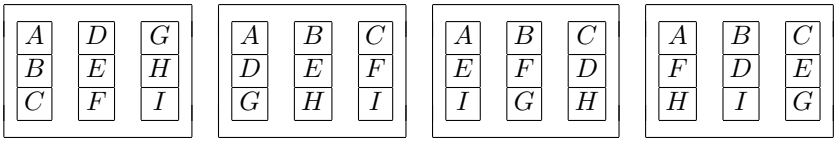

Figure 2: Resolved balanced incomplete-block design in Example 1.3: columns represent blocks, rectangles represent districts and letters represent varieties

tributed. The purpose of the experiment is to estimate the constants τ_i. Of course, this is impossible, because equation (1.1) is unchanged if a constant is added to every τ_i and subtracted from every β_j, but we aim to estimate differences such as $\tau_1 - \tau_2$, that is, to estimate the τ_i up to an additive constant.

Thus the two partitions have different roles. One (the partition B) is inherent, and we are usually not interested in the effects β_j of the different parts. The other (the partition L) has its parts allocated by the experimenter, and the purpose of the experiment is to find out what differences there are between its parts. Nonetheless, this paper will concentrate on the combinatorial relation between them. Before doing so, we give some examples with three partitions.

Example 1.3 Suppose that the twelve gardeners in Example 1.2 do not all live in the same neighbourhood. Instead, they are spread over four different districts, with three per district. If the first three blocks in Figure 1 represent the gardens in the first district, and so on, then each variety is grown once in each district, as shown in Figure 2. This is convenient if other people want to look at the different varieties during the course of the experiment.

Each block is contained within a single district, so the partition into blocks is a *refinement* of the partition into districts. Section 3 discusses refinement in more detail. On the other hand, the partitions into districts and into varieties have the property that each part of one (a district) meets each part of the other (a variety) in a single experimental unit. This is a special case of *strict orthogonality*, which is explained in Section 4.

The assumption about Y_ω might remain as in (1.1) or it might be

$$Y_\omega = \tau_{L(\omega)} + \beta_{B(\omega)} + \gamma_{D(\omega)} + \varepsilon_\omega, \tag{1.2}$$

where $D(\omega)$ is the district containing ω. Of course, if the β_j and γ_k are all constants then they are not estimable, because we can add a constant

A	B	C	D	E	F	G
C	D	E	F	G	A	B
D	E	F	G	A	B	C
E	F	G	A	B	C	D

Figure 3: Row–column design in Example 1.4: rows represent months, columns represent people and letters represent exercise regimes

to γ_1 and subtract it from β_j for all blocks j in district 1. However, it is sometimes assumed that the β_j are independent random variables with zero mean and the same variance σ_B^2. Section 6.2 discusses further the potential difficulty in an assumption like (1.2) when one partition is a refinement of another.

An incomplete-block design whose blocks can be grouped into collections each of which contains each variety just once, as in Example 1.3, is called *resolvable*. Section 15 gives more information about such designs.

Example 1.4 In order to assess the benefits of different exercise regimes, a health scientist asks seven healthy people to participate in an experiment over four months. Each month each person will be allocated one of seven exercise regimes. At the end of each month, the change in some measure of fitness, such as heart rate, will be recorded for each person.

Now each experimental unit is one person for one month. The partitions into months and into people are inherent, but the scientist chooses the partition into exercise regimes. Figure 3 shows one possible design for this experiment. The partitions into months and into people are strictly orthogonal to each other, as are the partitions into months and into exercise regimes. The partitions into people and into exercise regimes are both balanced with respect to each other.

Example 1.5 A small modification of Example 1.4 has five months, six people and ten exercise regimes. One possible design is shown in Figure 4, where rows represent months, columns represent people and letters represent exercise regimes.

Example 1.6 A modification of Example 1.2 has ten gardens of three vegetable patches each, and six varieties of lettuce. In addition, there are are five possible watering regimes. Each patch must have one variety of lettuce and one watering regime. The design in Figure 4 can be used, but now rows represent watering regimes, columns represent lettuce varieties and letters represent gardens.

H	J	I	G	F	E
J	I	H	C	B	D
D	F	A	J	G	C
A	B	G	E	D	I
E	A	C	B	H	F

Figure 4: Combinatorial design used in Examples 1.5 and 1.6

Denote by R, C and L the partitions into rows, columns and letters in the design in Figure 4. From the point of view of the statistician, the uses of this design in Examples 1.5 and 1.6 are quite different. In the former, the partitions R and C are inherent while L is at the choice of the experimenter; in the latter, L is inherent while the experimenter chooses R and C. However, in both cases it may be assumed that

$$Y_\omega = \alpha_{R(\omega)} + \phi_{C(\omega)} + \tau_{L(\omega)} + \varepsilon_\omega. \tag{1.3}$$

From a combinatorial point of view, Figure 4 simply shows a set with three partitions. The partitions R and C are strictly orthogonal to each other, while each of R and C is balanced with respect to letters. In fact, there is a third property, called *adjusted orthogonality*, that will be defined in Section 8.

For further explanation of how combinatorial design problems arise from statistically designed experiments, see [22, 33, 177, 204].

The remainder of this paper treats a combinatorial design as a collection of partitions of a finite set. Section 2 establishes some notation for partitions and their associated matrices and subspaces. Sections 3–5 discuss the three most important binary relations between partitions, all of which have been seen in the examples so far. Section 6 explains more about the background to equations (1.1)–(1.3). Section 7 discusses the relations between the subspaces defined by partitions, and shows that sometimes there is a need for a ternary relation. Sections 8 and 9 give more details of two important non-binary relations. These are used in Section 10, which considers possibilities for three partitions. This leads to several different types of combinatorial design, considered in the remaining sections. Each type is defined by three partitions, or is a simple generalization with more partitions but no need for any further non-binary relations.

2 Partitions on a finite set

Let Ω be a finite set of size e, where $e > 1$. The elements of Ω will be called *experimental units*, or just *units*. The rest of this paper deals with

partitions of Ω.

If F is such a partition, denote by n_F its number of parts. The $e \times n_F$ *incidence matrix* X_F has (ω, i)-entry equal to 1 if unit ω is in part i of F; otherwise, this entry is zero. Thus $X_F X_F^\top$ is the $e \times e$ *relation matrix* for F, with (ω_1, ω_2)-entry equal to 1 if ω_1 and ω_2 are in the same part of F, and equal to 0 otherwise. The $n_F \times n_F$ matrix $X_F^\top X_F$ is diagonal, with (i, i)-entry equal to the size of the i-th part of F.

Definition A partition is *uniform* if all of its parts have the same size.

Many statisticians, including Tjur [208, 209], call uniform partitions *balanced*, but this conflicts with the notion of balance introduced in Section 1. This terminology is discussed again in Section 9. Preece reviewed the overuse of the word *balance* in design of experiments in [155]. The adjectives *homogeneous* [44], *proper* [151] and *regular* [66] are also used.

If F is uniform, denote the size of all its parts by k_F. Then $n_F k_F = e$ and $X_F^\top X_F = k_F I_{n_F}$, where I_n is the identity matrix of order n.

Denote by $\mathbf{R}^\Omega$ the real vector space of dimension e whose coordinates are labelled by the elements of Ω, so that each vector may be regarded as a function from Ω to $\mathbf{R}$. If F is a partition of Ω, denote by V_F the subspace of $\mathbf{R}^\Omega$ consisting of vectors which are constant on each part of F. Then $\dim(V_F) = n_F$.

We assume the standard inner product on $\mathbf{R}^\Omega$. Denote by P_F the matrix of orthogonal projection onto V_F. Then P_F replaces the coordinate y_ω of any vector y by the average value of y_ν for ν in $F(\omega)$, which is the part of F containing ω. In fact, $P_F = X_F \left(X_F^\top X_F\right)^{-1} X_F^\top$. If F is uniform then $X_F X_F^\top = k_F P_F$.

Equations (1.1)–(1.3) all have the form

$$Y = \sum_{F \in \mathcal{F}} X_F \psi_F + \varepsilon, \tag{2.1}$$

where Y and ε are random vectors of length e, $\mathcal{F}$ is a set of partitions of Ω, and, for F in $\mathcal{F}$, ψ_F is a real vector of length n_F. Thus the expectation $\mathbb{E}(Y)$ of Y is in the subspace $\sum_{F \in \mathcal{F}} V_F$.

There are two trivial partitions on Ω, which are different when $e > 1$. The parts of the *equality* partition E are singletons, so $k_E = 1$, $n_E = e$ and $X_E = I_e = P_E$. At the other extreme, the *universal* partition U has a single part, so $n_U = 1$, $k_U = e$, $X_U X_U^\top = J_{ee}$ and $P_U = e^{-1} J_{ee}$, where J_{nm} denotes the $n \times m$ matrix with all entries equal to 1. Moreover, V_E is the whole space $\mathbf{R}^\Omega$, while V_U is the 1-dimensional subspace of constant vectors.

If F and G are two partitions of Ω, their $n_F \times n_G$ *incidence matrix* N_{FG} is defined by $N_{FG} = X_F^\top X_G$. The (i,j)-entry is the size of the intersection of the i-th part of F with the j-th part of G. In particular, $N_{EF} = X_F$.

Given a set $\mathcal{F}$ of partitions of Ω, denote by $\mathcal{A}_\mathcal{F}$ the algebra of $e \times e$ real matrices generated by the projection matrices P_F for F in $\mathcal{F}$, and denote by $\mathcal{J}_\mathcal{F}$ the algebra generated by the relation matrices $X_F X_F^\top$ for F in $\mathcal{F}$. These are the same if all partitions in $\mathcal{F}$ are uniform. James called $\mathcal{J}_\mathcal{F}$ the *relationship algebra* of $\mathcal{F}$ in [99], but it was shown in [100, 115] that $\mathcal{A}_\mathcal{F}$ is more useful for understanding the properties of $\mathcal{F}$ relevant to a designed experiment.

3 Refinement

Definition If F and G are partitions of Ω, then F is *finer* than G (equivalently, G is *coarser* than F) if every part of F is contained in a single part of G but at least one part of G is not a part of F. This relation is denoted $F \prec G$ or $G \succ F$.

In Example 1.3, $B \prec D$. If $F \prec G$ then $n_F > n_G$ and $V_G < V_F$.

Write $F \preccurlyeq G$ (or $G \succcurlyeq F$) to mean that either $F \prec G$ or $F = G$. Then $\preccurlyeq$ is a partial order. For every partition F, it is true that $E \preccurlyeq F \preccurlyeq U$ and $V_U \le V_F \le V_E$.

Proposition 3.1 *Let F and G be partitions of Ω. If $F \preccurlyeq G$ then $P_F P_G = P_G P_F = P_G$.*

As with any partial order, there is a choice about which of the two objects should be considered 'smaller'. Some statisticians write the refinement partial order in the opposite way to that used here. For example, see [31, 208, 209].

Since there are only a finite number of partitions of Ω, there is no difficulty with the next definition.

Definition Let F and G be partitions of Ω. The *infimum* $F \wedge G$ of F and G is the coarsest partition H satisfying $H \preccurlyeq F$ and $H \preccurlyeq G$; its parts are the non-empty intersections of a part of F and a part of G. Thus $F \wedge G = E$ if and only if no part of F intersects any part of G in more than one unit. The *supremum* $F \vee G$ of F and G is the finest partition K satisfying $F \preccurlyeq K$ and $G \preccurlyeq K$; its parts are the connected components of the graph with vertex-set Ω and an edge between ω_1 and ω_2 if $F(\omega_1) = F(\omega_2)$ or $G(\omega_1) = G(\omega_2)$.

Thus if $F \preccurlyeq G$ then $F \wedge G = F$ and $F \vee G = G$. In the design in Figure 4, $R \wedge C = R \wedge L = C \wedge L = E$ and $R \vee C = R \vee L = C \vee L = U$.

Proposition 3.2 *If F and G are partitions of Ω then $V_F \cap V_G = V_{F \vee G}$.*

4 Orthogonality

4.1 Definitions

As Preece noted in [154], the word *orthogonal* has many different meanings in the statistical literature. Here I use the terminology in [23, 25, 32, 208].

Proposition 3.2 shows that subspaces V_F and V_G can never be orthogonal to each other. This motivates the following definition, from [208].

Definition Let V and W be subspaces of $\mathbf{R}^\Omega$. Then V and W are *geometrically orthogonal* to each other if the subspaces $V \cap (V \cap W)^\perp$ and $W \cap (V \cap W)^\perp$ are orthogonal to each other.

Proposition 4.1 *Let F and G be partitions of Ω. The following statements are equivalent:*

(i) V_F *is geometrically orthogonal to* V_G*;*

(ii) $P_F P_G = P_G P_F$*;*

(iii) $P_F P_G = P_{F \vee G}$*;*

(iv) for every unit ω*, we have* $|F(\omega)|\,|G(\omega)| = |(F \wedge G)(\omega)|\,|(F \vee G)(\omega)|$.

The second statement above is sometimes called 'projectors commute', and the fourth 'proportional meeting within each class of the supremum'.

Definition Let F and G be partitions of Ω. Then F is *orthogonal* to G, written $F \perp G$, if $P_F P_G = P_G P_F$; and F is *strictly orthogonal* to G, written $F \underline{\perp} G$, if $P_F P_G = P_G P_F = P_U$.

Duquenne calls these two concepts *local orthogonality* and *orthogonality* respectively in [66]; the latter agrees with Gilliland's definition of orthogonality in [83]. Some authors split the definitions further according to whether or not $F \wedge G$ is uniform.

Proposition 3.1 shows that if $F \preccurlyeq G$ then $F \perp G$. In particular, all partitions are orthogonal to both E and U, and every partition is orthogonal to itself.

A B	B C	A C
C	A	B

Figure 5: A 2×3 row–column design with nine units and three letters, giving mutually orthogonal partitions into rows, columns and letters

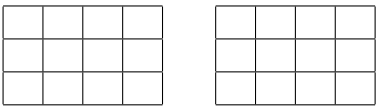

Figure 6: Two blocks, each of which is a 3×4 rectangle, so that there are 6 rows and 8 columns

In the design in Figure 4, $R \underline{\perp} C$. If R, C and L denote the partitions into rows, columns and letters in Figure 5, then $R \underline{\perp} C$, $R \underline{\perp} L$ and $C \underline{\perp} L$ even though R, $R \wedge C$ and $R \wedge L$ are not uniform, because the 'proportional meeting' condition in Proposition 4.1(iv) is satisfied for all pairs and all pairwise suprema are equal to U. If B, R and C denote the partitions into blocks, rows and columns in Figure 6, then $R \perp C$ but R is not strictly orthogonal to C because $R \vee C = B \neq U$.

Proposition 4.2 *Let F and G be partitions of Ω. Then $F \underline{\perp} G$ if and only if $N_{FG} = e^{-1}(X_F^\top X_F) J_{n_F n_G} (X_G^\top X_G)$.*

4.2 Orthogonal arrays

Definition An *orthogonal array* of strength two on Ω is a collection $\mathcal{F}$ of at least two uniform partitions of Ω with the property that every pair of distinct partitions is strictly orthogonal. Inductively, for $m \geq 3$, a collection $\mathcal{F}$ of at least m partitions of Ω is an orthogonal array of strength m if it is an orthogonal array of strength $m-1$ and, whenever $F_1, \ldots, F_m$ are distinct partitions in $\mathcal{F}$, the infimum $F_1 \wedge F_2 \wedge \cdots \wedge F_{m-1}$ is strictly orthogonal to F_m.

Figure 7 shows an orthogonal array of strength two with $e = 12$, $|\mathcal{F}| = 11$, and $n_F = 2$ for all F in $\mathcal{F}$. It is equivalent to that given by Plackett and Burman [142]. Replacing each 0 by -1 and adjoining a row of 1s gives a Hadamard matrix of order 12. The paper [142] inspired Rao to define orthogonal arrays and begin to develop a general theory of them in [181, 182].

F_1	0	0	1	0	0	0	1	1	1	0	1	1
F_2	1	0	0	1	0	0	0	1	1	1	0	1
F_3	0	1	0	0	1	0	0	0	1	1	1	1
F_4	1	0	1	0	0	1	0	0	0	1	1	1
F_5	1	1	0	1	0	0	1	0	0	0	1	1
F_6	1	1	1	0	1	0	0	1	0	0	0	1
F_7	0	1	1	1	0	1	0	0	1	0	0	1
F_8	0	0	1	1	1	0	1	0	0	1	0	1
F_9	0	0	0	1	1	1	0	1	0	0	1	1
F_{10}	1	0	0	0	1	1	1	0	1	0	0	1
F_{11}	0	1	0	0	0	1	1	1	0	1	0	1

Figure 7: Orthogonal array of strength two, consisting of 11 partitions of a set of size 12 into two parts: columns represent elements of the set, and each row shows one partition

For $n \geq 2$, the rows, columns and letters of any Latin square of order n give an orthogonal array of strength two on a set of size n^2, with three partitions into parts of size n. See [95] for many uses and constructions of orthogonal arrays, as well as more theory. Eendebak and Schoen maintain a catalogue on the web page [76].

From Finney [77] onwards, finite Abelian groups have been a fruitful source of orthogonal arrays, under the name *fractional factorial designs*. For $i = 1, \ldots, s$ let G_i be an Abelian group of order n_i, where $n_i \geq 2$. Let G be the product group $G_1 \times G_2 \times \cdots \times G_s$. Every complex irreducible character χ of G has the form $\chi = (\chi_1, \chi_2, \ldots, \chi_s)$ where χ_i is an irreducible character of G_i and $\chi(g_1, g_2, \ldots, g_s) = \chi_1(g_1)\chi_2(g_2)\cdots\chi_s(g_s)$. Let H be a subgroup of G, and let F_i be the partition of H defined by the values of the i-th coordinate. Then $\{F_1, \ldots, F_s\}$ forms an orthogonal array of strength m on H if and only if the only non-trivial characters χ of G whose restriction to H is trivial have non-trivial components χ_i for at least $m + 1$ values of i. For example, if $s = 3$, $n_1 = n_2 = n_3 = 7$ and G_i is $\mathbf{Z}_7$ written additively for $i = 1$, 2 and 3 then $\{F_1, F_2, F_3\}$ forms an orthogonal array of strength two on the subgroup $H = \{(g_1, g_2, g_3) : g_1 + g_2 + g_3 = 0\}$. Up to isotopism (permutations of the names of the parts of each partition), this is the Latin square obtained as the Cayley table of $\mathbf{Z}_7$.

Some papers, such as [61, 112, 141, 215], call an orthogonal array *regular* if and only if it is made from an Abelian group in this way. There are two problems with this. The first is that, in each experiment, the parts of F_i (such as varieties of lettuce) are unlikely to be labelled by the elements of a finite Abelian group. How is the statistician analysing the