# Block designs

## 1 Background

In a typical experiment, we have a set $\Omega$ of experimental units or plots, and (after some preparation) we make a measurement on each plot (for example, the yield of the plot). So the result of the experiment is a function from $\Omega$ to the real numbers, that is, an element of the vector space $\mathbb{R}^\Omega$. The dimension of this vector space is equal to $|\Omega|$, and a basis for the space consists of the characteristic functions of the elements of $\Omega$: these are the functions $f_\alpha$ given by

$$f_\alpha(\omega) = \begin{cases} 1 & \text{if } \omega = \alpha, \\ 0 & \text{otherwise.} \end{cases}$$

The statistician has two jobs:

- *design*, that is, specification of the preparations made before the measurement (for example, allocation of treatments to plots);
- *analysis*, that is, drawing conclusions about the treatments from the result of the experiment.

If all the plots were identical and independent, design of the experiment would be easy. We could allocate treatments to plots in any manner, though we might expect that some allocations are better than others.

Of course, things are not so simple, and the many forms of experimental design have been devised to cope with the fact that plots are not identical and independent. For example, in an agricultural experiment, we may have access to a number of plots on experimental farms in different parts of the country. We expect that fertility, drainage, etc., will vary from one plot to another, but that plots on the same farm will be more alike than plots on different farms.

A *block design* is the type of experimental design used to cope with this situation. That is, there is a partition $B$ of the set $\Omega$ into subsets called *blocks*, over which we have no control. We are free to decide which treatment to apply to each plot, that is, choose a second partition of $\Omega$ into subsets called *treatments*. Of all possible partitions, we want to choose one that will allow the maximum amount of information about treatment effects to be extracted from the result of the experiment.

Thus, a *block design* consists of a set $\Omega$ and two partitions of $\Omega$, the *block partition* $B$ and the *treatment partition* $T$. We let $n = |\Omega|$, $b = |B|$, and $v = |T|$.

(Some authors use $t$ for $|T|$, but $t$ has a different meaning in combinatorial design theory; think of $v$ for "varieties".)

A general reference for block designs is Dey [2].

## 2 Matrices

The partition $B$ of $\Omega$ into blocks can be described by a linear map $D_B$ from the vector space $\mathbb{R}^\Omega$ to the space $\mathbb{R}^B$, which maps a function $f \in \mathbb{R}^\Omega$ to $g \in \mathbb{R}^B$ given by

$$g(x) = \sum_{\omega \in x} f(\omega).$$

If we use the standard bases for the two vector spaces, the matrix representing $D_B$ has rows indexed by $\Omega$ and columns by $B$, and its $(\omega, x)$ entry is equal to 1 if $\omega \in x$, and 0 otherwise. It is an easy exercise to check that

$$D_B D_B^\top = K,$$

where $K$ is the diagonal matrix with $(x, x)$ entry equal to the number of plots in the block $x$.

Of course, the treatment partition $T$ is also represented by a linear transformation from $\mathbb{R}^\Omega$ to $\mathbb{R}^T$, and hence by a matrix $D_T$ satisfying

$$D_T D_T^\top = R,$$

where $R$ is the diagonal matrix with $(a, a)$ entry equal to the number of plots in the treatment $a$.

Let

$$N = D_T D_B^\top.$$

Then $N$ is the *incidence matrix* of the design. Its $(a, x)$ entry is equal to the number of plots in block $x$ which receive the treatment $a$.

## 3 Binary designs

The block design is called *binary* if each treatment occurs at most once in a block. In this case, $N$ is a zero-one matrix, and we can regard it as the incidence matrix of an incidence structure whose points and blocks are the treatments and blocks of the block design, a point and block being incident if the treatment occurs on some plot in the block. In this case, $N$ is the usual incidence matrix of the incidence structure. Thus, block designs in statistics are more general than those in combinatorics in this respect: they are not necessarily binary!

A binary design (in the sense of a set with two partitions) can be reconstructed from the corresponding incidence structure. The plots of the design are the *flags*, or incident point-block pairs, of the incidence structure. To each block we associate the set of flags containing that block; this gives the partition of the plots into blocks. Analogously, there is a partition of the plots corresponding to the points (which we re-name as "treatments").

The *incidence graph*, or *Levi graph*, of an incidence structure is the bipartite graph whose vertices are the points and blocks of the structure, two vertices adjacent if one is a point and the other a block incident with it. Thus, a binary block design is recovered from the Levi graph of its incidence structure by taking the plots to be the edges of the graph, and the treatments and blocks to correspond to the two bipartite sets of vertices (in some order).

# 4 Connectedness

To do any analysis on the experiment, we have to make some assumptions about the effects of treatments and blocks. We consider here only one simple case, a *fixed-effects linear model*. In this case, we assume that the result of the experiment (which is a vector in $\mathbb{R}^\Omega$) is given by

$$f = \mu + D_T^\top \tau + D_B^\top \beta + e,$$

where $\mu$ is a constant vector and $\tau \in \mathbb{R}^T$, $\beta \in \mathbb{R}^B$ are unknown vectors of *treatment effects* and *block effects* respectively. The components of the vector $e$ (the random errors) are assumed to be uncorrelated random variables all having mean 0 and the same variance $\sigma^2$. We wish to estimate the treatment effect vector $\tau$.

This estimation is done by the least-squares method. This gives rise to the following system of linear equations for $\tau$:

$$C\tau = Qf,$$

where

$$\begin{aligned} C &= R - NK^{-1}N^\top, \\ Q &= D_T - NK^{-1}D_B. \end{aligned}$$

The matrix $C = R - NK^{-1}N^\top$ is called the *information matrix* or *C-matrix* of the block design.

It is easily shown that $C$ is positive semidefinite. Moreover, the all-1 vector $j$ satisfies $Cj = 0$. (This corresponds to the fact that, from what we have said, $\tau$ can only be determined up to an additive constant.) Hence a necessary condition for a linear combination $p^\top \tau$ of the coefficients of $\tau$ to be *estimable* (that is, to be in the image of $C$) is that $p^\top j = 0$.

The block design is said to be *connected* if the rank of $C$ is $v-1$ (that is, the null space of $C$ is spanned by $j$). If this holds, then the image of $C$ consists of all vectors in $\mathbb{R}^T$ with coordinate sum 0. Such vectors are called *contrasts*. The space of contrasts is spanned by the *elementary contrasts* $\tau_i - \tau_j$.

The use of the word "connected" is explained by the next result, which is due to Chakrabarti [1].

**Theorem 1** *A block design is connected if and only if, for any two treatments $i$ and $j$, there is a sequence with first term $i$ and last term $j$, in which any two consecutive terms are a treatment and a block which contain a common plot.*

Hence, a binary design is connected if and only if its Levi graph is connected.

# 5 Balance

An unbiased estimator of a treatment contrast is a random variable whose expected value is the value we are looking for. A design is said to be *variance-balanced* if the variance of the best unbiased estimators of all normalised treatment contrasts are equal. Rao [5] showed that a connected block design is variance-balanced if and only if all the non-zero eigenvalues of $C$ are equal; equivalently:

**Theorem 2** *A connected block design is variance-balanced if and only if its information matrix has the form*

$$C = (a-b)I + bJ$$

*for some scalars $a$ and $b$, where $J$ is the all-1 matrix.*

This gives a combinatorial interpretation of variance balance. In particular, the constancy of the off-diagonal entries of $C$ means that, if each block is assigned a weight which is the reciprocal of its cardinality, then the sum of the weights of the blocks containing two distinct treatments (counted according to the multiplicities of the treatments in the blocks) is constant.

There is a related concept of *efficiency balance*, which we treat more briefly. The efficiency of a design is computed relative to a (possibly hypothetical) design on the same set in which treatment contrasts and block contrasts are orthogonal (and so can be estimated independently). Such an *orthogonal design* would satisfy $N = rk^\top/n$, where $r$ and $k$ are the vectors of treatment sizes and block sizes (the diagonals of $R$ and $K$ respectively).

We turn straight to the criterion for efficiency balance. The matrix $R$ is diagonal with positive entries (the sizes of the parts of the treatment partition), and so has a square root $R^{1/2}$ which is also diagonal with positive entries. Now the *M-matrix* of the design is defined to be the real symmetric matrix

$$M = R^{-1/2} N K^{-1} N^\top R^{-1/2}.$$

Since $C = R^{1/2}(I - M)R^{1/2}$, we see that all the eigenvalues of $M$ are at most 1. Moreover, it has the eigenvalue 1, and the multiplicity of this eigenvalue is equal to 1 if and only if the design is connected.

A connected block design is said to be *efficiency-balanced* if all the eigenvalues of its M-matrix apart from the trivial eigenvalue 1 are equal.

This condition also has a combinatorial interpretation, rather more complicated than that for variance balance. This condition can be extracted from the fact that the M-matrix has the form $(a-b)I + bR^{1/2}JR^{1/2}$ for some scalars $a$ and $b$.

# 6 Simplifying conditions

A block design is said to be *equireplicate* if all treatments contain the same number of plots; in other words, if the matrix $R$ is a scalar: $R = rI$.

Dey *et al.* [3] proved the following result.

**Theorem 3** *Any two of the following properties of a connected block design implies the third:*

*(a) variance-balanced;*

*(b) efficiency-balanced;*

*(c) equireplicate.*

To see that (a) and (b) are equivalent in the presence of (c) we note that, if $R = rI$, then $C$ and $M$ have the same eigenspaces, and their eigenvalues are related simply by $\lambda_C = r(1-\lambda_M)$.

A block design is called *uniform* if all parts of the block partition contain equally many plots. A binary block design is *pairwise balanced* or *combinatorially balanced* if any two treatments occur together in the same number of blocks. We note that a binary, pairwise balanced uniform design is necessarily equireplicate and connected, and is both variance-balanced and efficiency-balanced: we call such a design *balanced*.

We saw earlier that binary block designs can be regarded by incidence structures. We recall that a $t$-$(v,k,\lambda)$ design, or $t$-design, is an incidence structure with $v$ points, in which every block is incident with $k$ points and any $t$ points are incident with $\lambda$ blocks. Now

(a) a binary uniform design is a 0-design;

(b) a binary uniform equireplicate design is a 1-design;

(c) a binary uniform balanced design is a 2-design.

There is one further distinction made for binary block designs. Such a design is a *complete-block design* if every treatment occurs in every block, and is an *incomplete-block design* otherwise. In terms of incidence structures, a design is complete if $k = v$ and incomplete if $k < v$. In the terminology of the last section, a binary block design is orthogonal if and only if it is a complete-block design.

Thus, a *balanced incomplete-block design*, or BIBD in the statistical literature, is a 2-$(v,k,\lambda)$ design with $k < v$.

# 7 Existence

We now consider existence of BIBDs, and summarise some of the main results.

**Theorem 4** *(a) The parameters of a* 2-$(v,k,\lambda)$ *design with b blocks and replication number r satisfy* $vr = bk$ *and* $(v-1)\lambda = r(k-1)$.

*(b) If a* 2-$(v,k,\lambda)$ *design exists, then* $\lambda(v-1) \equiv 0 \pmod{k-1}$ *and* $\lambda v(v-1) \equiv 0 \pmod{k(k-1)}$.

*(c) Given k and* $\lambda$, *the necessary conditions of (b) are sufficient for all but finitely many values of v.*

*(d) If $k < v$ (and $\lambda > 0$) , then $b \geq v$.*

Part (a) is easy double-counting, and (b) follows immediately from (a). Part (c) was proved by Wilson, and is a consequence of his general theorem on PBD-closed sets [6]. Part (d) is Fisher's inequality [4], and indeed holds more generally for all non-orthogonal efficiency-balanced designs.

# References


[1] M. C. Chakrabarti, On the C-matrix in design of experiments, *J. Indian Statist. Assoc.* **1** (1963), 8–23.

[2] A. Dey, *Theory of Block Designs*, Wiley Eastern, New Delhi, 1986.

[3] A. Dey, M. Singh and G. M. Saha, Efficiency balanced block designs, *Commun. Statist.* (A) **10** (1981), 237–247.

[4] R. A. Fisher, An examination of the different possible solutions of a problem in incomplete blocks, *Ann. Eugenics* **10** (1940), 52–75.

[5] V. R. Rao, A note on balanced designs, *Ann. Math. Statist.* **29** (1958), 290–294.

[6] R. M. Wilson, An existence theory for pairwise balanced designs: I, Composition theorems and morphisms, *J. Combinatorial Theory* (A) **13** (1972), 220–245; II, The structure of PBD-closed sets and the existence conjectures, *ibid.* **13** (1972), 246–273; III, A proof of the existence conjectures, *ibid.* **18** (1975), 71–79.



Peter J. Cameron

August 7, 2002