The Annals of Statistics
1976, Vol. 4, No. 2, 419–422

OPTIMALITY OF TWO AND THREE FACTOR DESIGNS

BY K. R. SHAH,[1] D. RAGHAVARAO[2] AND C. G. KHATRI[1]

University of Waterloo, University of Guelph and Gujarat University

In a two-factor design, a design which is optimal for one factor is shown to be optimal jointly for both the factors with respect to each of A-, D-, and E-optimalities. As an interesting consequence we have that a linked block design is optimal for the estimation of treatment differences. Similar results are also obtained for a class of three factor designs.

1. Introduction and summary. Study of optimal designs in the current literature has been made with reference to a single factor in a two or three factor design with the additive model (Kiefer (1958, 1959)). In Block–Treatment designs or in Row–Column–Treatment designs where one is interested in estimating the treatment differences eliminating the environmental effects this approach is justifiable. However, as in factorial experiments with noninteracting factors there could be situations where one is interested in the estimation of main effects of all the factors. One may wonder if a design which is optimal for the estimation of the main effects of one factor will also be optimal for the joint estimation of the main effects of all the factors. In this paper, we show that a design which is optimal for one factor will be optimal for the other factor and will also be optimal jointly for both the factors with respect to each of A-, D- and E-optimalities as defined by Kiefer (1959). As an interesting consequence of this we note that a linked block design is optimal for the estimation of treatment differences. Similar results are also obtained for a class of three factor designs. The following notation will be used in the next section.

Let $\boldsymbol{\tau}$ denote an $n \times 1$ vector of parameters and let $\sigma^2 V$ be the matrix of variances and covariances for the estimator of $\boldsymbol{\tau}$. A design will be said to be D-better than another design for estimation of $\boldsymbol{\tau}$ if it has smaller value for det V. Similarly a design will be said to be A-better than another design if it has smaller value for trace V. Finally, a design will be said to be E-better if it has smaller value for maximum eigenvalue of V.

2. Optimality results. Consider a class of designs in which p additive factors $F_1, F_2, \cdots, F_p$ are applied so that factor F_i has s_i levels. Let $\Delta_i = \text{diag}\,(r_1{}^i, \cdots, r^i_{s_i})$ where $r_j{}^i$ denotes that number of observations with the jth level of the ith

Received January 1974; revised August 1975.

[1] These authors' work was partially supported by National Research Council of Canada Grant No. A 7272.

[2] Now at Temple University.

AMS 1970 *subject classification*. 62K05.

Key words and phrases. Optimal designs, A-, D-, and E-optimalities, main effects, BIB design, LB design.

factor, $j = 1, 2, \cdots, s_i$; $i = 1, 2, \cdots, p$. Let $\boldsymbol{\tau}_i$ denote the $s_i \times 1$ vector of effects of the s_i levels of the factor F_i and let $\mathbf{T}_i$ denote the corresponding vector of total yields for the levels of factor F_i. Let $N_{ij}(i \neq j)$ denote the $s_i \times s_j$ incidence matrix for the levels of factors F_i and F_j. Let μ denote the general mean and let E_{pq} denote the $p \times q$ matrix with all elements unity. The normal equations for the estimation of parameters are

$$\begin{bmatrix} n & E_{1s_1}\Delta_1 & \cdots & E_{1s_p}\Delta_p \\ \Delta_1 E_{s_1 1} & \Delta_1 & \cdots & N_{1p} \\ \vdots & \vdots & \vdots & \vdots \\ \Delta_p E_{s_p 1} & N_{p1} & \cdots & \Delta_p \end{bmatrix} \begin{bmatrix} \mu \\ \boldsymbol{\tau}_1 \\ \vdots \\ \boldsymbol{\tau}_p \end{bmatrix} = \begin{bmatrix} G \\ \mathbf{T}_1 \\ \vdots \\ \mathbf{T}_p \end{bmatrix} \tag{2.1}$$

where G is the grand total of all the observations and n is the total number of observations.

Eliminating μ we get the following matrix of coefficients for the equations for estimation of contrasts in $\boldsymbol{\tau}_i$'s

$$K = \begin{bmatrix} \bar{\Delta}_1 & \bar{N}_{12} & \cdots & \bar{N}_{1p} \\ \bar{N}_{21} & \bar{\Delta}_2 & \cdots & \bar{N}_{2p} \\ \vdots & \vdots & & \vdots \\ \bar{N}_{p1} & \bar{N}_{p2} & \cdots & \bar{\Delta}_p \end{bmatrix}$$

where $\bar{\Delta}_i = H_i \Delta_i H_i'$, $\bar{N}_{ij} = H_i N_{ij} H_j'$ and H_i is obtained by deleting the first row of a $s_i \times s_i$ orthogonal matrix for which all the elements in the first row are equal. We shall assume that the design is connected, i.e., rank of $K = \sum_{i=1}^{p} (s_i - 1)$.

The following theorems connect the D-optimality of the design for the estimation of $\boldsymbol{\tau}_1$ eliminating $\boldsymbol{\tau}_2, \cdots, \boldsymbol{\tau}_p$ with the D-optimality for the estimation of all $\boldsymbol{\tau}_i$'s.

THEOREM 1. *If two designs D_1 and D_2 have the same matrices Δ_i and N_{ij} for i, $j = 2, 3, \cdots, p$, D_1 will be D-better than D_2 for the estimation of $H_1\boldsymbol{\tau}_1$* iff *it is D-better than D_2 for the estimation of $H_1\boldsymbol{\tau}_1, \cdots, H_p\boldsymbol{\tau}_p$.*

PROOF. The matrix of coefficients for the normal equations for the estimation of $\boldsymbol{\tau}_1$ eliminating $\boldsymbol{\tau}_2, \boldsymbol{\tau}_3, \cdots, \boldsymbol{\tau}_p$ is $\bar{\Delta}_1 - FS^{-1}F'$ where $F = [\bar{N}_{12}| \cdots |\bar{N}_{1p}]$ and S is obtained by deleting the first $(s_1 - 1)$ rows and the first $(s_1 - 1)$ columns of K. Since $\det K = \det S . \det (\bar{\Delta}_1 - FS^{-1}F')$ it is clear that smaller values of $\det K^{-1}$ correspond to smaller values of $\det (\Delta_1 - FS^{-1}F')^{-1}$ when S is fixed.

It would be interesting to examine if a similar result holds for A-optimality. The result is not true with the same generality as in the case of D-optimality. However, for some important classes of designs the following theorem provides the answer.

THEOREM 2. *If for each of designs D_1 and D_2, $\bar{\Delta}_1$ and S are multiples of identity matrices of appropriate order, D_1 will be A-better than D_2 for the estimation of $H_1\boldsymbol{\tau}_1$* iff *D_1 is A-better than D_2 for the estimation of $H_1\boldsymbol{\tau}_1, H_2\boldsymbol{\tau}_2, \cdots, H_p\boldsymbol{\tau}_p$.*

PROOF. Let $\bar{\Delta}_1 = aI_{s_1-1}$ and $S = bI_q$ where $q = \sum_{i=2}^{p} (s_i - 1)$. We may write K^{-1} in the following form

$$K^{-1} = \begin{bmatrix} (aI - b^{-1}FF')^{-1} & -(abI - FF')^{-1}F \\ -F'(abI - FF')^{-1} & b^{-1}I + b^{-2}F'(aI - b^{-1}FF')^{-1}F \end{bmatrix}.$$

It is easy to verify that trace $K^{-1} = \text{trace}\,(aI - b^{-1}FF')^{-1}(1 + ab^{-1}) + b^{-1}(q - s_1 + 1)$.

Thus trace K^{-1} is minimized iff trace $(aI - b^{-1}FF')^{-1}$ is minimized. Since $aI - b^{-1}FF'$ is the matrix of coefficients in the normal equations for the estimation of $H_1\tau_1$ the result follows.

To examine the E-optimality we note that the characteristic polynomial of $aI - b^{-1}FF'$ is $\prod_{i=1}^{s_1-1} (a - b^{-1}\delta_i - \lambda)$ where $\delta_1, \cdots, \delta_{s_1-1}$ are the eigenvalues of FF'. It can be verified that the characteristic polynomial of K is

$$(b - \lambda)^{q-s_1+1} \prod_{i=1}^{s_1-1} \left\{\frac{a + b + ((a - b)^2 + 4\delta_i)^{\frac{1}{2}}}{2} - \lambda\right\} \times \left\{\frac{a + b - ((a - b)^2 + 4\delta_i)^{\frac{1}{2}}}{2} - \lambda\right\}.$$

Thus in each case the smallest eigenvalue will correspond to the smallest δ_i. Thus Theorem 2 also holds if we are dealing with E-optimality instead of A-optimality.

We shall now consider applications of these theorems to some standard settings. First we consider two factor designs for which $\Delta_1 = r_1 I_{s_1}$ and $\Delta_2 = r_2 I_{s_2}$. Theorems 1 and 2 indicate that for any two designs, D_1 and D_2 in this class D_1 is more efficient than D_2 for the joint estimation of the main effects of both the factors iff it is more efficient than D_2 for estimating the main effects of factor F_2. Interchanging the two factors we can also claim that this holds iff D_1 is more efficient than D_2 for estimating the main effects of F_1. Thus D_1 is more efficient than D_2 for estimation of F_1 iff it is more efficient than D_2 for estimation of the main effects of F_2.

As a special case we consider block designs where F_1 refers to treatments and F_2 refers to blocks. Thus if we have v treatments arranged in b blocks each of size k so that each treatment is repeated r times in the whole design $\Delta_1 = rI_v$ and $\Delta_2 = kI_b$ giving $\bar{\Delta}_1 = rI_{v-1}$ and $S = kI_{b-1}$.

In this class of designs, D_1 is more efficient than D_2 for the estimation of treatment effects iff it is more efficient for the estimation of block effects. It is well known that in this class of designs a balanced incomplete block design (BIBD) is optimal for the estimation of treatment effects. The above argument shows that the BIBD is also optimal for the estimation of block effects. Interchanging the roles of blocks and treatments we get the following corollary.

COROLLARY 1. *If the class of designs with parameters b, v, r, k contains a linked block* (LB) *design that design is A-, D-, and E-optimal for the estimation of treatment effects.*

Next we consider designs for comparison of v treatments arranged in s rows and s columns. Denoting the treatments by F_1 the rows by F_2 and the columns by F_3 we have $\Delta_2 = \Delta_3 = sI_s$ and $N_{23} = E_{ss}$. It is easy to see that if each of the v treatments occurs r times $\bar{\Delta}_1 = rI_{v-1}$ and $S = sI_{2s-2}$ and thus Theorem 2 can be applied. This gives us the following corollary.

COROLLARY 2. *In the class of equireplicate designs for comparing v treatments in s rows and s columns, D_1 is more efficient than D_2 for the joint estimation of treatment, row and column effects* iff *it is more efficient than D_2 for the the estimation of treatment effects only.*

Results of Theorem 2 do not necessarily hold when S is not a multiple of the identity matrix. This is demonstrated by the following simple example in which we consider the following two arrangements for comparing two treatments in rows and columns.

$$D_1 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 1 & 1 \end{bmatrix}, \quad D_2 = \begin{bmatrix} 1 & 1 & 1 & 1 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

It is easy to verify that for these two designs trace K^{-1} is not the same even though the difference between the two treatments is estimated with the same precision in each design. For each of these designs the matrix of coefficients for the estimation of treatments eliminating rows and columns is $\begin{pmatrix} 6 & -6 \\ -6 & 6 \end{pmatrix}$ and thus the two designs are equivalent for the purpose of estimation of treatment effects. The eigenvalues for the K-matrix of D_1 are $[4, 4, 4, 4, 4, 4, 4, 8, 8, 12 + 4(3)^{\frac{1}{2}}, 12 - 4(3)^{\frac{1}{2}}]$. For the K-matrix of D_2 the eigenvalues are $[4, 4, 4, 4, 4, 4, 8, 8, 8, 10 + (52)^{\frac{1}{2}}, 10 - (52)^{\frac{1}{2}}]$. Evidently trace K^{-1} is not the same for these two designs. Of course det K is the same for the two designs.

REFERENCES

KIEFER, J. (1958). On the nonrandomized optimality and randomized nonoptimality of symmetrical designs. *Ann. Math. Statist.* **29** 675–699.

KIEFER, J. (1959). Optimum experimental designs. *J. Roy. Statist. Soc. Ser. B.* **21** 272–319.

DEPARTMENT OF STATISTICS
UNIVERSITY OF WATERLOO
WATERLOO, CANADA

DEPARTMENT OF STATISTICS
TEMPLE UNIVERSITY
SCHOOL OF BUSINESS ADMINISTRATION
PHILADELPHIA, PENNSYLVANIA 19122

DEPARTMENT OF STATISTICS
GUJARAT UNIVERSITY
AHMEDABAD, GUJARAT
INDIA