Interchanges and Statistical Designs

Wen-Ai Jackson

Department of Pure Mathematics, The University of Adelaide,
Adelaide SA 5005, Australia

Abstract. Let D be a statistical design consisting of v varieties on b blocks. We say it is *binary* if a variety occurs at most once on every block. For varieties p and q on blocks x and y respectively, we can obtain a design D' from D by *interchanging* p and q (so p is on y and q is on x). We say that the exchange is *binary* if either $x = y$, or $p = q$, or: q not on x and p not on y. If D is binary then D' is also binary. We examine the equivalence classes for binary designs under binary interchanges. This work is relevant in the construction of certain classes of optimal or near-optimal designs.

1 Introduction

In [5] Venables and Eccleston considered a family of techniques to construct "optimal" or "near optimal" incomplete block designs. These designs may have extra prescribed properties such as resolvability, or being a row-column design. The techniques employ randomized search directions and at some stages allow the possibilities of taking "steps" in a direction of decreasing efficiency in an effort to avoid local optima (as in the simulated annealing algorithm of combinatorial optimization). These steps are in general called *interchanges* and the aim of this note is to examine a special class of interchanges (which we will call *binary* or b-interchanges) and determine which designs can be obtained from others using b-interchanges. The two advantages of using b-interchanges over interchanges (in an appropriate situation) is firstly, that there are in general fewer b-interchanges than interchanges at any one step, and secondly, that using b-interchanges preserves the property of the design being binary; in general this is not true for interchanges. For a given parameter set, the subclass of the designs which are binary is relatively small. A possible disadvantage of using b-interchanges over interchanges is that it may take more b-interchanges to transform one design into another.

2 Notation

For an introduction to the theory of statistical designs, see [4]. Let $\mathcal{V}$ be a set of v *varieties* and $\mathcal{B}$ a set of b *blocks*. A *design* is an allocation of varieties to blocks, where a variety may be allocated more than once to a block. Let n_{ij} denote the number of times variety i occurs in block j ($1 \leq i \leq v$, $1 \leq j \leq b$). The $v \times b$ matrix $N = [n_{ij}]$ is called the *incidence matrix* of the design. If for all i, j, $n_{ij} \in \{0, 1\}$ then the design is said to be *binary*. This is the case when each variety occurs at most once in every block, and a block can be considered to be a set of varieties. A design is *proper* if no variety is on every block and no block contains every variety.

Let r_i ($1 \leq i \leq v$) be the number of times variety i occurs in the design; similarly let k_j ($1 \leq j \leq b$) be the number of (not necessarily distinct) varieties in block j. If $\mathbf{r} = (r_1, \ldots, r_v)^T$ (where ^{T} denotes the transpose) and $\mathbf{k} = (k_1, \ldots, k_b)^T$, then

$$N\mathbf{1}_b = \mathbf{r} \qquad \text{and} \qquad N^T\mathbf{1}_v = \mathbf{k}$$

where $\mathbf{1}_u$ is the all ones column vector of size u.

If $\mathbf{r} = r\mathbf{1}_b$ then the design is said to be *equireplicate* with *replication number* r. If $\mathbf{k} = k\mathbf{1}_v$ then the design is said to be a block design with *block size* k. If variety p occurs in block x we write $p \in x$, otherwise we write $p \notin x$.

For a design D, let p, q be varieties on blocks x and y respectively. We call $E = [(p, x), (q, y)]$ an *interchange pair* (see [2]). Let $D(E)$ be the design which is the same as D except that in $D(E)$ variety p occurs in block y in place of that occurrence of q, and variety q occurs in block x in place of that occurrence of p. If $x = y$ or $p = q$ then $D(E) = D$ and the interchange is called *trivial*. Note that interchanging does not alter the number k_i of varieties on block i or the number r_j of blocks on variety j. Hence if N and N' are the incidence matrices for designs D and $D(E)$ respectively, we have

$$N\mathbf{1}_b = N'\mathbf{1}_b \quad \text{and} \quad N^T\mathbf{1}_v = N'^T\mathbf{1}_v \quad . \tag{1}$$

If D is binary then $D(E)$ is binary if and only if E is trivial or both $p \notin y$ and $q \notin x$. An interchange pair with the latter property is called a *binary*-interchange pair (or b-interchange pair, for short). So, after performing a b-interchange, a binary design is again a binary design. Further, a binary equireplicate block design remains a binary equireplicate block design under b-interchanges. If $E = [(p, x), (q, y)]$ is b-interchange pair for such a design D then $E' = [(p, y), (q, x)]$ is a b-interchange pair for $D(E)$ and $D(E)(E') = D$. So, performing a b-interchange is a reversible operation.

Note that if $E_1 = [(p_1, x_1), (p_2, x_2)]$ and $E_2 = [(q_1, y_2), (q_2, y_2)]$ are two b-interchange pairs for D with x_1, x_2, y_1 and y_2 distinct blocks, then E_2 is a b-interchange pair for $D(E_1)$ (and vice versa) and also $D(E_1)(E_2) = D(E_2)(E_1)$. If D' is obtained from D by a sequence of b-interchanges, then D can also be obtained from D' by a sequence of b-interchanges. Thus the relation $D \sim D'$ when D' is obtainable from D by a sequence of b-interchanges is an equivalence relation. The aim of this paper is to classify the equivalence classes of the collection of all binary designs with v varieties and b blocks.

An *m-resolvable* design ($m \geq 1$) is an equireplicate block design D with replication number r and whose blocks $\mathcal{B}$ are partitioned into sets $\mathcal{B}_1, \ldots, \mathcal{B}_{r/m}$ such that each $\mathcal{B}_i$ ($1 \leq i \leq r/m$) is an equireplicate block design with replication number m.

Result 1 ([5]) *1. The collection of all proper equireplicate block designs with given parameters v, k, r is an equivalence class under interchanging.*

2. The collection of all 1-resolvable designs with given parameters is an equivalence class under interchanging using blocks of the same class only.

The aim of this note is to prove the corresponding results for the binary designs (not necessarily block or equireplicate) using only b-interchanges.

3 Main result

For fixed v and b ($v, b \geq 1$), let $\mathcal{C}$ be the collection of all binary designs with v varieties and b blocks.

For $0 \leq m \leq v$, we say that binary designs *agree* on $\mathcal{V}_m = \{1, 2, \ldots, m\}$ if the $m \geq 1$ varieties $1, \ldots, m$ are in exactly the same blocks in both designs. If $m = 0$ then the condition is regarded as empty. Note that two designs agree on $\mathcal{V}_v$ if and only if they are identical.

Theorem 1 *Designs $C, D \in \mathcal{C}$ are in the same equivalence class if and only if their incidence matrices N_C and N_D satisfy: $N_C \mathbf{1}_b = N_D \mathbf{1}_b$ and $N_C^T \mathbf{1}_v = N_D^T \mathbf{1}_v$.*

Proof. From (1) in Sect. 2, the condition is necessary. We show that it is sufficient.

Suppose $C, D \in \mathcal{C}$ satisfy the conditions on their incidence matrices given in the theorem. So each block $x \in \mathcal{B}$ contains the same number k_x of varieties in both C and D; similarly each variety $p \in \mathcal{V}$ is in the same number r_p blocks in both C and D.

We will give an algorithm to show that C and D are equivalent. Suppose m satisfies $0 \leq m < v$. We show what to do at Step m below and by doing Step 0 to Step $v-1$ we will show that C and D are equivalent.

At Step m:

Suppose that

$$C_m \text{ and } D_m \text{ agree on } \mathcal{V}_m \quad . \tag{2}$$

Suppose further

$$\begin{aligned} \text{in } C_m \ &: \ \text{variety } m+1 \in s_1, \ldots, s_u, y_{u+1}, \ldots, y_{r_{m+1}}, \\ \text{in } D_m \ &: \ \text{variety } m+1 \in s_1, \ldots, s_u, z_{u+1}, \ldots, z_{r_{m+1}}, \end{aligned}$$

where $s_1, \ldots, s_u, y_{u+1}, \ldots, y_{r_{m+1}}, z_{u+1}, \ldots, z_{r_{m+1}}$ are all distinct blocks and possibly some of the subsequences are empty. Consider y_i and z_i ($u < i \leq r_{m+1}$). In D_m, variety $m+1 \notin y_i$. If $y_i \backslash (z_i \cup \mathcal{V}_m) \neq \emptyset$ in D_m then let $q_i \in y_i \backslash (z_i \cup \mathcal{V}_m)$. By (2) $y_i \cap \mathcal{V}_m = z_i \cap \mathcal{V}_m$ so $q_i \notin \mathcal{V}_m$. $E_i = [(m+1, z_i), (q_i, y_i)]$ is a b-interchange pair for D_m, and then in both C_m and $D_m(E_i)$ we have variety $m+1 \in y_i$ and variety $m+1 \notin z_i$.

Suppose now that in D_m we have $y_i \backslash (z_i \cup \mathcal{V}_m) = \emptyset$. By (2) this holds if and only if $y_i \subseteq z_i$. Consider the situation in C_m. We have variety $m+1 \notin z_i$ by the definition of z_i. So variety $m+1 \in y_i \backslash z_i$ and as $y_i \subseteq z_i$ we have $k_{y_i} \leq k_{z_i}$ and so $z_i \backslash y_i \neq \emptyset$. Let $q_i \in z_i \backslash y_i$, and by (2) $q_i \notin \mathcal{V}_m$. $F_i = [(m+1, y_i), (q_i, z_i)]$ is a b-interchange pair for C_m, and in both $C_m(F_i)$ and D_m we have variety $m+1 \in z_i$ and variety $m+1 \notin y_i$.

Calculate all the b-interchange pairs E_i or F_i obtained in this manner ($u < i \leq r_{m+1}$). None of the varieties in the b-interchange pairs include any varieties from $\mathcal{V}_m$. The b-interchange pairs are all on distinct blocks so we can perform them in any order: let C_{m+1} be the design obtained from C_m by the F_i's and let D_{m+1} the design obtained from D_m by the E_i's. So we have $C_m \sim C_{m+1}$, $D_m \sim D_{m+1}$ and

$$C_{m+1} \text{ and } D_{m+1} \text{ agree on } \mathcal{V}_{m+1} \quad .$$

Thus we have the conditions to apply Step $m+1$.

Let $C_0 = C$, $D_0 = D$. C_0 and D_0 satisfy (2) for $m = 0$, since in this case the condition is empty. Apply Step 0 as described above to obtain C_1, D_1 and then Step 1, Step 2 and so on until after Step $v-1$, C_v and D_v agree on $\mathcal{V}$, that is, $C_v = D_v$. So we have $C = C_0 \sim C_1 \sim \cdots \sim C_v = D_v \sim \cdots \sim D_0 = D$, that is, C and D are in the same equivalence class, as required. □

Corollary 1 *There is a one to one correspondence between equivalence classes of $\mathcal{C}$ and pairs $(\mathbf{r}, \mathbf{k})$ of vectors $\mathbf{r} = (r_1, \ldots, r_v)^T$ and $\mathbf{k} = (k_1, \ldots, k_b)^T$ with*

$$\sum_{i=1}^{v} r_i = \sum_{j=1}^{b} k_j \quad . \tag{3}$$

The correspondence is given by: $(\mathbf{r}, \mathbf{k})$ corresponds to collection of (binary) designs D with incidence matrix N_D satisfying

$$N_D \mathbf{1}_b = \mathbf{r} \quad \text{and} \quad N_D^T \mathbf{1}_v = \mathbf{k} \quad . \tag{4}$$

Proof. By counting the varieties of a design D in two ways, if D satisfies (4) then D also satisfies (3). The corollary now follows from Theorem 1. □

Corollary 2 *For a fixed v, k and r, the collection of all binary equireplicate block designs is an equivalence class under b-interchanges.*

Proof. Such designs D are characterised by their incidence matrix N_D satisfying $N_D \mathbf{1}_v = r\mathbf{1}_v$ and $N_D^T \mathbf{1}_b = k\mathbf{1}_b$. □

Corollary 3 *The collection of all binary m-resolvable designs with given parameters v, k, r is an equivalence class under b-interchanges involving blocks of the same class.*

Proof. The set of blocks in any given block class is a binary equireplicate block design with block size k and replication number r/m. The collection of these designs is an equivalence class under b-interchanges. □

Note that for a 1-resolvable design, every interchange involving blocks from the same class is a b-interchange.

4 Comments

For $t \geq 2$, a t–(v, k, λ) design D is a block design with v varieties, block size k such that every t distinct varieties is on λ blocks. Such a design has

$$b = \lambda \frac{v(v-1)\cdots(v-t+1)}{k(k-1)\cdots(k-t+1)}$$

blocks [3, Corollary 1.4, p7]. Also from [3, Theorem 1.2, p6] D is equireplicate with replication number $r = bk/v$. Thus D belongs to the equivalence class with v varieties, b blocks and with $\mathbf{k} = k\mathbf{1}_b$ and $\mathbf{r} = r\mathbf{1}_v$.

A further comment. By Corollary 2, the equivalence class containing an equireplicate block design with v varieties, b blocks, replication number r and block size k includes *all* designs with these parameters. In particular, if D and D' are isomorphic such designs (by *isomorphic* we mean there exist permutations α on $\mathcal{V}$, β on $\mathcal{B}$ with $P \in x$ in D if and only if $P^\alpha \in x^\beta$ in D') then D' can be obtained from D by b-interchanges. This does not necessarily hold for other equivalence classes. In other words, isomorphic designs are not necessarily in the same equivalence class. For example, consider the designs C and D whose incidence matrices are:

$$N_C = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \qquad N_D = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} .$$

They are isomorphic with α being the identity on $\mathcal{V}$ and $\beta{:}\,1 \mapsto 1$ and $2 \leftrightarrow 3$, where $\mathcal{B} = \{1, 2, 3\}$. As $N_C^T\mathbf{1}_v \neq N_D^T\mathbf{1}_v$ they are not in the same equivalence class (Theorem 1).

5 Acknowledgements

I am grateful to Bill Venables for introducing the problem to me, and for his encouragement and discussion.

References

[1] J. A. Eccleston and D. Whitaker: Multi-objective simulated annealing for the construction of contiguous row-column designs, July 1993, Centre of Statistics Research Report, The University of Queensland

[2] B. Jones and J. A. Eccleston: Exchange and interchange procedures to search for optimal designs, J. Royal Satist. Soc. B **42** (1980) 238–243

[3] D. R. Hughes and F. C. Piper: Design theory, Cambridge University Press, 1985.

[4] A. P. Street and D. J. Street: Combinatorics of experimental design, Clarendon Press, Oxford, 1987.

[5] W. N. Venables and J. A. Eccleston: Randomized search strategies for finding optimal or near optimal block and row-column designs, Australian J. Statist. **35** (1993) 371–382

(Received 3/1/94)