Comparison of test vs. control treatments using distance optimality criterion

Nripes K. Mandal[1], **Kirti R. Shah**[2] **and Bikas K. Sinha**[3]

[1] Department of Statistics, Calcutta University, 35, Ballygunge Circular Road, Calcutta 700019, India

[2] Department of Statistics and Actuarial Science, University of Waterloo, 200 University Av. West, Waterloo, N2L 3G1, Canada (e-mail: kshah@math.uwaterloo.ca)

[3] Stat-Math Division, Indian Statistical Institute, 203, B.T. Road, Calcutta 700035, India

Received: December 1999

Abstract. We consider the problem of comparison of one test treatment (τ_0) with a set of v control treatments ($\tau_1, \tau_2, \ldots, \tau_v$) using distance optimality [DS-optimality] criterion introduced by Sinha (1970) in some treatment-connected design settings. It turns out that the nature of DS-optimal designs is quite similar to that for the usual $A-$, $D-$ and $E-$ optimality criteria. However, the optimality problem is quite complicated in most situations. First we deal with the CRD model and derive DS-optimal allocations for a given set of treatments. The results are almost identical to the A-optimal allocations for such problems. Then we consider a block design set-up and examine the nature of DS-optimal designs. In the process, we introduce the method of weighted coverage probability and maximize the resulting expression to obtain an optimal design.

Key words: Okamoto Lemma, Weighted Coverage Probabilities, Completely Randomized Design, Block Designs, Complete Classes of Designs.

1 Introduction

Our purpose here is to apply the distance optimality [DS-optimality] criterion for inference on the vector of parameters given by

$$\boldsymbol{\eta} = (\tau_0 - \tau_1, \tau_0 - \tau_2, \ldots, \tau_0 - \tau_v)^t \tag{1.1}$$

where τ_0 refers to the effect of the test treatment and $\tau_1, \tau_2, \ldots, \tau_v$ refer to those of the control treatments.

These are known as treatment-control designs. In the literature, both types of comparisons, viz., comparing a set of test treatments with one or more control as well as comparing a set of controls with one or more test treatments, have been considered. See Majumdar (1996) and references therein.

Let $\hat{\boldsymbol{\eta}}$ denote the BLUE of $\boldsymbol{\eta}$ in a given design context. Then, according to the DS-optimality criterion, we seek to

$$\text{minimize } P[\|\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}\| > \varepsilon] \tag{1.2}$$

by a proper choice of the "design'' out of the class of competing designs, uniformly in $\varepsilon > 0$. In the above, $\| \cdot \|$ refers to the Euclidean distance.

This criterion was put forward by Sinha (1970) who established that under the normality assumption for the errors, in the completely randomized design setting with a given number (n) of experimental units:

$$y_{ij} = \mu_i + \varepsilon_{ij}, \tag{1.3}$$

$1 \leq i \leq v$, $1 \leq j \leq n_i$, $\sum_1^v n_i = n$, the design to minimize $P[\|\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}\| > \varepsilon]$ uniformly in $\varepsilon > 0$ is the completely symmetric design (when $n = \sum_1^v n_i$ is divisible by v). When n is not divisible by v, the most symmetric allocation (i.e., one for which n_i's differ at most by unity) is one of the prominent competitors. The structure of other competing designs was also given in the form of a complete class of designs. In the above, $\boldsymbol{\mu}$ is the vector of treatment effects in (1.3).

In the block design setting, the balanced incomplete block design (BIBD) minimizes $P[\|\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}\| > \varepsilon]$ uniformly in $\varepsilon > 0$, where $\boldsymbol{\eta} = \boldsymbol{P\tau}$ refers to a full set of orthonormal treatment contrasts.

Without any further reference to it, we shall assume that the errors are independently normally distributed.

It turns out that the usual canonical reduction of the problem leads to:

$$\text{maximise } P\left[\sum \lambda_i Z_i^2 < \varepsilon^2\right] \text{ uniformly in } \varepsilon > 0 \tag{1.4}$$

where λ_i's are the eigenvalues of the dispersion matrix of the BLUE's of the vector of parameters under consideration and Z_i's are i.i.d. $N(0,1)$. A result due to Okamoto (1960) – also to be found in Marshall and Olkin (1979) – is of immediate applicability in this context. This is stated below for the sake of completeness.

Okamoto Lemma:

$$P\left[\sum_1^k \lambda_i Z_i^2 < \varepsilon^2\right] \leq P[\lambda \chi_k^2 < \varepsilon^2] \tag{1.5}$$

where χ_k^2 refers to central χ^2-variate with k d.f. and

$$\lambda = (\Pi \lambda_i)^{1/k}. \tag{1.6}$$

Recently, Liski et al. (1998) initiated a study of optimal regression designs using the DS-optimality criterion. In Sinha (1970) as also in Liski et al. (1998) some other corollaries/generalizations of Okamoto's Lemma have been found useful. These are stated below without proof.

Corollary 1:

$$P\left[\sum_{1}^{k}\lambda_i Z_i^2 < \varepsilon^2\right] \leq P[\lambda'\chi^2_{k'} + \lambda''\chi^2_{k''} < \varepsilon^2] \tag{1.7}$$

where,

$$k = k' + k'', \lambda' = (\Pi_{i=1}^{k'}\lambda_i)^{1/k'} \quad \text{and} \quad \lambda'' = (\Pi_{i=k'+1}^{k}\lambda_i)^{1/k''}. \tag{1.8}$$

Generalization 1:

$$P[\lambda_1 Z_1^2 + \lambda_2 Z_2^2 < \varepsilon^2] \leq P[\mu_1 Z_1^2 + \mu_2 Z_2^2 < \varepsilon^2] \tag{1.9}$$

provided that

$$\lambda_1\lambda_2 \geq \mu_1\mu_2 \quad \text{and} \quad \lambda_1 \vee \lambda_2 \geq \mu_1 \vee \mu_2. \tag{1.10}$$

Generalization 2:

$$P[\lambda_1^*\chi^2_{v_1} + \lambda_2^*\chi^2_{v_2} < \varepsilon^2] \leq P[\lambda^{**}\chi^2_{v_1} + \lambda_2^{**}\chi^2_{v_2} < \varepsilon^2] \tag{1.11}$$

where

$$(\lambda_1^*)^{v_1}(\lambda_2^*)^{v_2} > (\lambda_1^{**})^{v_1}(\lambda_2^{**})^{v_2} \tag{1.12}$$

and

$$\lambda_1^* \vee \lambda_2^* > \lambda_1^{**} \vee \lambda_2^{**} > \lambda_1^{**} \wedge \lambda_2^{**} > \lambda_1^* \wedge \lambda_2^*. \tag{1.13}$$

For the model (1.3), in a situation where n is *not* divisible by v, the most symmetric allocation (msa) turns out to be DS-optimal in view of Generalization 1, once the complete class of designs is characterized as in Sinha (1970). This completely settles the problem dealt by Sinha (1970) who deduced the optimality of msa only when $n = \frac{v}{2} (\mathrm{mod}\, v)$. Recently, Liski et al. (1999) studied further properties of the distance optimality criterion.

In this paper we attempt a solution to the problem formulated in (1.2) in both the CRD and the Block design settings. It turns out that the DS-optimal designs do depend on the parameter ε^2 and, therefore, we need to bring the concept of weighted coverage probability and minimize it. In the CRD setting, the DS-optimal designs are not different from A-optimal designs. In the block design settings, designs with similar structure of the **C**-matrices as for the A-optimal designs are found to be DS-optimal.

2 Control-Treatment comparisons in a CRD model

As is well-known, there has been considerable amount of study in the characterization and construction of optimal design for inference on $\boldsymbol{\eta}$ in (1.1) w.r.t the usual optimality criteria such as $A-, D-$ and $E-$ optimality. We refer to

Hedayat et al. (1988) and Majumdar (1996) for a comprehensive review of these results.

In this section we confine ourselves to the DS-optimality criterion for a completely randomized design (CRD) model.

Let $n_0, n_1, \ldots, n_v$ be the allocation numbers subject to $\sum_0^v n_i = n$. We set $p_i = n_i/n$, $i = 0, 1, 2, \ldots, v$ so that $\sum_0^v p_i = 1$. In the sense of the approximate design theory, we shall seek optimal values of the p_i's so as to satisfy (1.2) uniformly in $\varepsilon > 0$.

It is evident that

$$D(\hat{\boldsymbol{\eta}}) = \sigma^2 \left[\boldsymbol{n}^{-\delta} + \frac{\boldsymbol{J}_{vv}}{n_0} \right] \tag{2.1}$$

where

$$\boldsymbol{n}^{-\delta} = \operatorname{diag}\left(\frac{1}{n_1}, \frac{1}{n_2}, \ldots, \frac{1}{n_v} \right) \quad \text{and} \quad \boldsymbol{J}_{vv} = ((1))_{v \times v} \tag{2.2}$$

It is easy to see that the eigenvalues of $D(\hat{\boldsymbol{\eta}})/\sigma^2$ satisfy

$$\Pi_1^v \lambda_i = |\boldsymbol{n}^{-\delta} + \boldsymbol{J}_{vv}/n_0| = \frac{1}{\Pi n_i} \left(\frac{n}{n_0} \right) \geq \left(\frac{v}{n - n_0} \right)^v \cdot \frac{n}{n_0}. \tag{2.3}$$

Further,

$$\begin{aligned} \lambda_{\max} &\geq \frac{\mathbf{1}'(\boldsymbol{n}^{-\delta} + \boldsymbol{J}/n_0)\mathbf{1}}{\mathbf{1}'\mathbf{1}} = \frac{\sum_1^v 1/n_i + v^2/n_0}{v} \\ &\geq \frac{v^2/(n - n_0) + v^2/n_0}{v} = \frac{nv}{n_0(n - n_0)}. \end{aligned} \tag{2.4}$$

In (2.3) and (2.4), "=" holds whenever $n_1 = n_2 = \cdots = n_v = (n - n_0)/v$ for every *fixed* n_0.

We now refer to the probability inequality (1.11). We set

$$\begin{aligned} &v_1 = 1, v_2 = v - 1, \lambda_1^* = \lambda_{\max}, \lambda_2^* = (\Pi \lambda_i / \lambda_{\max})^{1/(v-1)} \\ &\lambda_1^{**} = nv/n_0(n - n_0) \quad \text{and} \quad \lambda_2^{**} = v/(n - n_0). \end{aligned} \tag{2.5}$$

We now consider two cases. We first consider the case where $\lambda_2^* \geq \lambda_2^{**}$. In this case an application of Corollary 1 gives

$$\begin{aligned} P\left[\sum_1^v \lambda_i Z_i^2 < \varepsilon^2 \right] &\leq P[\lambda_{\max} \chi_1^2 + \lambda_2^* \chi_{v-1}^2 \leq \varepsilon^2] \\ &\leq P[\lambda_1^{**} \chi_1^2 + \lambda_2^{**} \chi_{v-1}^2 \leq \varepsilon^2]. \end{aligned} \tag{2.6}$$

Here, the last inequality follows by implication of events. If on the other hand, $\lambda_2^* < \lambda_2^{**}$, conditions (1.12) and (1.13) are satisfied and hence, the above result follows from application of Corollary 1 followed by that of Generalization 2. Thus, (2.6) holds in all cases.

Thus, for every fixed n_0, the allocation $\left(n_0, \frac{n-n_0}{v}, \ldots, \frac{n-n_0}{v}\right)$ is uniformly (in $\varepsilon > 0$) better than $(n_0, n_1, \ldots, n_v)$ for all choices of n_i's subject to $\sum_1^v n_i = n - n_0$. Hence for fixed n_0 the allocation $\left(n_0, \frac{n-n_0}{v}, \ldots, \frac{n-n_0}{v}\right)$ is essentially complete in the sense that it is uniformly better than any other allocation $(n_0, n_1, \ldots, n_v)$, $\sum_1^v n_i = n - n_0$.

Denoting the RHS of (2.6) by $P(n, v, n_0, \varepsilon)$ we have

$$P(n, v, n_0, \varepsilon) = P\left[\frac{nv}{n_0(n-n_0)}\chi_1^2 + \frac{v}{n-n_0}\chi_{v-1}^2 < \varepsilon^2\right]. \tag{2.7}$$

According to the approximate design theory, we try to choose $p_0 (= n_{0/n})$ optimally i.e. in such a way that we maximize

$$P(v, p_0, \varepsilon) = P[\chi_1^2 + p_0\chi_{v-1}^2 \leq p_0(1-p_0)\varepsilon^2] \tag{2.8}$$

for *every* given $\varepsilon > 0$. The above expression (2.8) follows from (2.7) where ε is used as a "generic" notation .

We carried out extensive computations to determine the optimum value of p_0 i.e., the value of p_0 which maximizes the coverage probability P for a given value of ε^2. The results of these computations are presented in the following table.

Table 2.1. Optimal values of p_0

$v = 5$			$v = 10$			$v = 15$		
ε^2	p_0	P	ε^2	p_0	P	ε^2	p_0	P
5	0.23	0.2622	15	0.21	0.4954	20	0.18	0.4439
10	0.28	0.6006	20	0.24	0.7108	30	0.20	0.7802
15	0.31	0.7962	25	0.26	0.8393	35	0.22	0.8738
20	0.32	0.9002	30	0.27	0.9129	40	0.25	0.9329
25	0.37	0.9523	35	0.28	0.9534	45	0.26	0.9642
30	0.37	0.9757	40	0.34	0.9766	55	0.32	0.9894
35	0.36	0.9885	45	0.32	0.9877	65	0.34	0.9977
40	0.40	0.9947	50	0.37	0.9957	75	0.37	0.9993
45	0.44	0.9976	55	0.34	0.9973	85	0.37	0.9999
50	0.42	0.9991						

The above computations were based on 10,000 simulations and hence the values of p_0 are only approximate. However, our computations indicated the following.

(a) The values of P are nearly constant in a broad range of values of p_0 which are close to the optimum value. This is especially true for large values of P.
(b) As ε^2 increases both the optimum value of p_0 and the corresponding value of P increases at least initially. After that, optimum value of p_0 appears to decrease whereas the value of P continues to increase.
(c) No single value of p_0 maximizes P for all values of ε^2.
(d) Since A-optimal allocations satisfy the relation $n_0 = \sqrt{v}/n_1$ [vide Majumdar (1996)], the DS-optimal allocations are almost the same as A-optimal allocations.
(e) Since DS-optimal allocation tends to be E-optimal as ε^2 tends to infinity [vide Erkki et al. (1999)], it is possible to obtain approximately E-optimal allocations by choosing ε. These computations indicate that it would be desirable to choose a value of p_0 which maximizes P averaged w.r.t. an appropriate weight function for ε. We address this problem in the next section.

3 Weighted coverage probabilities

We shall consider the coverage probabilities discussed in the previous section *averaged* w.r.t. ε using a probability distribution for ε. Prompted by the expression (2.7), we shall start with the p.d.f. of ε^2 as that of χ^2_2 i.e., $\frac{1}{2}e^{-\varepsilon^2/2}$ for $0 < \varepsilon < \infty$. Subsequently, we shall show that our results can be applied for a large class of density functions for ε.

The weighted coverage probability is obtained by integrating the RHS of (2.7) w.r.t. ε using the p.d.f. for ε^2 given above. Let $\lambda_1 = v/np_0(1-p_0)$ and $\lambda_2 = v/n(1-p_0)$. The weighted coverage probability is given by

$$\bar{P}_{v,\lambda_1,\lambda_2} = \int_0^\infty P(\lambda_1\chi^2_1 + \lambda_2\chi^2_{v-1} < \varepsilon^2)e^{-\varepsilon^2/2}d(\varepsilon^2/2). \tag{3.1}$$

This may be written as

$$\bar{P}_{v,\lambda_1,\lambda_2} = \int_0^\infty \left[\iint_{\lambda_1\chi^2_1+\lambda_2\chi^2_{v-1}<\varepsilon^2} dF(\chi^2_1)\,dF(\chi^2_{v-1})\right]e^{-\varepsilon^2/2}d(\varepsilon^2/2)$$

$$= \int_0^\infty dF(\chi^2_1)\int_0^\infty dF(\chi^2_{v-1})\int_{\sqrt{T}}^\infty e^{-\varepsilon^2/2}d(\varepsilon^2/2)$$

where $T = \lambda_1\chi^2_1 + \lambda_2\chi^2_{v-1}$. This is easily seen to be

$$\int_0^\infty \int_0^\infty e^{-T/2}\, dF(\chi_1^2)\, dF(\chi_{v-1}^2)$$

$$= \frac{1}{2^{v/2}\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{v-1}{2}\right)} \int_0^\infty (\chi_1^2)^{-1/2} e^{-(1+\lambda_1)\chi_1^2/2}\, d\chi_1^2$$

$$\times \int_0^\infty (\chi_{v-1}^2)^{(v-3)/2} e^{-(1+\lambda_2)\chi_{v-1}^2/2} d\chi_{v-1}^2.$$

Evaluating the integrals we get

$$\bar{P}_{v,\lambda_1,\lambda_2} = \{(1+\lambda_1)(1+\lambda_2)^{v-1}\}^{-1/2}. \tag{3.2}$$

In the limiting case where n becomes indefinitely large, $\lambda_1 = \lambda_2 = 0$ and in the limit, $\bar{P} = 1$.

We have seen that

$$(\bar{P})^{-2} = (1+\lambda_1)(1+\lambda_2)^{v-1}. \tag{3.3}$$

We regard $\bar{P}^{-2}$ as a function of p_0 and try to minimize it w.r.t. p_0. For this, we define $\phi(p_0) = \ell n(\bar{P}^{-2})$ and set $\phi'(p_0) = 0$. This gives

$$\frac{2p_0 - 1}{p_0^2(1+\lambda_1)} + \frac{v-1}{1+\lambda_2} = 0$$

On simplification, this gives

$$\Delta\{(2p_0 - 1) + p_0(v-1)\} + (1-p_0)\{(2p_0-1) + p_0^2(v-1)\} = 0,$$

where $\Delta = v/n$.

Take limit when $n \to \infty$ i.e. when $\Delta \to 0$. This gives

$$(1-p_0)\{(2p_0-1) + p_0^2(v-1)\} = 0. \tag{3.4}$$

It can be easily seen that only admissible root of this is $p_0 = (1+\sqrt{v})^{-1}$.

We now consider the case when $\Delta > 0$. We note that λ_2 is small compared to v and so $(1+\lambda_2)^{v-1}$ can be approximed by $1 + (v-1)\lambda_2$. Under this approximation, we have

$$(\bar{P})^{-2} = (\Delta(v-1) + (1-p_0))(\Delta + p_0 - p_0^2)/p_0(1-p_0)^2.$$

We set a value for $\bar{P}$ and search for a value of p_0 which will maximize $\Delta = v/n$ and thereby minimize the value of n required to attain $\bar{P}$.

The table given below gives the optimum values of p_0 and the corresponding values of n when $\bar{P} = .9$ or $.95$. We give these for $v = 2, 3, 5, 10$ and

15. We also give the common value of n_i, the replication number for the control treatments.

$\bar{P} = .9$

v	p_0	n	n_0	n_1
2	0.409	52	22	15
3	0.360	102	36	22
5	0.304	241	71	34
10	0.236	809	189	62
15	0.202	1671	336	89

$\bar{P} = .95$

v	p_0	n	n_0	n_1
2	0.412	110	46	32
3	0.363	214	79	45
5	0.306	503	153	70
10	0.238	1676	396	128
15	0.204	3453	708	183

Relative efficiency of two designs having proportions p_0 and p_0^* for the test treatment may be measured by the ratio of the values n and n^* required to attain the same average coverage probability $\bar{P}$. This would in fact depend upon $\bar{P}$. However, if we use the approximation used above we get this measure of relative efficiency as

$$e_{p_0, p_0^*} = \frac{p_0(1-p_0)(1+(v-1)p_0^*)}{p_0^*(1-p_0^*)(1+(v-1)p_0)}. \tag{3.5}$$

Remark 1: It is interesting to note that the allocations under the DS-optimal designs displayed above in both the Tables do satisfy the approximate relation: $\frac{n_0}{n_1} = \sqrt{v}$. In other words, the DS-optimal designs are approximately A-optimal. [vide Majumdar (1996)].

Remark 2: We can easily incorporate a scaling factor in the p.d.f. for ε. If the p.d.f. for ε is $\varepsilon e^{-\varepsilon^2/2\delta}/\delta$, the value of n is given by $n_\delta = n/\delta$. Thus, small values of δ will lead to larger values of n_δ.

Remark 3: One way to generalize the form of the probability distribution of ε would be to use a mixture distribution for ε^2 i.e., a mixture of scaled χ_2^2 p.d.f.'s given by $\frac{1}{2}\sum_i \frac{w_i}{\delta_i} e^{-\varepsilon^2/2\delta_i}$. The coverage probability would now be given by

$$\bar{P} = \sum_i w_i \delta_i^{v/2} (\delta_i + \lambda_1)^{-1/2} (\delta_i + \lambda_2)^{-(v-1)/2}.$$

For given w_i's and δ_i's an analysis similar to the above can be carried out. This will be somewhat more complex.

Remark 4: If for some design settings $\lambda_1\chi_1^2 + \lambda_2\chi_{v-1}^2$ is replaced by $\lambda_1\chi_{v_1}^2 + \lambda_2\chi_{v_2}^2$, we can use the same technique for computing $\bar{P}$.

4 Control-Treatment comparisons in treatment-connected designs

For a treatment-connected design d, let $\boldsymbol{C}_d$ denote the usual $\boldsymbol{C}$-matrix which is also known as the information matrix for varietal contrasts. We shall formulate the current inference problem in terms of the matrix $\boldsymbol{C}_d$. Clearly,

$$\boldsymbol{\eta} = \boldsymbol{L}\boldsymbol{\tau}, \boldsymbol{L} = (\mathbf{1}, -\boldsymbol{I}), \boldsymbol{\tau} = (\tau_0, \tau_1, \ldots, \tau_v)^t \tag{4.1}$$

so that

$$\hat{\boldsymbol{\eta}} = \boldsymbol{L}\hat{\boldsymbol{\tau}}, D_d(\hat{\boldsymbol{\eta}}) = (\boldsymbol{L}\boldsymbol{C}_d^+\boldsymbol{L}^t)\sigma^2 = \sigma^2\Sigma_d. \tag{4.2}$$

where $\boldsymbol{A}^+$ denotes the Moore-Penrose inverse of the matrix $\boldsymbol{A}$.

Let $\lambda_1, \ldots, \lambda_v$ denote the eigenvalues of $\Sigma_d = \boldsymbol{L}\boldsymbol{C}_d^+\boldsymbol{L}^t$ so that (1.2) can be written as (1.4). We will write $\lambda_{\max}(d)$ for $\lambda_{\max}$ of Σ_d.

We note that

$$\lambda_{\max}(d) \geq \frac{\boldsymbol{x}^t(\boldsymbol{L}\boldsymbol{C}_d^+\boldsymbol{L}^t)\boldsymbol{x}}{\boldsymbol{x}^t\boldsymbol{x}} \quad \text{for any } \boldsymbol{x} \neq \boldsymbol{o}.$$

Taking $\boldsymbol{x} = \mathbf{1}$, we get

$$\begin{aligned}\lambda_{\max}(d) &\geq \frac{\mathbf{1}^t(\boldsymbol{L}\boldsymbol{C}_d^+\boldsymbol{L}^t)\mathbf{1}}{v} \\ &= \frac{(v, -1, -1, \ldots, -1)\boldsymbol{C}_d^+(v, -1, -1, \ldots, -1)^t}{v} \\ &= C_{doo}^+(v+1)^2/v. \end{aligned} \tag{4.3}$$

It is not difficult to verify that

$$C_{doo}^+ C_{doo} \geq (v/(v+1))^2 \tag{4.4}$$

so that

$$\lambda_{\max}(d) \geq \frac{(v+1)^2}{v} \cdot \left(\frac{v}{v+1}\right)^2 \frac{1}{C_{doo}} = \frac{v}{C_{doo}}. \tag{4.5}$$

Let d^* be a design related to d for which $\boldsymbol{C}_{d^*}$ has the structure

$$C_{d^*} = \left[\begin{array}{c|cccc} a & -a/v & \cdots & \cdots & -a/v \\ \hline -a/v & b/v & -c/v & \cdots & -c/v \\ -a/v & -c/v & b/v & \cdots & -c/v \\ \vdots & \vdots & \vdots & & \vdots \\ -a/v & -c/v & -c/v & \cdots & b/v \end{array}\right] \tag{4.6}$$

where $a = C_{doo}$ and $b = tr(\boldsymbol{C}_d) - C_{doo}$.

We claim that d^* improves over d uniformly in $\varepsilon > 0$ in terms of increasing the coverage probability whenever the condition

$$tr(\boldsymbol{C}_d) \geq 2\boldsymbol{C}_{d_{oo}} \tag{4.7}$$

is satisfied. The precise statement and proof are given below.

Theorem 1: *Let d be a design for estimation of* $\boldsymbol{\eta}$ *in (4.1) in a given design setting. Let* d^* *be a design satisfying (4.6). Then, under (4.7)*

$$P_d(\|\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}\| < \varepsilon) \leq P_{d^*}(\|\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}\| < \varepsilon) \tag{4.8}$$

uniformly in $\varepsilon > 0$.

Proof: Observe that

$$\sigma^{-2}\mathrm{D}_{d^*}(\hat{\boldsymbol{\eta}}) = \Sigma_{d^*} = \boldsymbol{L}\boldsymbol{C}_{d^*}^{+}\boldsymbol{L}^t = (\alpha - \beta)\boldsymbol{I}_v + \beta\boldsymbol{J}_{vv} \tag{4.9}$$

where,

$$\alpha = \frac{1}{a} + \frac{(v-1)^2}{vb-a}, \quad \beta = \frac{1}{a} - \frac{(v-1)}{vb-a}. \tag{4.10}$$

Thus, the eigenvalues of Σ_{d^*} are

$$\alpha + (v-1)\beta = v/a \quad \text{with multiplicity } 1$$

$$\text{and} \quad \alpha - \beta = \frac{v(v-1)}{vb-a} \quad \text{with multiplicity } (v-1). \tag{4.11}$$

Whenever (4.7) holds i.e., $b \geq a$, it turns out $\lambda_{\max}(d^*) = v/a$ so that in view of (4.5),

$$\lambda_{\max}(d) \geq \lambda_{\max}(d^*). \tag{4.12}$$

We now argue exactly as in the case of CRD indicated after (2.5).

Case I: $\Pi'\lambda_i(d) \geq \Pi'\lambda_i(d^*) = (\lambda_{\min}(d^*))^{v-1}$. Here Π' (Σ') indicates product (sum) over all eigen values excluding the maximum.

We have

$$P_d[\|\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}\| < \varepsilon] = P\left[\lambda_{\max}(d)Z_1^2 + \Sigma'\lambda_i(d)Z_i^2 \leq \frac{\varepsilon^2}{\delta^2}\right] \tag{4.13}$$

$$\leq P\left[\lambda_{\max}(d^*)Z_1^2 + \Sigma'\lambda_i(d)Z_i^2 \leq \frac{\varepsilon^2}{\delta^2}\right] \tag{4.14}$$

(by implication of events, using (4.12))

$$\leq P\left[\lambda_{\max}(d^*)Z_1^2 + (\Pi'\lambda_i(d))^{1/(v-1)}\chi^2_{v-1} \leq \frac{\varepsilon^2}{\delta^2}\right] \tag{4.15}$$

$$\leq P\left[\lambda_{\max}(d^*)Z_1^2 + (\Pi'\lambda_i(d^*))^{1/(v-1)}\chi^2_{v-1} \leq \frac{\varepsilon^2}{\delta^2}\right] \tag{4.16}$$

(by implication of events)

$$= P_{d^*}[\|\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}\| < \varepsilon] \quad \text{by (4.11).} \tag{4.17}$$

Case II: $\Pi'\lambda_i(d) < \Pi'\lambda_i(d^*) = (\lambda_{\min}(d^*))^{v-1}$.

Here we trace our steps as in Case I. The proof follows from Corollary 1 and Generalization 2 after noting that

(i) $|\boldsymbol{L}\boldsymbol{C}_d^+\boldsymbol{L}'| \geq |\boldsymbol{L}\boldsymbol{C}_{d^*}^+\boldsymbol{L}'|$, (Condition (1.12) in Generalization 2)

and

(ii) $\lambda_{\max}(\Sigma_d) \geq \lambda_{\max}(\Sigma_{d^*}) \geq \lambda_{\min}(\Sigma_{d^*}) \geq (\Pi'\lambda_i(d))^{1/(v-1)}$

hold here. For verification of (i), one may refer to Shah and Sinha (1989, pp 132–134). This completes the proof of the theorem.

If we restrict our attention to the class of designs of the type d^*, the problem reduces to that of choosing $\boldsymbol{C}_{doo}$ so as to maximize

$$P\left[\frac{v}{a}\chi_1^2 + \frac{v(v-1)}{vb-a}\chi^2_{v-1} < \varepsilon^2\right] \tag{4.18}$$

where $a = C_{doo}$ and $tr(\boldsymbol{C}_d) = a + b$, in situations where $b > a$.

As is noted before, the solution is very much ε^2-dependent and, hence we can maximize appropriate weighted average probability by using a suitable weight distribution for ε^2.

We will now specialize to specific design-settings and examine appropriate designs for distance optimality.

BLOCK DESIGNS

We start with the following block design set up. There are B blocks each of size $k(<v)$ and we have $v+1$ treatments of which the control treatment has effect τ_0 and the other v treatments have respective effects $\tau_1, \tau_2, \ldots, \tau_v$. As before the parametric functions of interest are as in $\boldsymbol{\eta}$ given by (1.1) and the optimality criterion to be used is the one given by (1.2).

Let $\mathscr{C}_{R_0,B,v,k}$ denote the class of *binary* connected block designs (d) under consideration for given B, v and k where R_0 is the replication number for the test treatment.

We note that for all designs (d) in $\mathscr{C}_{R_0,B,v,k}$, $C_{doo} = R_0\left(1 - \frac{1}{k}\right)$ and hence

$$\lambda_{\max}(d) \geq \frac{vk}{R_0(k-1)}. \tag{4.19}$$

At this stage we introduce a design $d^* \in \mathscr{C}_{R_0,B,v,k}$ which has the following structure:

$$d^*: \quad \begin{array}{r|c|c|} \hline & 0 & \text{BIBD in} \\ & 0 & \\ B_0 \text{ blocks} & 0 & 1,2,\ldots,v \\ & 0 & \\ & 0 & \\ \hline B_1 \text{ blocks} & \multicolumn{2}{c|}{\text{BIBD in } 1,2,\ldots,v} \\ \hline B_2 \text{ blocks} & \multicolumn{2}{c|}{\text{BIBD in } 0,1,2,\ldots,v} \\ \hline \end{array} \tag{4.20}$$

It is readily seen that C_{d^*} assumes the structure as in (4.6). Let r_0 denote the replication number for a control treatment in the B_0 blocks of the first BIBD and let r_i $(i = 1, 2)$ denote the corresponding replication numbers in the BIBDs comprising of B_i blocks $(i = 1, 2)$. This gives us the relations

$$R_0 = B_0 + r_2, r_0 v = R_0(k-1), r_1 v = B_1 k, r_2(v+1) = B_2 k. \tag{4.21}$$

Further, in (4.6), a and b will have the following expressions:

$$a = R_0(k-1)/k, b = v(r_0 + r_1 + r_2)(k-1)/k$$
$$\text{and} \quad b = a + (v-1)c. \tag{4.22}$$

We note that

$$b = \frac{v(k-1)(r_0 + r_1 + r_2)}{k} > \frac{vr_0 + r_2(k-1)}{k} = a \tag{4.23}$$

so that for $\sum_{d^*}$,

$$\lambda_{\max}(d^*) = \frac{v}{a} = \frac{vk}{R_0(k-1)}. \tag{4.24}$$

Hence, in view of Theorem 1 it is enough to concentrate only on designs of the type d^*.

At this stage, we note that d^* in (4.20) is thus demonstrated to be superior to *any* design d providing the same replication number R_0 for the control treatment. We now go ahead to examine if an optimal choice of R_0 can be made for further improvement.

We write $T = Tr(\boldsymbol{C}_{d^*}) = a + b$ which can be seen to be the same for all d^* within $C_{R_0, B, v, k}$ and for all values of R_0. We also write $\gamma = b/T$. Since $a < b$ (as seen in (4.23)) we have $\frac{1}{2} < \gamma < 1$. Thus, (4.18) reduces to

$$P\left[\frac{v}{T(1-\gamma)}\chi_1^2 + \frac{v(v-1)}{T[(v+1)\gamma - 1]}\chi_{v-1}^2 < \varepsilon^2\right] \tag{4.25}$$

which can also be written as

$$P[\chi_1^2 + p_0\chi_{v-1}^2 < \varepsilon^2 \Delta(p_0)] \tag{4.26}$$

where $p_0 = (v-1)(1-\gamma)/[(v+1)\gamma - 1]$ and $\Delta(p_0) = 1 - \gamma = vp_0/[(v+1)p_0 + v - 1]$. We also note that $\Delta(p_0) = 1 - \gamma = \dfrac{B_0 + r_2}{k(B_0 + B_1 + B_2)} = \dfrac{R_0}{n}$ which is the proportion of experimental units receiving the test treatment. In (4.25) and (4.26), we have used the same ε^2 as a generic notation.

We carried out extensive computations to determine the optimum value of p_0 for a fixed value of ε. Again we find that different values of ε give different values of p_0 and hence a uniformly optimal design does not exist. One can determine the value of p_0 to maximize $\bar{P}$, the weighted coverage probability as in Section 3.

As in the case of CRD we can average over the p.d.f. of ε^2 which is assumed to be that a χ^2 with 2 d.f. to get weighted coverage probability. This gives

$$\bar{P} = \{(1 + \lambda_1^*)(1 + \lambda_2^*)^{v-1}\}^{-1/2}.$$

where $\lambda_1^* = v/T(1-\gamma)$ and $\lambda_2^* = v(v-1)/T[(v+1)\gamma - 1]$. Let $\phi(\gamma) = -2\ell n\bar{P}$. We wish to choose γ to maximize $\bar{P}$ or to minimize $\phi(\gamma)$. Setting $\phi'(\gamma) = 0$ yields

$$A\gamma^2 + B\gamma + C = 0$$

where

$$A = T(v-1)^2(v+1) - T(v+1)^2$$

$$B = -2T(v+1)(v^2 - 2v) - v^2(v^2 - 1)$$

and

$$C = T[(v-1)^2(v+1) - 1] + v^3(v-1).$$

This would give a value of γ which will maximize $\bar{P}$ for fixed values of T and v. We do not find a simple limiting pattern as we did for a CRD. We present here values of γ_{opt} and max $\bar{P}$ for various values of n when $k = 3$ and $v = 5$. It can be seen from these that for fixed values of n, v, and k, the desired values of $\bar{P}$ may not be attainable.

Table 4.1. γ_{opt} and maximal $\bar{P}$ for given n

n	γ_{opt}	max $\bar{P}$
30	.7192	.4236
60	.7036	.6285
75	.7	.6843
150	.6936	.8202
225	.6894	.8744
300	.6879	.9034

As in the case of a CRD, we considered a fixed value of $\bar{P}$ and computed the necessary values of T for many values of γ. Since T relates to the size of the experiment, the best value of γ is the one which gives the smallest value of T. The following table gives the results of some calculations along these lines. We have taken $\bar{P} = .9$ and .95.

Table 4.2. Optimal values of n and R_0
$\bar{P} = 0.90$

k	v	γ	T	n	R_0	B
3	5	0.6881	192	289	90	96
	6	0.7078	271	407	119	136
	7	0.7238	363	544	150	181
	8	0.7372	467	700	184	233
	9	0.7487	582	874	220	291
	10	0.7587	710	1065	257	355
4	7	0.7238	363	484	134	121
	8	0.7372	467	622	164	155
	9	0.7487	582	776	195	194
	10	0.7587	710	947	228	237
5	10	0.7587	710	887	214	177
	11	0.7675	849	1062	247	212
	12	0.7753	1000	1251	281	250

$P = 0.95$

k	v	γ	T	n	R_0	B
	5	0.6858	400	601	188	200
	6	0.7055	564	846	249	282
	7	0.7217	753	1129	314	376
3	8	0.7352	967	1450	384	483
	9	0.7468	1206	1808	458	603
	10	0.7569	1469	2203	536	734
	7	0.7217	753	1004	279	251
4	8	0.7352	967	1289	341	322
	9	0.7468	1206	1607	407	402
	10	0.7569	1469	1958	476	490
	10	0.7569	1469	1836	446	367
5	11	0.7658	1756	2195	514	439
	12	0.7737	2067	2584	585	517

We note that in all cases, the values of n and R_0 are rather large.

Design d^* (with $B_2 = 0$) is identified as a Balanced Treatment Incomplete Block (BTIB) design defined by Bechhofer and Tamhane (1981). See also Hedayat et al. (1988). We note that any design d^* having the form of $\boldsymbol{C}_{d^*}$ as given in (4.6) would have the same optimality properties. Thus, *any* BTIB design would have the same property as d^*. For computations, we would have to take an appropriate parametrization of C_{doo} in terms of the design parameters.

Thus, if there exists a BTIB with approximately the above values of n and R_o for given v and k, it is optimal for the given coverage probability. The optimality is in the class of binary designs with the same values of B, v and k. The design is optimal in the sense of maximum averaged coverage probability.

GENERAL SETTING

The results of the previous sections reveal a pattern which may be useful in a more general setting. Consider a setting in which the $\boldsymbol{C}$-matrix for a design d^* has the form given by (4.6). This will lead to the Σ_{d^*} matrix of the form given in (4.9). This will give the eigenvalues for Σ_{d^*} as given in (4.11). Further, for any competing design d, whenever (4.12) will hold, the analysis of the previous section demonstrating superiority of d^* over the competitor d will remain valid. What remains to be done is to identify a design d^* for the particular setting. For a value of ε which is of interest to us, we try to compute the optimum value of C_{d^*oo} and then try to see if a design d^* having $\boldsymbol{C}_{d^*}$ of the form (4.6) exists. Here, the combinatorial structure of the setting plays a very important role. We have seen how optimal designs may be obtained for zero- and one-way classification designs. For two-way designs the families of designs given in Hedayat et al. (1988) provide A-optimal designs. The class of designs available here is not very rich and hence we do not carry out computations as we did in the other two cases.

Acknowledgement. We sincerely thank two anonymous referees for their critical comments which helped in the revision of an earlier version of the manuscript.

References

Bechofer RE, Tamhane AC (1981) Incomplete block designs for comparing treatments with a control. General Theory. Technometrics 23:45–57 (Corrigendum: 1982 24: pp 71)

Hedayat AS, Jacroux M, Majumdar D (1988) Optimal designs for comparing test treatments with controls. Statistical Science 3:462–491

Liski E, Luoma A, Mandal NK, Sinha BK (1998) Pitman nearness, distance criterion and optimal regression designs. Calcutta Stat. Assoc. Bull. 48:179–184

Liski E, Luoma A, Zaigraev A (1999) Distance optimality design criterion in linear models. Metrika 49:193–211

Majumdar D (1996) Optimal and efficient treatment-control designs. In: S Ghosh and CR Rao, ed. Handbook of Statistics 13: North-Holland, Amsterdam 1007–1054

Marshall AW, Olkin I (1979) Inequalities. Theory of Majorization and its Applications. Academic Press, NY

Okamoto M (1960) An inequality for the weighted Sum of χ^2 variates. Bull. Mathematical Statistics 9:69–70

Sinha BK (1970) On the optimality of some designs. Calcutta Stat. Assoc. Bull. 19:1–22