Journal of Statistical Planning and Inference 139 (2009) 2823 – 2835

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

Optimum covariate designs in partially balanced incomplete block (PBIB) design set-ups

Ganesh Dutta[a], Premadhis Das[b], Nripes K. Mandal[c,*]

[a] Department of Statistics, Basanti Devi College, 147/B, Rashbehari Avenue, Kolkata 700029, India
[b] Department of Statistics, University of Kalyani, Kalyani 741235, India
[c] Department of Statistics, University of Calcutta, 35 B.C. Road, Kolkata 700019, India

ARTICLE INFO

Article history:
Received 29 November 2007
Received in revised form
4 October 2008
Accepted 9 January 2009
Available online 21 January 2009

MSC:
primary 62K05
secondary 62K10

Keywords:
Block designs
Covariates
Optimal designs
Orthogonal arrays
Hadamard matrices
Khatri–Rao product
Kronecker product

ABSTRACT

The use of covariates in block designs is necessary when the covariates cannot be controlled like the blocking factor in the experiment. In this paper, we consider the situation where there is some flexibility for selection in the values of the covariates. The choice of values of the covariates for a given block design attaining minimum variance for estimation of each of the parameters has attracted attention in recent times. Optimum covariate designs in simple set-ups such as completely randomised design (CRD), randomised block design (RBD) and some series of balanced incomplete block design (BIBD) have already been considered. In this paper, optimum covariate designs have been considered for the more complex set-ups of different partially balanced incomplete block (PBIB) designs, which are popular among practitioners. The optimum covariate designs depend much on the methods of construction of the basic PBIB designs. Different combinatorial arrangements and tools such as orthogonal arrays, Hadamard matrices and different kinds of products of matrices viz. Khatri–Rao product, Kronecker product have been conveniently used to construct optimum covariate designs with as many covariates as possible.

1. Introduction

Consider the non-stochastic controllable covariates model in a block design set-up

$$(\mathbf{Y}, \mu\mathbf{1}_N + \mathbf{X}_1\boldsymbol{\beta} + \mathbf{X}_2\boldsymbol{\tau} + \mathbf{Z}\gamma, \sigma^2\mathbf{I}), \tag{1.1}$$

where μ is the intercept term, σ^2 is the common variance of the observations, $\boldsymbol{\beta}^{b\times 1}$, $\boldsymbol{\tau}^{v\times 1}$ and $\gamma^{c\times 1}$ correspond, respectively, to the vectors of block effects, treatment effects and covariate effects. $\mathbf{Y}$ is the uncorrelated observation vector of order $N\times 1$, $\mathbf{X}_1$, $\mathbf{X}_2$, respectively, being the incidence matrices of block effects, treatment effects and $\mathbf{Z}$ is the design matrix corresponding to the covariate effects and $\mathbf{1}_N$ is a vector of order N with all elements unity. For the covariates, it is assumed without loss of generality, the (location-scale)-transformed version: $|z_{ij}|\leqslant 1$. It is evident that for orthogonal estimation of treatment and block effect contrasts on one hand and covariate effects on the other, it is necessary and sufficient that

$$\mathbf{Z}'\mathbf{X}_1 = \mathbf{0}, \quad \mathbf{Z}'\mathbf{X}_2 = \mathbf{0} \tag{1.2}$$

* Corresponding author.
E-mail address: mandalnk2001@yahoo.co.in (N.K. Mandal).

0378-3758/$ - see front matter
doi:10.1016/j.jspi.2009.01.009

and for most efficient estimation of each of the regression parameters the following condition must also be satisfied (cf. Pukelsheim, 1993):

$$\mathbf{Z}'\mathbf{Z} = N\mathbf{I}_c. \tag{1.3}$$

A Covariate design **Z** *is said to be optimum if for given* $\mathbf{X}_1$, $\mathbf{X}_2$ *it satisfies* (1.2) *and* (1.3). It is evident from (1.2) and (1.3) that (i) the columns of **Z** should be orthogonal to the columns of $\mathbf{X}_1$ and $\mathbf{X}_2$ and (ii) the columns of **Z** must be mutually orthogonal among themselves and the elements of **Z** must be ± 1.

The problem of choice of covariates in an experimental design set-up was earlier considered by Lopes Troya (1982a, b), Liski et al. (2002), Das et al. (2003), Rao et al. (2003), Dutta (2004) and Dutta et al. (2007, 2009). Lopes Troya (1982a, b) first considered the problem of choice of the **Z** matrix in a completely randomised design (CRD) model. Das et al. (2003) extended it to the set-up of randomised block designs (RBDs) and to some series of balanced incomplete block designs (BIBDs). Rao et al. (2003) re-visited the problem in CRD and RBD set-ups and identified the solutions as mixed orthogonal arrays, thereby providing further insights and some new solutions. As mentioned earlier, the choice of covariate designs depends heavily on the block design set-up as is evidenced from (1.2). In the case of incomplete block (IB) designs, the allocation of treatments to the plots of the blocks depends on the method of construction of designs. Das et al. (2003) considered symmetric balanced incomplete block designs (SBIBDs) with parameters $b = v$, $r = k$, λ constructed through Bose's difference method and some BIBDs with repeated blocks. Dutta (2004) also considered some series of BIBDs obtained through Bose's difference technique together with some arbitrary BIBDs. Thereafter, Dutta et al. (2007) considered a series of SBIBDs which is obtained through projective geometry. However, as is well known, there are different methods of construction leading to different IB designs. It is very difficult to construct optimum covariate design for arbitrary IB design and the choice of the **Z** matrix varies from method to method. It seems that the partially balanced incomplete block (PBIB) design set-up has not been considered previously though such design is very often used in practice. Moreover, the series of PBIB designs considered in this article are obtained not only through the method of differences but also are obtained by other methods as described by Bose et al. (1953), Zelen (1954) and Vartak (1954).

Following Das et al. (2003), each column of the **Z** matrix can be recast to a **W**-matrix where the element in the ith row and jth column of $\mathbf{W}^{(s)}$ is $z_{ij}^{(s)}$, $z_{ij}^{(s)}$ being the element of **Z** corresponding to the jth treatment in the ith block of the design for the sth covariate. Corresponding to the block-treatment classification, optimality conditions (1.2) and (1.3) in terms of **W**-matrices reduce to:

(C_1) each **W**-matrix has all column-sums equal to zero;
(C_2) each **W**-matrix has all row-sums equal to zero;
(C_3) the grand total of all the entries in the Hadamard product (vide Rao, 1973) of any two distinct **W**-matrices reduces to zero.

Henceforth, the conditions C_1–C_3 *together are referred to as the single condition C*.

It is to be noted that a covariate design **Z** for c covariates is equivalent to c **W**-matrices which are convenient to work with.

Definition 1.1. With respect to model (1.1), the c **W**-matrices corresponding to the c covariates are said to be optimum if they satisfy the condition C.

Remark 1.1. It is to be noted that if $c = 1$, only the conditions C_1 and C_2 are to be satisfied by the **W**-matrix to be optimum.

In an IB design set-up, the basic principle of constructing optimum **W**-matrices i.e. the optimum covariate design is to convert the incidence matrix **N** of the IB design with parameters b, v, r and k by placing judiciously ± 1's in the non-zero cells of the incidence matrix so that the **W**-matrices satisfy the condition C mentioned above.

2. Optimum covariate designs

In this section, construction of **W**-matrices satisfying the condition C in different series of PBIB designs, obtainable through different constructional methods, are given.

2.1. Singular group divisible (SGD) design set-up

It had been shown in Bose et al. (1953) that if in a BIBD with parameters v^*, b^*, r^*, k^* and λ^* each treatment is replaced by a group of n treatments, a SGD design can be obtained with parameters

$$v = nv^*, \quad b = b^*, \quad r = r^*, \quad k = nk^*, \quad m = v^*, \quad n = n, \quad \lambda_1 = r^*, \quad \lambda_2 = \lambda^*. \tag{2.1}$$

Here m stands for the number of groups in the corresponding association scheme. It will be seen that **W**-matrices for such a SGD design with parameters in (2.1) can be constructed and the construction of **W** in this case does not depend on the method of construction of the corresponding BIBD. Precise statement follows.