ELSEVIER

Journal of Statistical Planning and Inference 53 (1996) 237–251

journal of statistical planning and inference

Bayes A-optimal designs for comparing test treatments with a control

Chao-Ping Ting*, Yu-Yun Yang

Graduate Institute of Statistics, National Chengchi University, No. 64, Chi-Nam Road Sec. 2, Wenshan, Taiwan 116, Republic of China

Received 10 February 1994; revised 27 June 1995

Abstract

The problem of comparing v test treatments simultaneously with a control treatment when k, $v \geqslant 3$ is considered. Following the work of Majumdar (1992), we use exact design theory to derive Bayes A-optimal block designs and optimal Γ-minimax designs for a more general prior assumption for the one-way elimination of heterogeneity model. Examples of robust optimal designs, highly efficient designs, and the comparisons of the approximate optimal designs that are derived by our methods and by some other existing rounding-off schemes when using Owen's procedure are also provided.

AMS classifications: Primary 62K05; secondary 62K10

Keywords: A-optimal designs; Bayes experimental designs; Block designs; BTBD; Optimal Γ-minimax designs; Robust designs

1. Introduction

Comparing test treatments with a control treatment (or standard treatment) is a commonly encountered problem in many industrial, agricultural, and biological experiments. Examples of such treatments are different manufacturing processes, different fertilizers, or different prescriptions. Experimenters may wish to compare a set of test treatments, say v test treatments, indexed by $1, 2, \ldots, v$, with a control treatment, indexed by 0, and decide whether the currently used one (control treatment) should be replaced by one of the new ones. In order to reduce the variability from known nuisance sources, block designs become one of the most frequently used designs in practice.

Let τ_i denote the effect of treatment i, $i = 0, 1, \ldots, v$, and $\tau_i - \tau_0$ be the ith test treatment-control contrast that we are interested in. These contrasts are to be

* Corresponding author.

SSDI 0378-3758(95)00146-8

estimated by their best linear unbiased estimators (BLUE) $\hat{\tau}_i - \hat{\tau}_0$. The optimality criterion we employ here is the A-optimality, wherein the optimal design is obtained by minimizing the trace of the covariance matrix of $\hat{\tau}_i - \hat{\tau}_0$ over all designs. There are many well-known results concerning the A-optimality of certain block designs. A general exposition and a brief review can be found in Hedayat et al. (1988).

Since the technology of storaging data is progressing dramatically, one can reasonably assume that the experimenters have gained some knowledge of some of the treatments from past experiences. Designs utilizing prior information might affect number of replications of the $v + 1$ treatments in the experiment. Our interest in this paper is to investigate the effect on designs when a substantial amount of knowledge about the control treatment is introduced into the model. Owen (1970) was the first to address this problem. Following his work, Giovagnoli and Verdinelli (1983, 1985) continued to focus on the field of Bayes optimal designs, and have extended the priors to be three-stage hierarchical.

However, all of the above-mentioned papers are in the setup of approximate design theory, i.e. the n_{ij}, the number of times treatment i occurs in block j, are presupposed to be real numbers. The benefit of this is that powerful theorems applicable to a large class of prior distributions can be derived. Nevertheless, it is usually necessary to use some 'round off' strategies in order to implement the designs. Not only do we need to determine which rounding strategy is best, there may be no rounding strategy that leads to an exact design that is optimal.

Based on practical situation, we consider the problem in the exact design theory setup, i.e. the n_{ij} must be integers. Majumdar (1988) and Stufken (1991) are in this spirit. Majumdar (1992) investigated Bayes A-optimal block designs and optimal Γ-minimax block designs for a large class of prior distributions when $k \leqslant v$. In this paper, we extend Majumdar's (1992) consideration to a more general situation, and focus on finding optimal designs for all $k, v \geqslant 3$. The results are applicable to cases when $k > v$ as well as $k \leqslant v$. Moreover, Owen's (1970) results are confirmed by our procedure when the n_{ij} obtained by his method are integers. An illustrative example is provided.

Robustness of Bayes experimental designs is very desirable. We investigate the robustness of Bayes optimal designs for a slightly more general class of prior distributions than Majumdar (1992). It is not surprising to learn that Bayes optimal designs are very robust against departures from the given prior distribution. In the latter part of this paper, we also observe that Balanced Treatment Block designs are quite robust.

Assumptions, definitions, and notation are presented in Section 2. Section 3 contains the identification of Bayes A-optimal and optimal Γ-minimax designs. Robustness and other properties of these designs are given in Section 4. Examples of robust optimal designs, highly efficient designs, and the comparisons of the approximate optimal designs derived by our methods with those derived by some other existing schemes are given in Section 4 as well.

2. Preliminaries

Let $D(v+1, b, k)$ denote the set of all possible block designs with $v+1$ treatments arranged in b blocks of size k each, $k, v \geqslant 3$. For a design $d \in D(v+1, b, k)$, the model we use is the usual additive linear model without interactions,

$$Y = X_{1d}\theta + X_2\xi + \varepsilon,$$

where Y is the vector of observations which are arranged block by block, $\theta = ((\tau_i - \tau_0))$ is the $v \times 1$ vector of the v test treatment-control contrasts. $\xi = ((\mu + \tau_0 + \beta_j))$ is the $b \times 1$ vector which indicate the performances of the control treatment in b blocks, τ_i is the effect of treatment i, $i = 0, \ldots, v$, β_j is the effect of block j, $j = 1, \ldots, b$ and ε is the vector of random error. Then $X_{1d} = ((\delta_{hi}))$, $h = 1, \ldots, bk$, and $i = 1, \ldots, v$, where $\delta_{hi} = 1$ if the hth observation receives treatment i, and 0 otherwise. And $X_2 = I_b \otimes \mathbf{1}_k$, where $\mathbf{1}_k$ is a $k \times 1$ vector of 1's and I_b is the $b \times b$ identity matrix.

Suppose the distribution of errors and the prior distributions are multivariate normal:

$$Y \mid \theta, \xi \sim \mathrm{N}(X_{1d}\theta + X_2\xi, \Sigma),$$

$$\begin{pmatrix}\theta \\ \xi\end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix}\mu_\theta \\ \mu_\xi\end{pmatrix}, \begin{pmatrix}T & 0 \\ 0 & B\end{pmatrix}\right).$$

Then the posterior distribution of θ is

$$\theta \mid y, T, B \sim \mathrm{N}(\mu_d, C_d^{-1}),$$

where

$$C_d = X'_{1d}(\Sigma + X_2BX'_2)^{-1}X_{1d} + T^{-1},$$

and

$$C_d\mu_d = X'_{1d}(\Sigma + X_2BX'_2)^{-1}(y - X_2\mu_\xi) + T^{-1}\mu_\theta.$$

Under squared error loss $L(\hat{\theta}, \theta) = (\hat{\theta} - \theta)'(\hat{\theta} - \theta)$, the Bayes estimator $\hat{\theta}$ of θ is the posterior mean μ_d, and the posterior expected loss is the trace of $C_d^{-1}(\mathrm{tr}(C_d^{-1}))$. A Bayes A-optimal design d^* is then defined as one that satisfies

$$\mathrm{tr}(C_{d^*}^{-1}) = \min_{d \in D(v+1, b, k)} \mathrm{tr}(C_d^{-1}).$$

Next we confine the prior covariance matrices Σ, T and B to the following:

$$\Sigma = \{I_b \otimes ((1-\gamma_1)I_k + (\gamma_1 - \gamma_2)J_k) + \gamma_2 J_{bk}\}\sigma^2, \tag{2.1}$$

$$T = ((\upsilon_1 - \upsilon_2)I_v + \upsilon_2 J_v)\sigma^2, \qquad B = ((a - a\rho)I_b + a\rho J_b)\sigma^2, \tag{2.2}$$

where J_n is an $n \times n$ matrix of 1's, and $-1 < \gamma_1, \gamma_2, \rho < 1$, $a > 0$, and $\upsilon_1 > |\upsilon_2|$.

Let $N_d = ((n_{dij})$ denote the $v \times b$ test treatment vs block incidence matrix. Then by Majumdar (1992) and under regularity conditions,

$$a(1-\rho) + \gamma_1 - \gamma_2 \neq 0,$$

$$1 - \gamma_1 + k(\gamma_1 - \gamma_2) + ka(1-\rho) \neq 0,$$

$$1 - \gamma_1 + k(\gamma_1 - \gamma_2) + ka(1-\rho) + bk(\gamma_2 + a\rho) \neq 0,$$

C_d can be written as follows:

$$\sigma^2(1-\gamma_1)C_d = \mathrm{diag}(r_{d1}, \ldots, r_{dv}) - (k+p)^{-1} N_d N_d' - q\boldsymbol{r}_d\boldsymbol{r}_d' + \sigma^2(1-\gamma_1)T^{-1}, \tag{2.3}$$

where

$$p = (1-\gamma_1)(a(1-\rho) + (\gamma_1 - \gamma_2))^{-1},$$

$$q = (1-\gamma_1)(\gamma_2 + a\rho)(1 - \gamma_1 + k(\gamma_1 - \gamma_2) + ka(1-\rho))^{-1} \times (1 - \gamma_1 + k(\gamma_1 - \gamma_2) + ka(1-\rho) + bk(\gamma_2 + a\rho))^{-1},$$

$$r_{di} = \sum_{j=1}^{b} n_{dij}, \text{ the replication of treatment } i, \quad \forall i,$$

$$\boldsymbol{r}_d = (r_{d1}, \ldots, r_{dv})'.$$

In this paper, we focus on finding optimal designs whose prior covariances are as in (2.1) and (2.2), and whose posterior covariance C_d^{-1} is determined by (2.3). At the end of this section, two needed definitions from Jacroux and Majumdar (1989) are stated.

Definition 1. A design d is a Balanced Treatment Block design (BTBD) if

$$\lambda_{d01} = \cdots = \lambda_{d0v} = \lambda_0,$$

$$\lambda_{d12} = \cdots = \lambda_{dv-1,v} = \lambda_1,$$

where $\lambda_{dii'} = \sum_{j=1}^{b} n_{dij}n_{di'j}$, for $i \neq i'$, $i, i' = 0, \ldots, v$.

A special group of BTBDs whose optimality property, when their is no prior information available, has been studied by many researchers is defined in the following.

Definition 2. A BTBD$(v, b, k; t, s)$ is a BTBD in $D(v+1, b, k)$ with the additional properties that

$$n_{d01} = \cdots = n_{d0s} = t + 1, \qquad n_{d0,s+1} = \cdots = n_{d0b} = t,$$

$$|n_{dij} - n_{di'j'}| \leqslant 1, \quad \text{for } (i,j) \neq (i',j'), \ i, i' = 1, \ldots, v, \quad \text{and} \quad j, j' = 1, \ldots, b.$$

3. Optimal designs

Theorem 1. *Suppose* (2.1) *and* (2.2) *hold, and* $1 > \gamma_1 \geqslant \gamma_2, \gamma_2 + a\rho \geqslant 0$, *and* $\upsilon_2 \geqslant 0$. *Let* $\eta_1 = (1-\gamma_1)(\upsilon_1 - \upsilon_2)^{-1}$, $\eta_2 = (1-\gamma_1)(\upsilon_1 + (v-1)\upsilon_2)^{-1}$, $U = by - z$, $V = by^2 - 2yz + z$, $h(y,z) = -bv[U/bv]^2 + (2U - bv)[U/bv] + U$, *where* y *and* z *are two nonnegative integers, and*

$$g(y,z) = \frac{v(v-1)^2(k+p)}{(v-1)(k+p)U - vh(y,z) + V + v(v-1)(k+p)\eta_1} + \frac{v(k+p)}{(k+p)U - q(k+p)U^2 - V + v(k+p)\eta_2},$$

$$\Delta = \{(y,z)\colon y = [(k+p+1)/(2(bq(k+p)+1))] + 1, \ldots, k;\ z = 0, 1, \ldots, b\},$$

where $[\cdot]$ *is the greatest integer function. Then*

(i) *if* $(k+p+1)/(2(bq(k+p)+1)) < k$, *and* $g(y^*, z^*) = \min\{g(y,z)\colon (y,z) \in \Delta\}$, *then a* $BTBD(v, b, k; k - y^*, z^*\}$, *if it exists, is Bayes A-optimal in* $D(v+1, b, k)$.

(ii) *if* $(k+p+1)/(2(bq(k+p)+1)) \geqslant k$, *then a* $BTBD(v, b, k; 0, 0)$, *if it exists, is Bayes A-optimal in* $D(v+1, b, k)$.

Proof. The proof of this theorem is long and we use some of the techniques in Majumdar (1992), hence is put in the Appendix.

Remark. When $k \leqslant v$, then $[U/bv] = 0$, Theorem 1 reduces to Theorem 3.1 of Majumdar (1992).

As to the optimal Γ-minimax designs, suppose the range of priors is

$$\Gamma = \{T, B\colon \upsilon_m I_v \sigma^2 \leqslant T \leqslant \upsilon_M I_v \sigma^2,\ a_m I_b \sigma^2 \leqslant B \leqslant a_M I_b \sigma^2\}, \tag{3.1}$$

where '$\leqslant$' means nonnegative definite ordering of matrices. For a design d, the risk of the Γ-minimax rule is given by $\mathrm{tr}(C_d^{-1}(\Gamma))$, where

$$C_d(\Gamma) = X'_{1d}(\Sigma + \sigma^2 a_M X_2 X'_2)^{-1} X_{1d} + \sigma^{-2}\upsilon_M^{-1} I_v.$$

An optimal Γ-minimax design (Majumdar, 1992) d^* is defined by minimizing the risk, i.e.

$$\mathrm{tr}(C_{d^*}^{-1}(\Gamma)) = \min_{d \in D(v+1,b,k)} \mathrm{tr}(C_d^{-1}(\Gamma)).$$

To find optimal Γ-minimax designs, a_m and υ_m can be taken to be zeroes without loss of generality. Under (2.1), (3.1), and $1 > \gamma_1 \geqslant \gamma_2 \geqslant 0$, optimal Γ-minimax designs can be obtained by applying Theorem 1 with $\rho = \upsilon_2 = 0$, $a = a_M$, $\upsilon_1 = \upsilon_M$.

For simplicity of the subsequent discussion of design robustness, we consider priors with $\upsilon_1 \to \infty$. This may arise when v test treatments are new and unanalyzed so that the prior information on the performances of the new treatments relative to the control treatment is vague. Additionally, the errors are assumed to be uncorrelated between blocks and positively correlated within blocks, i.e. $0 \leqslant \gamma_1 < 1, \gamma_2 = 0$. This is more general than Majumdar (1992) where the errors are uncorrelated between as well as within blocks. Under this specialization and with a slightly different notation, Corollary 1 is derived.

Corollary 1. *Suppose* $\Sigma = \sigma^2 I_b \otimes ((1-\gamma_1)I_k + \gamma_1 J_k)$. *Let*

$$g_0(y,z;\pi) = (v(v-1)^2(k+\pi^{-1}))((v-1)(k+\pi^{-1})U - vh(y,z) + V)^{-1}$$
$$+ (v(k+\pi^{-1}))((k+\pi^{-1})U - V)^{-1},$$

and

$$\Delta_0(\pi) = \{(y,z)\colon y = [(k+\pi^{-1}+1)/2]+1, \ldots, k;\ z = 0, \ldots, b\},$$

where $\pi = (a+\gamma_1)/(1-\gamma_1)$. *Let* $d^* \in D(v+1,b,k)$ *denote a Bayes A-optimal design for* $T = \upsilon_1 I_v \sigma^2$, $\upsilon_1 \to \infty$, $B = aI_b\sigma^2$, *and optimal Γ-minimax design according to the prior covariance class* $\{T,B\colon T = \upsilon_1 I_v\sigma^2,\ \upsilon_1 \to \infty,\ O_b \leqslant B \leqslant aI_b\sigma^2\}$, *where* O_b *is a* $b \times b$ *matrix of zeroes. Then*

(i) *if* $\pi > 1/(k-1)$, *and* $g_0(y_\pi^*, z_\pi^*; \pi) = \min\{g_0(y,z;\pi)\colon (y,z) \in \Delta_0(\pi)\}$, *then we can take* d^* *as a* $BTBD(v,b,k;k-y_\pi^*,z_\pi^*)$, *if it exists.*

(ii) *if* $\pi \leqslant 1/(k-1)$, *then we can take* d^* *as a* $BTBD(v,b,k;0,0)$, *if it exists.*

Next, let us examine the behavior of $g_0(y,z;\pi)$ by the following lemma from Jacroux and Majumdar (1989) with a slight modification.

Lemma 2. *Given values of* v,b,k *and* π. *Suppose* $\pi > 1/(k-1)$ *and* $y \in \{[(k+\pi^{-1}+1)/2] + 1, \ldots, k\}$.

(i) *For every value of* y, *there exists a* $z_0 \in [0,b]$ *which is a function of* y, *such that* $g_0(y,z;\pi)$ *is* (a) *decreasing in* z *for* $z \in [0,z_0]$ *and increasing in* z *for* $z \in (z_0,b]$, *or* (b) *increasing in* z *for* $z \in [0,b]$ *when* $z_0 = 0$, *or* (c) *decreasing in* z *for* $z \in [0,b]$ *when* $z_0 = b$;

(ii) (a) $g_0(y,b-1;\pi) \leqslant g_0(y,b;\pi)$ *implies* $g_0(y-1,0;\pi) \leqslant g_0(y-1,1;\pi)$ for $y \in \{[(k+\pi^{-1}+1)/2]+2, \ldots, k\}$. (b) $g_0(y,1;\pi) \leqslant g_0(y,0;\pi)$ *implies* $g_0(y+1,b;\pi) \leqslant g_0(y+1,b-1;\pi)$ for $y \in \{[(k+\pi^{-1}+1)/2]+1, \ldots, k-1\}$.

Lemma 2 guarantees the existence of y_π^* and z_π^* for each π, and $g_0(y,z;\pi)$ is minimized either by a unique pair (y_π^*, z_π^*) or by two pairs (y_π^*, z_π^*) or (y_π^*, z_π^*+1). Since $r_{d0} = b(k-y) + z$, we define the 'optimal' replication of the control treatment for each π as $r_0^*(\pi)$, where

$$r_0^*(\pi) = \begin{cases} b(k-y_\pi^*) + z_\pi^* & \text{if } g_0(y,z;\pi) \text{ is uniquely minimized at } y_\pi^*, z_\pi^*, \\ b(k-y_\pi^*) + z_\pi^* + 1 & \text{if } g_0(y_\pi^*, z_\pi^*;\pi) = g_0(y_\pi^*, z_\pi^*+1;\pi) = \min\limits_{y,\,z\in\Delta_0} g_0(y,z;\pi). \end{cases}$$

Table 1

π	$r_0^*(\pi)$	π	$r_0^*(\pi)$	π	$r_0^*(\pi)$
1/8.5	1	1/5.5	+ 8	1/2.5	17
1/8	3	1/5	11	1/2	19
1/7.5	+ 4	1/4.5	+ 12	1/1.5	+ 20
1/7	5	1/4	+ 12	1/1	21
1/6.5	+ 8	1/3.5	15	1/0.5	+ 24
1/6	+ 8	1/3	16		

Note: '+' denotes that a BTBD(4, 4, 18; t, s) with $r_0^*(\pi)$ replications of the control exists.

For Σ, T, B as sated in Corollary 1, using Owen's (1970) procedure one can obtain that for $\pi \geqslant \sqrt{v}/k$ the best $n_{dij} = (k + \pi^{-1})/(v + \sqrt{v})$, $n_{d0j} = k - vn_{dij}$, $i = 1, \ldots, v$, $j = 1, \ldots, b$. For cases when these n_{dij} and n_{d0j} are integers, Owen's result is indeed confirmed by the above Corollary 1. The following is an example.

Example 3.1. For $v = 4$, $b = 4$, $k = 18$, there is only one value of $\pi \geqslant \sqrt{v}/k$ such that the n_{dij} and n_{d0j} obtained by Owen's procedure are integers, i.e. when $\pi = \frac{1}{6}$, the best n_{dij} is 4, and n_{d0j} is 2, hence $r_{d0} = 8$, $r_{di} = 16$, $i = 1, \ldots, 4$. We examine 17 different values of π, and the corresponding $r_0^*(\pi)$ values obtained by Corollary 1 are listed in Table 1.

One can see that when $\pi = 1/6$, $r_0^*(1/6) = 8$ which is the same as Owen's result.

4. Robustness and approximation

In Example 3.1, one can see that for many different values of π the optimal r_{d0} values are the same. For example, when $\pi = \frac{1}{6.5}, \frac{1}{6}$, and $\frac{1}{5.5}$, the optimal r_{d0} values are all 8. Thus, if one can show that when $\frac{1}{6.5} \leqslant \pi \leqslant \frac{1}{5.5}$, $r_0^*(\pi) = 8$, then BTBD(4, 4, 18; 2, 0) is robust optimal for all π between $\frac{1}{6.5}$ and $\frac{1}{5.5}$.

This section is proceeded under the situation of Corollary 1 where a and γ_1 are considered through $\pi = (a + \gamma_1)/(1 - \gamma_1)$.

Theorem 3. *Suppose $\pi_1 < \pi_2$ are two nonnegative real numbers. Then $r_0^*(\pi_1) \leqslant r_0^*(\pi_2)$.*

Proof. See the Appendix.

For $\alpha = (k + \pi^{-1})^{-1}$, let $q_0(y, z; \pi) = q(y, z; \alpha)$, where $q(y, z; \alpha) \propto g_0(y, z + 1; \pi) - g_0(y, z; \pi)$ is as defined in the Appendix. One can see that if $q_0(y, z;\ \infty) \leqslant 0$, there exists exactly one π such that $q_0(y, z; \pi) = 0$, i.e. both $r_0^*(\pi)$ and $r_0^*(\pi) + 1$ minimize g_0. For the other π's, there is only one $r_0^*(\pi)$ minimize g_0. Since Theorem 3 is equivalent to either one of the following two statements, i.e. if $q_0(y, z; \pi_2) > 0$, then $q_0(y, z; \pi_1) > 0$, or if $q_0(y,z;\pi_1) < 0$, then $q_0(y,z;\pi_2) < 0$, hence the following statements can be written.

$$g_0(y,z;\pi_1) \leqslant g_0(y,z+1;\pi_1) \text{ implies } g_0(y,z;\pi) < g_0(y,z+1;\pi) \quad \text{for all } \pi < \pi_1, \tag{4.1}$$

$$g_0(y,z;\pi_2) \geqslant g_0(y,z+1;\pi_2) \text{ implies } g_0(y,z;\pi) > g_0(y,z+1;\pi) \quad \text{for all } \pi > \pi_2. \tag{4.2}$$

Therefore, if $\pi_1 < \pi_2$, and $r_0^*(\pi_1) = r_0^*(\pi_2) = r_0^*$, then by (4.1), (4.2) and Theorem 3, $r_0^*(\pi) = r_0^*$ for all $\pi \in [\pi_1, \pi_2]$.

Let $r_0^*(\infty) = b(k - y_\infty^*) + z_\infty^*$ denote the value of $r_0^*(\pi)$ when $\pi \to \infty$. Define π_e be such that $q_0(y_\infty^*, z_\infty^* - 1; \pi_e) = 0$. Then for $\pi \in [\pi_e, \infty)$, $r_0^*(\pi) = r_0^*(\pi_e) = r_0^*(\infty)$, and for $\pi \in [0, \pi_e)$, $r_0^*(\pi) < r_0^*(\infty)$.

Let $\tilde{\pi}_i$, $i = 1, \ldots, r_0^*(\pi_e) - 1$ be such that $q_0(\tilde{y}_i^*, \tilde{z}_i^*; \tilde{\pi}_i) = 0$, $r_0^*(\tilde{\pi}_i) = i$. (The existence of $\tilde{\pi}_i$'s can be verified by Majumdar (1992), Corollary 4.3.) Define $\tilde{\pi}_0 = 0$, $\tilde{\pi}_{r_0^*(\pi_e)} = \pi_e$. It can easily be seen that $\tilde{\pi}_0 < \tilde{\pi}_1 < \cdots < \tilde{\pi}_{r_0^*(\pi_e)} = \pi_e$. Let $S_i = [\tilde{\pi}_i, \tilde{\pi}_{i+1})$, $i = 0, \ldots, r_0^*(\pi_e) - 1$, and $S_{r_0^*(\pi_e)} = [\pi_e, \infty)$.

Theorem 4. *Given vlaues of v, b, and k, there exist intervals* $S_0, S_1, \ldots, S_{r_0^*(\pi_e)}$ *such that* $\bigcup_{i=0,\ldots,r_0^*(\pi_e)} S_i = [0, \infty)$, *and a* $BTBD(v,b,k;t,s)$ *with* $bt + s = i$, *if exists, is Bayes A-optimal and optimal* Γ*-minimax for each* $\pi \in S_i$, $i = 0, 1, \ldots, r_0^*(\pi_e)$.

Proof. The proof is a direct consequence of all the above.

Therefore, for the inverval S_i where BTBD$(v,b,k;t,s)$ with $bt + s = i$ exists, it is robust optimal over all $\pi \in S_i$. As to the remaining intervals, it is unavoidable to look for an approximately optimal designs whose $\mathrm{tr}(C_d^{-1})$ are close enough to $g_0(y_\pi^*, z_\pi^*; \pi)$. Designs that are derived by the following two methods might be highly efficient.

Method 1. Construct a design which is or is 'closest' to a BTBD with r_{d0} close to i for each $\pi \in S_i$ such that $(bk - r_{d0})/v$ is an integer. That is, let $\bar{i}$ *be the smallest* i, and $\hat{i}$ be the largest i such that $\hat{i} < i < \bar{i}$ and $(bk - \bar{i})/v$, $(bk - \hat{i})/v$ both are integers. Construct designs which are or are 'closest' to BTBDs with $r_{d0} = \bar{i}$ and $\hat{i}$. Choose the one with the smaller $\mathrm{tr}(C_d^{-1})$.

Method 2. Construct a design which is 'closest' to a BTBD with $r_{d0} = i$ for each $\pi \in S_i$.

In the application of Owen's procedure, there are many systematic schemes to round off the n_{dij} and n_{d0j} and to get exact approximate optimal designs. The following two approximation schemes are considered.

Scheme 1. Round off n_{dij} to the nearest or second nearest integer such that $\sum_{i=1}^{v} n_{dij} \leqslant k$, $\forall j$, and construct a BTBD.

Scheme 2. Round off n_{d0j} to the nearest integer, and construct a design with $|n_{dij} - n_{di'j'}| \leqslant 1$, for $(i,j) \neq (i',j')$, $i, i' = 1, \ldots, v$, and $j, j' = 1, \ldots, b$, which is or is closest to a BTBD.

From the many examples that we have examined, we find out that when the values of v, b, and k are close, Methods 1 and 2 usually do better than Schemes 1 and 2. For the cases when b is large compared to v and k, Method 1 usually outperforms the others, and when k is large compared to b and v, Method 1 and Scheme 2 do better. Two illustrative examples are given in the following.

Example 4.1. For $v = 4$, $b = 5$, and $k = 6$. We use the above two methods and two schemes to approximate the optimal designs for those S_i whose BTBD$(v, b, k; t, s)$'s do not exist. Let N_i denote the incidence matrix of the $v + 1$ treatments of the approximately optimal designs that are constructed by the above methods and schemes with $r_{d0} = i$.

$$N_0 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 2 & 2 & 2 \\ 2 & 2 & 2 & 1 & 1 \\ 2 & 1 & 1 & 2 & 1 \\ 1 & 2 & 1 & 1 & 2 \end{bmatrix}, \qquad N_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 2 & 2 & 2 \\ 1 & 2 & 2 & 1 & 1 \\ 2 & 1 & 1 & 2 & 1 \\ 1 & 2 & 1 & 1 & 2 \end{bmatrix},$$

$$N_2 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 2 & 2 & 1 \\ 1 & 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 1 & 2 \\ 2 & 2 & 1 & 1 & 1 \end{bmatrix}, \qquad N_3 = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 2 & 2 & 1 \\ 1 & 1 & 1 & 2 & 2 \\ 1 & 2 & 1 & 1 & 2 \\ 2 & 1 & 1 & 1 & 1 \end{bmatrix},$$

$$N_4 = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 1 & 2 \\ 1 & 2 & 1 & 1 & 1 \\ 2 & 1 & 1 & 1 & 1 \end{bmatrix}, \qquad N_5 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 & 1 \\ 1 & 1 & 1 & 2 & 2 \end{bmatrix},$$

$$N_7 = \begin{bmatrix} 2 & 2 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 & 1 \\ 1 & 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 1 & 2 \end{bmatrix}, \qquad N_8 = \begin{bmatrix} 2 & 2 & 2 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 1 & 2 \end{bmatrix},$$

$$N_9 = \begin{bmatrix} 2 & 2 & 2 & 2 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 2 \end{bmatrix}$$

In Table 2 of efficiencies for different points of π for the above two methods and schemes are given, where efficiency is measured by $(1 - \gamma_1)^{-1}\operatorname{tr}(C_d^{-1})/\min g_0$.

Table 2

π	$r_0^*(\pi)$	$\min g_0$	Method 1	Method 2	Scheme 1	Scheme 2
0.395	1	0.859853	0.997798(2)	**0.998087**(1)	0.880513(10)	0.997338(0)
0.415	2	0.873992	**0.999465**(2)	**0.999465**(2)	0.890537(10)	0.995091(0)
0.440	3	0.889347	**0.999705**(2)	0.997686(3)	0.900847(10)	0.989830(0)
0.475	4	0.907325	0.996271(2)	**0.997059**(4)	0.912002(10)	0.994865(5)
0.515	5	0.924347	0.989368(2)	**0.997600**(5)	0.921623(10)	**0.997600**(5)
0.885	+ 6	1.029393	**1.000000**(6)	**1.000000**(6)	0.975177(10)	0.997863(5)
0.985	7	1.047921	**0.999891**(6)	0.997442(7)	0.983668(10)	0.994988(5)
1.120	8	1.066726	**0.996760**(6)	0.996684(8)	0.990881(10)	0.988526(5)
1.320	9	1.086366	0.996565(10)	**0.997572**(9)	0.996565(10)	0.977019(5)
1.655	10	1.107193	**1.000000**(10)	**1.000000**(10)	**1.000000**(10)	**1.000000**(10)
∞	+ 10					

Note: (i) '+' denotes that a BTBD$(4, 5, 6; t, s)$ with $r_0^*(\pi)$ replications of the control exists.

(ii) The integers in the parentheses are the r_{d0} values derived by the corresponding methods or schemes to construct the approximate optimal designs.

(iii) Bold-faced values denote the highest efficiency among the four.

(iv) $\pi \to \infty$ indicates the situation where there is no prior information.

Table 3

π	$r_0^*(\pi)$	$\min g_0$	Method 1	Method 2	Scheme 1	Scheme 2
0.535	1	0.326769	**0.999871**(0)	0.999627(1)	0.969256(12)	**0.999871**(0)
0.545	2	0.329109	**0.999636**(3)	0.999632(2)	0.973154(12)	0.999283(0)
0.555	+ 3	0.331313	**1.000000**(3)	**1.000000**(3)	0.976668(12)	0.998295(0)
0.570	4	0.334361	**0.999812**(3)	0.999614(4)	0.981210(12)	0.996080(0)
0.580	5	0.336269	0.999477(6)	**0.999617**(5)	0.983899(12)	0.994267(0)
0.595	+ 6	0.338916	**1.000000**(6)	**1.000000**(6)	0.987345(12)	0.990971(0)
0.615	7	0.342109	**0.999822**(6)	0.999600(7)	0.991043(12)	0.985726(0)
0.630	8	0.344316	0.999373(9)	**0.999608**(8)	0.993336(12)	0.981363(0)
0.650	+ 9	0.346994	**1.000000**(9)	**1.000000**(9)	0.995721(12)	0.974951(0)
0.675	10	0.349983	**0.999931**(9)	0.999589(10)	0.997825(12)	0.997825(12)
0.700	11	0.352657	0.999195(12)	**0.999592**(11)	0.999195(12)	0.999195(12)
0.735	+ 12	0.355908	**1.000000**(12)	**1.000000**(12)	**1.000000**(12)	**1.000000**(12)
6.395	13	0.466192	**0.999998**(12)	0.998238(13)	**0.999998**(12)	**0.999998**(12)
8.935	14	0.474108	**0.998467**(12)	0.998278(14)	**0.998467**(12)	**0.998467**(12)
15.730	+ 15	0.482605	**1.000000**(15)	**1.000000**(15)	0.994945(12)	0.994945(12)
92.985	16	0.491172	**0.999999**(15)	0.997997(16)	0.988804(12)	0.988804(12)
∞	16					

Example 4.2. For $v = 3$, $b = 12$, and $k = 4$. We use the above two methods and two schemes to approximate the optimal designs for those S_i whose BTBD$(v, b, k; t, s)$'s do not exists. Since the incidence matrices N_i are constructed the same way as in the above example, hence are omitted. Efficiencies are given in Table 3.

Remark. In the above two examples, the best n_{dij}'s of the corresponding intervals of π are all real numbers, the optimality property thus cannot be proved by Owen's method, but can be proved so by our Corollary 1.

5. Conclusion

Even though the efficiencies in the columns titled Method 1 and Method 2 in the above two tables are higher in most cases, but these two methods do not lead to much of an improvement over Scheme 2. Due to the relative simplicity of Scheme 2, and from the practitioners' point of view, i.e. without running a computer program to search for the best $r_0^*(\pi)$, we thus suggest to use Scheme 2 to approximate the optimal designs.

Appendix

Proof of Theorem 1. By permuting and averaging over the v test treatments in C_d^{-1}, one can obtain

$$\sigma^{-2}(1-\gamma_1)^{-1}\operatorname{tr}(C_d^{-1}) \geqslant v(v-1)^2(k+p)(\phi+v(v-1)(k+p)\eta_1)^{-1} + v(k+p)(\psi+v(k+p)\eta_2)^{-1},$$

where

$$\phi=(v-1)(k+p)U_d - v\sum_{j=1}^{b}\sum_{i=1}^{v} n_{dij}^2 + V_d - vq(k+p)\sum_{i=1}^{v} r_{di}^2 + q(k+p)U_d^2,$$

$$\psi=(k+p)U_d - V_d - q(k+p)U_d^2, \qquad U_d=\sum_{i=1}^{b} r_{di}, \qquad V_d=\sum_{j=1}^{v}\left(\sum_{i=1}^{v} n_{dij}\right)^2.$$

Now since

$$\min\sum_{i=1}^{v} r_{di}^2 = U_d^2/v, \tag{A.1}$$

and the minimum value occurs when $r_{di}=U_d/v$, $i=1,\dots,v$. And

$$\min\sum_{j=1}^{b}\sum_{i=1}^{v} n_{dij}^2 = -bvm^2+(2U_d-bv)m+U_d, \tag{A.2}$$

where $m=[U_d/bv]$, and the minimum value occurs when the n_{dij} are as equal as possible, i.e. $n_{dij}=m$ or $m+1$. Then by (A.1), (A.2) and the fact that $p\geqslant 0$ and $q\geqslant 0$ (since $1>\gamma_1\geqslant\gamma_2$, $\gamma_2+ap\geqslant 0$), one can obtain

$$\sigma^{-2}(1-\gamma_1)^{-1}\operatorname{tr}(C_d^{-1}) \geqslant \bar{g}(U_d, V_d),$$

where

$$\bar{g}(U_d, V_d) = v(v-1)^2(k+p)(\bar{\phi} + v(v-1)(k+p)\eta_1)^{-1} + v(k+p)(\psi + v(k+p)\eta_2)^{-1},$$

$$\bar{\phi} = (v-1)(k+p)U_d - v(-bvm^2 + (2U_d - bv)m + U_d) + V_d.$$

Let us further examine the relationship between V_d and $\bar{g}(U_d, V_d)$ for fixed value of U_d by taking the derivative of $\bar{g}$ with respect to V_d. One then has

$$(\partial/\partial V_d)\bar{g}(U_d, V_d) = v(k+p)\{(\bar{\phi} + v(v-1)(k+p)\eta_1)(\psi + v(k+p)\eta_2)\}^{-1} \times \{(\bar{\phi} + v(v-1)(k+p)\eta_1)^2 - ((v-1)\psi + v(v-1)(k+p)\eta_2)^2\} > 0,$$

since $\bar{\phi} > (v-1)\psi$, and $\eta_1 \geqslant \eta_2$ (since $\upsilon_2 \geqslant 0$).

Therefore, for fixed value of U_d, $\bar{g}(U_d, V_d)$ achieves its minimum value when V_d is small or when $|\sum_{i=1}^{v} n_{dij} - \sum_{i=1}^{v} n_{dij'}| \leqslant 1$, $\forall j \neq j'$. Let $y = [U_d/b] + 1$, $z = by - U_d$. We then have

$$\bar{g}(U_d, V_d) \geqslant g(y, z).$$

Note that $\sigma^{-2}(1-\gamma_1)^{-1}\mathrm{tr}(C_d^{-1}) = g(y,z)$ if d is a BTBD$(v,b,k;k-y,z)$.

The rest of the proof is devoted to the examination of the behaviour of g in terms of y and z, respectively. Let

$$w_1(y,z) = (v-1)(k+p)U - vh(y,z) + V,$$

$$w_2(y,z) = (k+p)U - q(k+p)U^2 - V.$$

Through some straightforward but tedious calculation, we derive the following:

(a) $w_1(y,z)$ is decreasing in z for fixed value of y, and is increasing in y for fixed value of z,

(b) $(\partial/\partial z)w_2(y,z) \leqslant 0$ if and only if $y \leqslant y_u$, where

$$y_u = (\tfrac{1}{2}(k+p+1) + qz(k+p))(bq(k+p)+1)^{-1}.$$

And for fixed value of z,

$$(\partial^2/\partial y^2)w_2(y,z) = -2qb^2(k+p) - 2b < 0,$$

hence $w_2(y,z)$ in concave in y and is maximized when $y = y_{\max}$, where

$$y_{\max} = \tfrac{1}{2}(k+p)(bq(k+p)+1)^{-1} + z/b.$$

Combining (a), (b), and the fact that $g(y,b) = g(y-1,0)$, it is sufficed to examine the value of (y,z) in Δ to find the minimum value of g. The proof of Theorem 1 has now been completed.

Proof of Theorem 3. The proof of this theorem basically follows the proof of Theorem 4.2 of Majumdar (1992), but the procedure is much involved since $m = [U/bv] \neq 0$. From Lemma 2 and the fact that $g_0(y, b; \pi) = g_0(y-1, 0; \pi)$, proving the theorem is equivalent to prove the following statement.

$$\text{If } g_0(y,z;\pi_2) < g_0(y,z+1;\pi_2), \quad \text{then } g_0(y,z;\pi_1) < g_0(y,z+1;\pi_1), \tag{A.3}$$

for $y \in \{[(k+1)/2]+1, \ldots, k\}$, and $z \in \{0, 1, \ldots, b-1\}$.

Let $U_1 = by - (z+1)$, $V_1 = by^2 - 2y(z+1) + z + 1$, and $\alpha = (k + \pi^{-1})^{-1}$, then one can show that

$$0 < \frac{1}{2y-1} < \frac{1}{k} \leqslant \frac{U}{V} < \frac{U_1}{V_1} < \frac{(v-1)U_1}{vh(y,z+1) - V_1}, \tag{A.4}$$

$$U/V < ((v-1)U)/(vh(y,z) - V). \tag{A.5}$$

Further define

$$A_1 = \begin{cases} v - 1 - \alpha v(2m-1) + \alpha(2y-1), & \text{if } [U_1/bv] = m-1, \\ v - 1 - \alpha v(2m+1) + \alpha(2y-1), & \text{if } [U_1/bv] = m, \end{cases}$$

$$A_2 = U_1 - \alpha V_1, \qquad A_3 = U - \alpha V, \qquad B_1 = \alpha(2y-1) - 1,$$

$$B_2 = (v-1)U_1 - \alpha(vh(y,z+1) - V_1), \qquad B_3 = (v-1)U - \alpha(vh(y,z) - V).$$

Then

$$\begin{aligned} v^{-1}(g_0(y,z+1;\pi) - g_0(y,z;\pi)) &= (A_2A_3B_2B_3)^{-1}((v-1)^2 A_1A_2A_3 - B_1B_2B_3) \\ &= (A_2A_3B_2B_3)^{-1} q(y,z;\alpha), \quad \text{say.} \end{aligned}$$

It is easy to see that A_2, A_3, B_2, and B_3 are positive for all $\alpha \in [0, 1/k]$. Hence the sign of $v^{-1}(g_0(y,z+1;\pi) - g_0(y,z;\pi))$ is the same as that of $q(y,z;\alpha)$.

Next we investigate the behavior of $q(y,z;\alpha)$ in (A.4). Through some calculation one can obtain the following.

$$q(y,z;\alpha) > 0 \quad \text{for } 0 \leqslant \alpha \leqslant 1/(2y-1), \tag{A.6}$$

$$q(y,z;\alpha) < 0 \quad \text{for } U/V \leqslant \alpha \leqslant U_1/V_1. \tag{A.7}$$

To study the behavior of $q(y,z;\alpha)$ in the remaining intervals, separate discussion is necessary.

Case I: If $[U_1/bv] = m-1$.

Then $q(y,z;((v-1)U_1)/(vh(y,z+1) - V_1)) > 0$. Since $q(y,z;\alpha)$ is cubic in α, by (A.6), (A.7), and the previous statement, we conclude that $q(y,z;\alpha)$ has at least one root in $(U/V, \infty)$, and exactly one root in $[0, U/V]$. Statement (A.3) is thus proved, hence the theorem.

Case II: If $[U_1/bv] = m$.

The proof of Case II is more involved. The following two inequalities are claimed in aid of our discussion:

$$U^2/bv < h(y,z) < (U^2/bv) + (bv/4), \tag{A.8}$$

$$U^2/b \leqslant V < (U^2/b) + (b/4). \tag{A.9}$$

For $[U_1/bv] = m$, the $mbv < U < U+1 = U_1 < (m+1)bv$, and $A_1 = v - 1 + \alpha(2y - 1 - 2mv - v)$.

(a) Suppose $2y - 1 - 2mv - v \geqslant 0$. It is obvious that $q(y,z;((v-1)U_1)/(vh(y,z+1) - V_1)) > 0$. Then by the same statement as in Case I, the theorem is proved.

(b) Suppose $2y - 1 - 2mv - v < 0$. Let $q(y,z;\alpha) = q_1(y,z;\alpha) - q_2(y,z;\alpha)$, where $q_1(y,z;\alpha) = (v-1)^2 A_1 A_2 A_3$, and $q_2(y,z;\alpha) = B_1 B_2 B_3$. $q_1(y,z;\alpha)$ is cubic in α, and has three roots U/V, U_1/V_1, and $(v-1)/(2mv + v - 2y + 1)$, respectively. Through some computation, one can show that $U/V < U_1/V_1 < (v-1)/(2mv + v - 2y + 1)$, and $q_1(y,z;\alpha) > 0$, for $0 \leqslant \alpha < U/V$, and $U_1/V_1 < \alpha < (v-1)/(2mv + v - 2y + 1)$; $q_1(y,z;\alpha) < 0$, for $U/V < \alpha < U_1/V_1$. Hence $q_1(y,z;\alpha)$ is nonincreasing in α in $[0, U/V]$.

$q_2(y,z;\alpha)$ is also a cubic function in α, we denote its three roots by $x_1 = 1/(2y-1)$, $x_2 = (v-1)U_1/((vh(y,z+1) - V_1)$, $x_3 = (v-1)U/(vh(y,z) - V)$. One can show that $q_2(y,z;U_1/V_1) > 0$, and $q_2(y,z;\alpha) < 0$, for $0 < \alpha < x_1$. By (A.8) and (A.9), we can also obtain $x_3 > 1 > U_1/V_1$, $x_2 > 1 > U_1/V_1$. Define

$$q_2(y,z;\alpha_{\max}) = \max_{x_1 \leqslant \alpha \leqslant x_2 \,(\text{or } x_3)} q_2(y,z;\alpha).$$

By taking the derivative of $q_2(y,z;\alpha)$ with respect to α, one derives

$$\alpha_{\max} = \begin{cases} \frac{1}{3}(x_1 + x_2 + x_3 - (x_2^2 - x_1(x_3 - x_1) - x_3(x_2 - x_3) - x_1x_2)^{1/2} \\ \quad \text{if } x_2 \geqslant x_3, \\ \frac{1}{3}(x_1 + x_2 + x_3 - (x_3^2 - x_1(x_2 - x_1) - x_2(x_3 - x_2) - x_1x_3)^{1/2} \\ \quad \text{if } x_2 < x_3 \end{cases}$$

And it can easily be shown that $\alpha_{\max} > m(v-1)/v > 1/k$, regardless of whether $x_2 \geqslant x_3$ or $x_2 < x_3$. Thus $q_2(y,z;\alpha)$ is nondecreasing in α for $\alpha \in [0, 1/k]$. Hence $q(y,z;\alpha) = q_1(y,z;\alpha) - q_2(y,z;\alpha)$ is nonincreasing in α in $[0, 1/k]$, and the theorem is proved. Now statement (A.3) is completely verified and Theorem 3 is thus proved.

References

Cheng, C.S. and C.F. Wu (1980). Balanced repeated measurements designs. *Ann. Statist.* **8**, 1272–1283.

Cheng, C.S., D. Majumdar, J. Stufken and T.E. Ture (1988). Optimal step type designs for comparing test treatments with a control. *J. Amer. Statist. Assoc.* **83**, 377–482.

Giovagnoli, A. and I. Verdinelli (1983). Bayes D-optimal and E-optimal block designs. *Biometrika* **70**, 695–706.

Giovagnoli, A. and I. Verdinelli (1985). Optimal block designs under a hierarchical linear model. In: I.M. Bernardo, M.H. DeGroot, D.V. Lindly and A.F.M. Smith, Eds., *Bayesian Statistics* 2. North-Holland, Amsterdam, 655–662.

Hedayat, A.S., M. Jacrous and D. Majumdar (1988). Optimal designs for comparing test treatments with controls. *Statist. Sci.* **3**, 462–491.

Jacroux, M. and D. Majumdar (1989). Optimal block designs for comparing test treatments with a control when $k > v$. *J. Statist. Plann. Inference* **23**, 381–396.

Majumdar, D. (1988). Optimal block designs for comparing new treatments with a standard treatment. In: Y. Dodge, V.V. Fedorov and H.P. Wynn, Eds., *Optimal Design and Analysis of Experiments*. North-Holland, Amsterdam, 15–27.

Majumdar, D. (1992). Optimal designs for comparing test treatments with control utilizing prior information. *Ann. Statist.* **20**, 216–237.

Majumdar, D. and W.I. Notz (1983). Optimal incomplete block designs for comparing test treatments with a control. *Ann. Statist.* **11**, 258–266.

Owen, R.J. (1970). The optimal design of a two-factor experiment using prior information. *Ann. Math. Statist.* **41**, 1917–1934.

Stufken, J. (1991). Based A-optimal and efficient block designs for comparing test treatments with a standard treatment. *Comm. Statist. Theory Methods* **20**, 3849–3862.

Toman, B. and W.I. Notz (1991). Bayesian optimal experimental design for treatment-control comparisons in the presence of two-way heterogeneity. *J. Statist. Plann. Inference* **27**, 51–63.

Ting, C.P. and W.I. Notz (1988). A-optimal complete block designs for treatment-control comparisons. In: Y. Dodge, V.V. Fedorov and H.P. Wynn, Eds., *Optimal Design and Analysis of Experiments*. North-Holland, Amsterdam, 29–37.

Ting, C.P. and Y.Y. Yang (1993). Bayes A-optimal designs for comparing test treatments with a control when $k > v$. Technical Report No. 93-01, Graduate Institute of Statistics, National Chengchi University.