Journal of Statistical Planning and Inference 140 (2010) 2771–2777

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Optimal efficiency balanced designs and their constructions

Sun Tianyao, Tang Yu *

School of Mathematical Sciences, Suzhou University, China

ARTICLE INFO

Article history:
Received 5 November 2009
Received in revised form
20 March 2010
Accepted 22 March 2010
Available online 30 March 2010

MSC:
62K10

Keywords:
Block design
Combinatorial design
Efficiency balanced design
Experimental design
Optimal

ABSTRACT

Efficiency balanced design can be used to increase the efficiency of an experimental arrangement. In this manuscript, two classes of non-proper binary efficiency balanced designs are discussed. Properties of optimal efficiency balanced designs in these two classes are provided and construction methods of designs satisfying these properties are also given. Combined with known results and construction techniques in combinatorial design theory, two infinite classes of optimal efficiency balanced designs with certain parameters are obtained.

1. Introduction

The modern approach to the experimental design was introduced by Fisher in 1920s'. It is based on three basic principles which are "replication", "randomization" and "local control". The main purpose of local control is to increase the efficiency of a design by decreasing the experimental error. The general way to conduct "local control" designs is to use "blocks". Consider v treatments arranged in b blocks with the j th block size $k_{(j)}$ (j= 1,2,3,…,b). Denote $r_{(i)}$ (i=1,2,…,v) as the number of times for the i th treatment occurring in those b blocks. The design can be described by a $v \times b$ incidence matrix N=(n_{ij}), where n_{ij} represents the number of units in the j th block receiving the i th treatment (j=1,2,3,…,b; i=1,2,…,v). When n_{ij} can take any of the values 0,1,…,$q-1$, the design is called a q-ary block design. If q=2, the design is called binary. When $r_{(1)}=r_{(2)}=\cdots=r_{(v)}$ or $k_{(1)}=k_{(2)}=\cdots=k_{(b)}$, the design is said to be equireplicate or proper, respectively.

The matrix $C=\mathbf{r}^{\delta}-N\mathbf{k}^{-\delta}N'$ is often called the "coefficient matrix", where $\mathbf{r}^{\delta}$ and $\mathbf{k}^{\delta}$ are the diagonal matrices of treatment replications and block sizes, i.e., $\mathbf{r}^{\delta}=\operatorname{diag}\{r_{(1)},r_{(2)},\ldots,r_{(v)}\}$ and $\mathbf{k}^{\delta}=\operatorname{diag}\{k_{(1)},k_{(2)},\ldots,k_{(b)}\}$, while $\mathbf{k}^{-\delta}$ is the inverse matrix of $\mathbf{k}^{\delta}$. It is clear that $C\,1_v=0$, so $rank(C)\le v-1$ and C is a positive semidefinitive matrix. A design is said to be connected if and only if $rank(C)=v-1$ (Das and Kageyama, 1991). A connected design also means that for any two treatments i and i', there exists a chain of treatments $i=i_0,i_1,\ldots,i_m=i'$, such that any two consecutive treatments in the chain occur together in at least one block. We always assume that the design is connected in the following discussion.

Denote $\mathbf{r}^{-\delta}$ as the inverse matrix of $\mathbf{r}^{\delta}$. The eigenvalues of $\mathbf{r}^{-\delta}C$ have been extensively investigated by existing literature. Following James and Wilkinson (1971), for any connected design, the $v-1$ nonzero eigenvalues of $\mathbf{r}^{-\delta}C$ are called the canonical efficiency factors, denoted by e_i's (i=1,2,…,$v-1$). There are many statistical implications between canonical

* Corresponding author.
E-mail address: ytang@suda.edu.cn (T. Yu).

0378-3758/$ - see front matter
doi:10.1016/j.jspi.2010.03.039

efficiency factors and basic contracts (Caliński and Kageyama, 2000, Section 3.4). When those e_i's are all equal, the design is called efficiency balanced (EB). If the design is equireplicate, efficiency balanced is equivalent to variance balanced, which means every elementary contrast of treatments is estimated with the same variance. It has been showed in Williams (1975) that a proper connected design is efficiency balanced if and only if the incidence matrix N satisfied

$$NN' = k(1-e)\mathbf{r}^{-\delta} + \frac{e}{b}\mathbf{r}\mathbf{r}', \tag{1}$$

where e is the common eigenvalue of $\mathbf{r}^{-\delta}C$, with multiplicity $v-1$. When it comes to a non-proper connected design, the condition should be modified as follows

$$N\mathbf{k}^{-\delta}N' = (1-e)\mathbf{r}^{\delta} + \frac{e}{n}\mathbf{r}\mathbf{r}'. \tag{2}$$

The properties for a class of proper binary optimal EB designs and a class of generalized proper binary optimal EB designs were considered by Das and Kageyama (1991) and Das (1998), respectively. Both of these two papers dealt with the proper case, and to our best knowledge, little work has been reported on non-proper optimal EB designs. However, as stated in Kageyama (1976), non-proper designs may be reasonable choices in some circumstances. The present manuscript will mainly discuss the properties for two classes of non-proper binary optimal EB designs. From a practical viewpoint, there should be restrictions on the number of experimental units, as the cost of experiment is an important problem we must concern. Throughout the manuscript, we aim to provide optimal designs in classes of EB designs, which have the same number of treatments, the same experimental units, the same block sizes and the same replication numbers. We denote by $EBD(v_1,v_2,n,k_1,k_2,r_1,r_2)$ the class of connected binary EB design having $v=v_1+v_2$ treatments and block sizes k_1,k_2 ($k_1<k_2$), with $v_1(v_2)$ replication numbers $r_1(r_2)$ ($r_1<r_2$). Particularly, when $r_1=r_2=r$, we denote it by $EBD(v,n,k_1,k_2,r)$.

2. Character of optimal EB design

The harmonic mean of nonzero eigenvalues of $\mathbf{r}^{-\delta}C$ represented by e is said to be the efficiency factor of the design. If the design is connected, then $e=(v-1)/\sum_{j=1}^{v-1} e_i^{-1}$. Clearly,

$$e \le (v-1)^{-1}\sum_{i=1}^{v-1} e_i = (v-1)^{-1}\,\mathrm{tr}(\mathbf{r}^{-\delta}C) \tag{3}$$

the equality holds if and only if the design is EB.

Definition 2.1. An efficiency balanced design is said to be optimal if it maximizes the efficiency factor in a class of EB designs with the same parameters.

There are many other definitions of optimality (for example, E-optimality, A-optimality, D-optimality), see Caliński and Kageyama (2003, Section 7.1). Uddin (1996) discussed E-optimal designs in the class $D(v,b_1,b_2,k_1,k_2)$ of incomplete block designs for v treatments in b_1 blocks of size k_1 each and b_2 blocks of size k_2 each. Das (1998) gave the optimality of GB-EB design in a class of block designs which had the same number of treatments, the same number of blocks, the same block sizes and the same replication numbers, denote by $D(v_1,v_2,b,k,r_1,r_2)$. While here an optimal EB design in Definition 2.1 is restricted to a class of EB designs with the same parameters.

It is well known that a binary proper connected EB design is a balanced incomplete block design (BIBD), which is equireplicate (Caliński and Kageyama, 2003, Section 8.2). However, when coming to the non-proper case, it will be more complicated. In this section, we will first consider the optimal EB design in the class $EBD(v,n,k_1,k_2,r)$, then generalize the idea to the class $EBD(v_1,v_2,n,k_1,k_2,r_1,r_2)$.

Take the trace of (2), we can get $e=[n-\mathrm{tr}(N\mathbf{k}^{-\delta}N')]/(n-n^{-1}\mathbf{r}'\mathbf{r})$. If the design is binary, $\mathrm{tr}(N\mathbf{k}^{-\delta}N')=\sum_{i=1}^{v}\sum_{j=1}^{b} n_{ij}^2/k_j=\sum_{i=1}^{v}\sum_{j=1}^{b} n_{ij}/k_j=b$, then

$$e=\frac{n-b}{n-n^{-1}\mathbf{r}'\mathbf{r}}. \tag{4}$$

Let b_1 and b_2 be numbers of blocks with size k_1 and k_2, respectively. The facts that $k_1b_1+k_2b_2=n$ and $b_1+b_2=b$ imply $b=(1-k_1/k_2)b_1+n/k_2$ and

$$e=\frac{(k_1-k_2)n}{k_2(n^2-\mathbf{r}'\mathbf{r})}b_1+\frac{(k_2-1)n^2}{k_2(n^2-\mathbf{r}'\mathbf{r})}. \tag{5}$$

What Eq. (5) says is that an EB design with the minimum b_1 is an optimal EB design which maximizes the efficiency factor in a class of EB designs with the same number of treatments, the same number of experimental units, the same two block sizes and the same replication numbers, under the condition $k_1<k_2$.

Remark. Notice that two designs in the same class of EB designs may have different efficiency factors. For example, let $\mathcal{A}_1=\{(1,2,3,4)\}$ and $\mathcal{A}_2=\{(1,2),(1,3),(1,4),(2,3),(2,4),(3,4)\}$. Take $\mathcal{B}_1$ as four copies of $\mathcal{A}_1$ and one copy of $\mathcal{A}_2$, and $\mathcal{B}_2$ as one copy of $\mathcal{A}_1$ and two copies of $\mathcal{A}_2$. Simple calculation shows that $\mathcal{B}_1$ and $\mathcal{B}_2$ both belong to the class EBD (4,28,2,4,7), while the efficiency factors of these two designs are 0.8571 and 0.7143, respectively. Thus, the design $\mathcal{B}_1$ should be better with