The Annals of Statistics
1978, Vol. 6, No. 6, 1262–1272

OPTIMAL DESIGNS FOR THE ELIMINATION OF MULTI-WAY HETEROGENEITY

BY CHING-SHUI CHENG
University of California, Berkeley

The purpose of this paper is to study optimal designs for the elimination of multi-way heterogeneity. The C-matrix for the n-way heterogeneity setting when $n > 2$ is derived. It turns out to be a natural extension of the known formulas in the lower dimensional case. It is shown that under some regularity, the search for optimal designs can be reduced to that in a lower-way setting. Youden hyperrectangles are defined as higher dimensional generalizations of balanced block designs and generalized Youden designs. When all the sides are equal, they are called Youden hypercubes. It is shown that a Youden hyperrectangle is E-optimal and a Youden hypercube is A- and D-optimal. The latter is quite interesting since it is not always true in two-way settings.

1. Introduction. The optimality of symmetric designs for the elimination of heterogeneity has been substantially studied by Kiefer (1958, 1959, 1971, 1975) for the one-way and two-way settings. In this series of important papers, he made clear the role played by symmetry and developed some powerful methods for proving optimality. Using his methods, we continue the investigation in the multi-way situation.

In an n-way heterogeneity setting, we are given v varieties and an n-dimensional hyperrectangle of size $b_1 \times b_2 \times \cdots \times b_n$, where b_i is the number of levels of the ith factor. There are $b_1 b_2 \ldots b_n$ cells in this hyperrectangle. We can coordinatize them by the n-tuples of integers $(i_1, i_2, \cdots, i_n)$ with $1 \leqslant i_j \leqslant b_j$. The usual additive model (no interactions) specifies the expectation of an observation on variety i in the cell $(j_1, \cdots, j_n)$ to be $\alpha_i + \sum_{k=1}^n \beta_{j_k}{}^{(k)}$, where α_i and $\beta_{j_k}{}^{(k)}$ are the effects of variety i and the j_kth level of factor k, respectively. Also, we assume that all the observations are uncorrelated with common variance. The usual restriction to one observation per cell is unnecessary. Instead, suppose there are t observations taken in each cell, then a *design* is an allocation of the v variety labels $1, 2, \cdots, v$ into these cells with t varieties (not necessarily all different) in each cell. We denote such a setting by $E(v; b_1, \cdots, b_n; t)$. If there exists a subset S of $\{1, 2, \cdots, n\}$ such that $v|(\prod_{l \neq i} b_l)t$, $\forall i \in S$, then we say that $E(v; b_1, \cdots, b_n; t)$ is *regular relative to the factors in* S. A setting is called *completely regular* if it is regular relative to $\{1, 2, \cdots, n\}$.

For a specified design d, let N_{di} be the incidence matrix between the v varieties and the b_i levels of the ith factor, that is, the (s, u)th element of N_{di} is the total

Received May 1977; revised November 1977.
[1]This paper is part of the author's doctoral dissertation written at Cornell University. Research sponsored in part by NSF Grant MCS-75-22481.
AMS (1970) *subject classifications*. Primary 62K05; Secondary 62K10.
Key words and phrases. D-optimality, A-optimality, E-optimality, universal optimality, C-matrix, Youden hyperrectangle, Youden hypercube, regular settings.

number of times that variety s appears in the cells with u as the ith coordinate. Also, let $\operatorname{diag}(a_1, \cdots, a_k)$ be the diagonal matrix with diagonal elements $a_1, \cdots, a_k$, and J_{k_1, k_2} be the $k_1 \times k_2$ matrix with all entries equal to 1. We will also write $J_k = J_{k,k}$. Then for a design d, the coefficient matrix of the normal equation for estimating $(\alpha_1, \cdots, \alpha_v; \beta_1^{(1)}, \cdots, \beta_{b_1}^{(1)}; \cdots; \beta_1^{(n)}, \cdots, \beta_{b_n}^{(n)})'$ is

$$\begin{bmatrix} \operatorname{diag}(r_{d1}, \cdots, r_{dv}) & N_{d1} & N_{d2} & \cdots & N_{dn} \\ N'_{d1} & (\Pi_{j\neq 1} b_j) t I_{b_1} & (\Pi_{j\neq 1,2} b_j) t J_{b_1, b_2} & & (\Pi_{j \neq 1, n} b_j) t J_{b_1, b_n} \\ N'_{d2} & (\Pi_{j\neq 1,2} b_j) t J_{b_2, b_1} & & & \\ \vdots & \vdots & & & \\ N'_{dn} & (\Pi_{j\neq 1,n} b_j) t J_{b_n, b_1} & \cdots & \cdots & (\Pi_{j\neq n} b_j) t I_{b_n} \end{bmatrix}, \tag{1.1}$$

where r_{di} is the number of replications of variety i.

If we are interested in the variety effects only, then the coefficient matrix of the reduced normal equation for the variety effects is

$$C_d = \operatorname{diag}(r_{d1}, \cdots, r_{dv}) - (N_{d1}, \cdots, N_{dn}) E^-(N_{d1}, \cdots, N_{dn})', \tag{1.2}$$

where E is the matrix obtained by deleting the first v rows and v columns of (1.1), and E^- is a generalized inverse of E, i.e., E^- is a matrix such that $EE^-E = E$.

This matrix C_d is called the *C-matrix* or the *information matrix* of the design d. A *ψ-optimal design d^** is one which minimizes a functional ψ of C_d over all possible designs. The well-known A-, D- and E-criteria are defined in the same way as in Kiefer (1975).

In his 1958 paper Kiefer generalized balanced incomplete block designs (BIBD) and Youden squares to balanced block designs (BBD) and generalized Youden designs (GYD), and proved that in the one-way heterogeneity setting, any BBD is A-, D- and E-optimal over all block designs with the same parameter values. Later, in Kiefer (1975), a striking result on the universal optimality of BBD's was established. In the same paper, it was also proved that in the 2-way heterogeneity setting, a GYD is A- and E-optimal, and except for the case $v = 4$, it is D-optimal.

We say that a design d is *balanced in direction i*, or, *balanced relative to factor i* $(1 \leqslant i \leqslant n)$, if d reduces to a BBD when we consider the union of all the cells with the same ith coordinate as a block. Then a natural higher dimensional generalization of BBD and GYD is a design d^* which is balanced in each of the n directions. We call such a design a *Youden hyperrectangle* (YHR). If all b_i's are equal, it is called a *Youden hypercube* (YHC). We denote a Youden hyperrectangle in $E(v; b_1, \cdots, b_n; t)$ by YHR $(v; b_1, \cdots, b_n; t)$. When $b_1 = \cdots = b_n = b$, the corresponding Youden hypercube is abbreviated as YHC $(v; b^n; t)$.

The purpose of the present research is to extend the optimality of BBD's and GYD's to Youden hyperrectangles. In order to investigate optimality, we have to compute the C-matrix explicitly. This is done in Section 2. It turns out to be a natural extension of the known formulas for $n = 1$ and 2. In Section 3 we begin with regular settings because they are easier to deal with. In Section 4, the E-optimality of general (not necessarily regular) Youden hyperrectangles is established. We consider A- and D-criteria in Section 5, where we are able to prove the D- and A-optimality of any Youden hypercube for $n \geqslant 3$. It is interesting to note that Kiefer (1975) has proved that any YHC $(4; b^2; 1)$ is not D-optimal. This peculiarity thus occurs only for $n = 2$.

Much of the material in this paper refers to Kiefer (1975). The reader is expected to have a copy of that paper in hand while reading the current paper.

The construction of Youden hyperrectangles is treated in Cheng (1977).

2. Computation of the C-matrix. The derivation of the C-matrix involves taking generalized inverses of some matrices with zero row and column sums. Choosing appropriate generalized inverses can significantly simplify the computation. In the present case, we will use the generalized inverse which also has zero row and column sums. The reason for this choice will become clear as we carry out the computation.

Let $m = \prod_{j=1}^n b_j$, then the matrix E defined in Section 1 can be written as:

$$(2.1) \qquad E = mt\begin{bmatrix} b_1^{-1}I_{b_1} & (b_1b_2)^{-1}J_{b_1,b_2} & \cdots & (b_1b_n)^{-1}J_{b_1,b_n} \\ (b_1b_2)^{-1}J_{b_2,b_1} & b_2^{-1}I_{b_2} & \cdots & (b_2b_n)^{-1}J_{b_2,b_n} \\ \vdots & \vdots & & \vdots \\ (b_1b_n)^{-1}J_{b_n,b_1} & \cdots & \cdots & b_n^{-1}I_{b_n} \end{bmatrix}.$$

We have the following

LEMMA 2.1. *For any positive integer b and any real number $r \neq 0$, $r(I_b - b^{-1}J_b)$ and $r^{-1}(I_b - b^{-1}J_b)$ are generalized inverses of each other.*

PROOF. Follows from inspection or from the idempotence of $I_b - b^{-1}J_b$. □

Thus, $r^{-1}(I_b - b^{-1}J_b)$ is the generalized inverse of $r(I_b - b^{-1}J_b)$ with zero row and column sums.

LEMMA 2.2. *For any positive integers $b_1, \cdots, b_n$ and t,*

$$(2.2) \qquad (mt)^{-1}\operatorname{diag}(b_1I_{b_1}, b_2I_{b_2} - J_{b_2}, \cdots, b_nI_{b_n} - J_{b_n})$$

is a generalized inverse of the E matrix of (2.1).

PROOF. This can be proved by induction. The case $n = 1$ is clear. Now, suppose it is true for $n - 1$ with $n > 1$. Partition (2.1) as $\begin{bmatrix} A & B \\ B' & D \end{bmatrix}$, where D is $b_n \times b_n$, then we have

$$\begin{bmatrix} A & B \\ B' & D \end{bmatrix}^- = \begin{bmatrix} A^- + A^-BQ^-B'A^- & -A^-BQ^- \\ -Q^-B'A^- & Q^- \end{bmatrix},$$

where $Q = D - B'A^-B$; e.g., see Searle (1971), page 27.

By the induction hypothesis, $(mt)^{-1}\operatorname{diag}(b_1 I_{b_1}, b_2 I_{b_2} - J_{b_2}, \cdots, b_{n-1} I_{b_{n-1}} - J_{b_{n-1}})$ is a generalized inverse of A.

Each of the matrices $b_h I_{b_h} - J_{b_h}$ has zero row and column sums, therefore $J_{b_n, b_h}(b_h I_{b_h} - J_{b_h}) = 0$, $\forall h$. It follows that

$$Q = mtb_n^{-1} I_{b_n} - mt(b_1 b_n)^{-1} J_{b_n, b_1}\left[(mt)^{-1} b_1 I_{b_1}\right] mt(b_1 b_n)^{-1} J_{b_1, b_n}$$
$$- \Sigma_{h=2}^{n-1} mt(b_h b_n)^{-1} J_{b_n, b_h}\left[(mt)^{-1}(b_h I_{b_h} - J_{b_h})\right] mt(b_h b_n)^{-1} J_{b_h, b_n}$$
$$= mtb_n^{-1}\left[I_{b_n} - b_n^{-1} J_{b_n}\right].$$

Hence by Lemma 2.1, we can choose Q^- to be $(mt)^{-1}[b_n I_{b_n} - J_{b_n}]$, which has zero row and column sums. Then since each row of B has constant components, we have $BQ^- = 0$. Therefore

$$A^- B Q^- B' A^- = -A^- B Q^- = -Q^- B' A^- = 0. \tag{2.3}$$

Consequently,

$$\begin{bmatrix} A & B \\ B' & D \end{bmatrix}^- = \begin{bmatrix} A^- & 0 \\ 0 & Q^- \end{bmatrix},$$

which is (2.2). □

From Lemma 2.2 and (1.2), we immediately get the following.

THEOREM 2.1. *For a design d in $E(v; b_1, \cdots, b_n; t)$,*

$$C_d = \operatorname{diag}(r_{d1}, \cdots, r_{dv}) - (mt)^{-1} b_1 N_{d1} N'_{d1}$$
$$- (mt)^{-1} \Sigma_{h=2}^{n} b_h N_{dh}\left(I_{b_h} - b_h^{-1} J_{b_h}\right) N'_{dh}$$
$$= \operatorname{diag}(r_{d1}, \cdots, r_{dv}) - (mt)^{-1} \Sigma_{h=1}^{n} b_h N_{dh} N'_{dh}$$
$$+ (n-1)(mt)^{-1}\left[r_{di} r_{dj}\right]_{v \times v},$$

where $[r_{di} r_{dj}]_{v\times v}$ is the $v \times v$ matrix whose (i, j)th entry is $r_{di} r_{dj}$, and $m = \Pi_{i=1}^{n} b_i$.

Note that this is consistent with the well-known forms of C_d when $n = 1$ or 2.

Before we end this section, we would also like to write down the whole reduced normal equation which is indispensable to the analysis of the design, although from an optimum design theoretic point of view, we are only interested in the *coefficient matrix* of this equation.

For each i with $1 \leqslant i \leqslant v$, let T_{di} be the sum of all the observations on variety i. Similarly, for any h with $1 \leqslant h \leqslant n$ and any j with $1 \leqslant j \leqslant b_h$, let B_{dhj} be the sum of all the observations in the $\Pi_{l \neq h} b_l$ cells with j as the hth coordinate. Write $T_d = (T_{d1}, \cdots, T_{dv})'$, $B_{dh} = (B_{dh1}, \cdots, B_{dh, b_h})'$. Then it can be shown that the reduced normal equation for the variety effects under a design d in

$E(v;\, b_1, \cdots, b_n;\, t)$ is

$$C_d \begin{bmatrix} \alpha_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \alpha_v \end{bmatrix} = T_d - (mt)^{-1}\Sigma_{h=1}^{n} b_h N_{dh} B_{dh} + (n-1)(mt)^{-1} \begin{bmatrix} g_d r_{d1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ g_d r_{dv} \end{bmatrix}, \tag{2.4}$$

where g_d is the sum of all the $b_1 b_2 \ldots b_n t$ observations.

Again, this generalizes the known equations for $n = 1$ and 2. From this equation it is easy to write down the ANOVA table for a Youden hyperrectangle.

3. Optimal designs in regular settings. Suppose we have a setting $E(v;\, b_1, \cdots, b_n;\, t)$ which is regular relative to a set S of $n - k$ factors, say $S = \{k+1, \cdots, n\}$ with $k < n$. If we ignore the $n - k$ factors in S, then the original setting is reduced to $E(v;\, b_1, \cdots, b_k;\, b_{k+1} \cdots b_n t)$, i.e., a k-way heterogeneity setting with $b_{k+1} \cdots b_n t$ observations in each cell. A design in $E(v;\, b_1, \cdots, b_n;\, t)$ can be automatically considered as a design in $E(v;\, b_1, \cdots, b_k;\, b_{k+1} \cdots b_n t)$. We use C_d^S to denote the C-matrix of d when the factors in S are ignored. Then by Theorem 2.1,

$$C_d = C_d^S - (mt)^{-1}\Sigma_{h=k+1}^{n} b_h N_{dh}\left(I_{b_h} - b_h^{-1} J_{b_h}\right) N'_{dh}. \tag{3.1}$$

Each of the matrices $N_{dh}(I_{b_h} - b_h^{-1} J_{b_h}) N'_{dh}$ is nonnegative definite. Therefore, if ψ is a *nonincreasing criterion* in the sense that $\psi(C) \leqslant \psi(D)$ whenever $C - D$ is nonnegative definite, then $\psi(C_d^S) \leqslant \psi(C_d)$ for any design d in $E(v;\, b_1, \cdots, b_n;\, t)$.

Let d^* be a design which is balanced relative to the factors in S. Then by the regularity assumption, $N_{d^*h} = mt(vb_h)^{-1} J_{v,\, b_h}$, $\forall h = k+1, \cdots, n$. It follows that

$$(mt)^{-1} b_h N_{d^*h}\left(I_{b_h} - b_h^{-1} J_{b_h}\right) N'_{d^*h} = 0, \qquad \forall h > k,$$

$C_{d^*} = C_{d^*}^S$ and certainly $\psi(C_{d^*}) = \psi(C_{d^*}^S)$. Furthermore, if d^* is ψ-optimal when considered as a design in $E(v;\, b_1, \cdots, b_k;\, b_{k+1} \cdots b_n t)$, then for any other design d in $E(v;\, b_1, \cdots, b_n;\, t)$, $\psi(C_{d^*}) = \psi(C_{d^*}^S) \leqslant \psi(C_d^S) \leqslant \psi(C_d)$.

In summary, we have the following

THEOREM 3.1. *Let $E(v;\, b_1, \cdots, b_n;\, t)$ be a setting which is regular relative to $S = \{k+1, \cdots, n\}$ with $k < n$. Also, assume d^* is a design which is balanced relative to the factors in S, and is ψ-optimal with respect to a nonincreasing criterion ψ when considered as a design in $E(v;\, b_1, \cdots, b_k;\, b_{k+1} \cdots b_n t)$. Then d^* is also ψ-optimal in $E(v;\, b_1, \cdots, b_n;\, t)$.*

The proof of Theorem 3.1 gives a neat argument showing that under regularity the search for optimal designs can be reduced to that in a lower-way setting. Note that A-, E- and D-criteria are nonincreasing. From a decision theoretic point of view, we are only interested in nonincreasing criteria. (See Kiefer (1958)).

By the definition of Youden hyperrectangles and Theorem 2.1, the C-matrix of a Youden hyperrectangle is completely symmetric, i.e., all the diagonal elements are equal, and all the off-diagonal elements are equal. Combining Proposition 1 of Kiefer (1975), the fact that a BBD is trace maximal, and our Theorem 3.1, we get

COROLLARY 3.1.1. *If $E(v; b_1, \cdots, b_n; t)$ is regular relative to $n-1$ factors, then any Youden hyperrectangle* YHR$(v; b_1, \cdots, b_n; t)$ *is universally optimal*, i.e., *it minimizes all functions $\Phi : \mathcal{B}_{v,0} \to (-\infty, +\infty]$ satisfying*

(a) *Φ is convex,*
(b) *$\Phi(bC)$ is nonincreasing in the scalar $b \geqslant 0$,*
(c) *Φ is invariant under each permutation of rows and (the same on) columns,*

where $\mathcal{B}_{v,0}$ is the set of all $v \times v$ nonnegative definite matrices with zero row and column sums.

This includes Kiefer's (1975) result on the universal optimality of regular generalized Youden designs as a special case.

An immediate application of Corollary 3.1.1 is the following

COROLLARY 3.1.2. *If there exists a* YHR$(v; b_1, \cdots, b_n; t)$ *and v is a prime number, then it is universally optimal.*

Kishen (1949) defined a v-sided m-fold Latin hypercube of the rth order with $r < m$, which is exactly a YHR in the completely regular setting

$$E\big(v^r; \underbrace{v, \cdots, v}_{m}; 1\big).$$

Accordingly, by Corollary 3.1.1, any Latin hypercube is universally optimal.

By Fisher's inequality on balanced incomplete block designs, it is easily seen that if there exists a YHR$(v; b_1, \cdots, b_n; t)$ and $b_i < v$ for some i, then $v|(\Pi_{l \neq i} b_l)t$ and hence $E(v; b_1, \cdots, b_n; t)$ is regular relative to the single factor i. In particular, if $b_1 = b_2 = \cdots = b_n$, then it is completely regular. Therefore, we have

COROLLARY 3.1.3. *If there exists a* YHC$(v; b^n; t)$ *and $b < v$, then it is universally optimal.*

Similarly,

COROLLARY 3.1.4. *If there exists a design d^* which is balanced in direction n and $b_n < v$, then for any nonincreasing criterion ψ, d^* is ψ-optimal in $E(v; b_1, \cdots, b_n; t)$ if it is ψ-optimal in $E(v; b_1, \cdots, b_{n-1}; b_n t)$.*

And

COROLLARY 3.1.5. *Assume* $E(v;\, b_1, \cdots, b_n;\, t)$ *is regular relative to* $\{2, 3, \cdots, n\}$. *If there exists a design* d^* *which is balanced relative to* $\{2, \cdots, n\}$ *and is an* MB GD PBBD *of type* 1 *in the setting* $E(v;\, b_1;\, b_2 \cdots b_n t)$, *then it is optimal* w.r.t. *any generalized criterion of type* 1 *over all possible designs in* $E(v;\, b_1, \cdots, b_n;\, t)$. *If such a design exists, then any design which is optimal in* $E(v;\, b_1, \cdots, b_n;\, t)$ w.r.t. *any particular type* 1 *criterion* (*not a generalized one*) *should also be of this sort.*

There is an obvious type 2 analogue of Corollary 3.1.5. For the notions of type i criteria and MB GD PBBD of type i, see Cheng (1978).

For example, the following design is optimal w.r.t. any generalized type 1 criterion in $E(4;\, 2, 4;\, 1)$:

$$\begin{matrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{matrix}.$$

4. E-optimality of Youden hyperrectangles. Let d^* be a YHR $(v;\, b_1, \cdots, b_n;\, t)$. Then the C-matrix of d^* is completely symmetric. As in page 340 of Kiefer (1975), we define

$$c(r) = \max_{\{d\,:\, r_{dj} = r\}} c_{djj}. \tag{4.1}$$

Then

$$g(r) \overset{\text{def}}{=} mtc(r) = mtr - \Sigma_{l=1}^{n} b_l h(r, b_l) + (n-1)r^2, \tag{4.2}$$

where

$$\begin{aligned} h(r, k) &= \min_{\{\Sigma_1^k n_i = r\}} \Sigma_1^k n_i^2 \\ &= [r - k\,\mathrm{int}(r/k)][1 + \mathrm{int}(r/k)]^2 \\ &\quad + [k - r + k\,\mathrm{int}(r/k)][\mathrm{int}(r/k)]^2 \\ &= -k[\mathrm{int}(r/k)]^2 + (2r - k)[\mathrm{int}(r/k)] + r. \end{aligned} \tag{4.3}$$

In the above expression, $\mathrm{int}(r/k)$ is the largest integer $\leqslant r/k$.

By the last expression of (4.3), we have $h(r+1, k) - h(r, k) = 1 + 2\,\mathrm{int}(r/k)$. Hence

$$\begin{aligned} \Delta(r) \overset{\text{def}}{=} g(r+1) - g(r) &= mt + (n-1)(2r+1) \\ &\quad - \Sigma_{l=1}^{n} b_l[1 + 2\,\mathrm{int}(r/b_l)]. \end{aligned} \tag{4.4}$$

Then by the same argument as in Section 3.2 of Kiefer (1975), a sufficient condition for the E-optimality of a $\mathrm{YHR}(v;\, b_1, \cdots, b_n;\, t)$ is that $\Delta(r) \geqslant 0$ for $r < \bar{r}$, where $\bar{r} = v^{-1}mt$.

By Corollary 3.1.4, if $b_i < v$ for some i, then d^* is E-optimal in $E(v;\, b_1, \cdots, b_n;\, t)$ if it is E-optimal in the corresponding $(n-1)$-way heterogeneity setting with factor i ignored. Therefore, in our E-optimality proof, we may assume $b_i \geqslant v \forall i$. Also, by Corollary 3.1.2, it suffices to consider the case $v \geqslant 4$. Then from (4.4),

$$\Delta(r) \geqslant mt + (n-1)(2r+1) - \Sigma_{i=1}^n b_i - 2nr$$
$$= mt - \Sigma_{i=1}^n b_i - 2r + (n-1).$$

If $r < \bar{r}$, then $r \leqslant \bar{r} - 1$, and hence

$$\Delta(r) \geqslant mt - \Sigma_{i=1}^n b_i - 2mt/v + 2 + (n-1)$$
$$\geqslant m/2 - \Sigma_{i=1}^n b_i + n + 1.$$

The last inequality is true because $v \geqslant 4$ and $t \geqslant 1$.

Since $b_i \geqslant 4 \ \forall i$, we can write b_i as $4 + \varepsilon_i$ with $\varepsilon_i \geqslant 0$. Then

$$m = \Pi_{i=1}^n b_i$$
$$= \Pi_{i=1}^n (4 + \varepsilon_i) > 4^n + 4^{n-1}\Sigma_{i=1}^n \varepsilon_i.$$

So

$$\Delta(r) > 2 \cdot 4^{n-1} + (2 \cdot 4^{n-2} - 1)\Sigma_{i=1}^n \varepsilon_i - 4n + n + 1$$
$$> 2(4^{n-1} - 3n/2)$$
$$> 0 \quad \text{for} \quad n \geqslant 2.$$

Therefore, we conclude

THEOREM 4.1. *If a Youden hyperrectangle in $E(v; b_1, \cdots, b_n; t)$ exists, then it is E-optimal.*

5. *D*- and *A*-optimality of Youden hypercubes. In this section, we will prove the *D*- and *A*-optimality of Youden hypercubes. Later, the difficulty for general Youden hyperrectangles will be indicated.

Let d^* be a YHC $(v; b^n; t)$ with $n \geqslant 3$. As before, in our optimality proof, we may assume $b \geqslant v$, $v \nmid b^{n-1}t$, and v is not a prime.

As in Kiefer (1975), we write $[C, D]$ for an interval of successive integers. Let $\mathfrak{N} = \{k : 0 \leqslant k \leqslant b^n t, b|k\}$, and $\mathfrak{M} = \{k \in \mathfrak{N} : k \leqslant b^n t/2\}$. If $C, D \in \mathfrak{N}$, $C < D$, and no integer between C and D is in $\mathfrak{N}$, we call $[C, D]$ an elementary interval. The elementary interval $[C_0, D_0]$ containing $\bar{r} = b^n t/v$ is called the basic interval.

Then all the properties of g listed on page 344 of Kiefer (1975) can be established, i.e., we have

(i) For each elementary interval $[C, D]$, $\Delta(r)$ is linear in r and increasing for $C \leqslant r < D$, i.e., g is a convex quadratic on each elementary interval.

(ii) g is increasing in each elementary interval $[C, D]$ with $D \leqslant D_0$.

(iii) g is symmetric about $b^n t/2$.

(iv) If $C_1, C_2 \in \mathfrak{N}$ with $C_1 < C_2$, then $\Delta(C_1) \geqslant \Delta(C_2)$ and $\Delta(C_1 - 1) \geqslant \Delta(C_2 - 1)$.

(v) g is nondecreasing on $\mathfrak{M}$ and nonincreasing on the remainder of $\mathfrak{N}$.

Therefore, by the argument on page 345 of Kiefer (1975) and especially (3.17) on that page, a YHC $(v;\, b^n;\, t)$ is D-optimal if

$$(5.1)\qquad \log g(r+1) - \log g(r) \quad \text{is nonincreasing} \quad \text{for} \quad C_0 \leqslant r < D_0,$$

i.e., if

$$(5.2)\qquad g^2(r) - g(r-1)g(r+1) \geqslant 0 \qquad \text{for} \quad C_0 < r < D_0.$$

Substituting $g(r-1) = g(r) - \Delta(r-1)$ and $g(r+1) = g(r) + \Delta(r)$, this becomes

$$(5.3)\qquad C_0 < r < D_0 \Rightarrow 0 \leqslant \Delta(r)\Delta(r-1) + g(r)[\Delta(r-1) - \Delta(r)].$$

Let $\Gamma_0(r) = \Delta(r)\Delta(r-1) + g(r)[\Delta(r-1) - \Delta(r)]$. Also, let $B = C_0/b$, which is $\operatorname{int}(r/b)$ when $C_0 \leqslant r < D_0$. Then from (4.2) and (4.3), we can write $g(r) = (n-1)r^2 + \beta r + \alpha$ on the interval $[C_0, D_0 - 1]$ with

$$(5.4)\qquad \alpha = nC_0^2 + \sigma C_0,$$

and

$$(5.5)\qquad \beta = \pi - \sigma - 2nC_0,$$

where $\pi = b^n t$, and $\sigma = nb$.

Since $v | b^n t$ and $v \nmid b^{n-1}t$, there is a prime number p such that $p^n | v$. Consequently, we have

$$(5.6)\qquad v \geqslant 2^n \geqslant 2n + 2.$$

The last inequality is true because $n \geqslant 3$.

On the interval $[C_0 + 1, D_0 - 1]$, we have

$$(5.7)\qquad \Gamma_0(r) = 2(n-1)^2 r^2 + 2(n-1)\beta r - 2\alpha(n-1) + \beta^2 - (n-1)^2.$$

Consider the right side of (5.7) as a differentiable function of real r, then

$$(5.8)\qquad \frac{d}{dr}\Gamma_0(r) = 2(n-1)[2(n-1)r + \beta] = 2(n-1)g'(r) > 0 \quad \text{for } r \quad \text{with} \quad C_0 \leqslant r \leqslant D_0 - 1, \quad \text{by (ii).}$$

Therefore, a sufficient condition for (5.3) is

$$(5.9)\qquad \Gamma_0(C_0 + 1) \geqslant 0.$$

From (5.7) we get

$$(5.10)\qquad \Gamma_0(C_0+1) = (2n+2)C_0^2 + [-4n + 4 + 4\sigma - (2n+2)\pi]C_0 + (n - 1 + \pi - \sigma)^2.$$

Let

$$Q(x) = (2n + 2)x^2 + [-4n + 4 + 4\sigma - (2n + 2)\pi]x + (n - 1 + \pi - \sigma)^2.$$

With this notation, showing that $Q(C_0) \geqslant 0$ establishes (5.9), and hence (5.3). We do this by seeing that $C_0 \leqslant \pi/(2n + 2)$, that Q is decreasing on $(-\infty, \pi/(2n + 2))$ and that $Q(\pi/(2n + 2)) \geqslant 0$.

The first two facts follow easily from $C_0 < \bar{r}$, (5.6), and the linearity of Q'. For the third,

$$\begin{aligned}(2n + 2)Q(\pi/(2n + 2)) &= \pi^2 + [-4n + 4 + 4\sigma - (2n + 2)\pi]\pi \\ &\quad + (n - 1 + \pi - \sigma)^2(2n + 2) \\ &= \pi^2 + 4n(n - 1 - \sigma)\pi + (n - 1 - \sigma)^2(2n + 2) \\ &\geqslant \pi[\pi + 4n(n - 1 - \sigma)] \\ &\geqslant b^n + 4n(n - 1 - \sigma) \\ &\geqslant b^n - 4n^2 b \\ &\geqslant b^{n-1} - 4n^2 \\ &\geqslant (2n + 2)^{n-1} - 4n^2,\end{aligned}$$

which is $\geqslant 0$ whenever $n \geqslant 3$.

This proves (5.9). Therefore we have

THEOREM 5.1. *If $n \geqslant 3$, and there exists a* YHC$(v; b^n; t)$, *then it is D-optimal and hence A-optimal.*

This is a little bit surprising at the first look. Kiefer (1975) proved that any YHC$(4; b^2; 1)$ is not D-optimal. Why is $n = 2$ special? The crucial point is (5.6). Our optimality proof strongly depends on the condition that $v \geqslant 2n + 2$. This is not satisfied by $v = 4$ and $n = 2$, but for $n \geqslant 3$, it is always true if there is a nonregular YHC$(v; b^n; t)$.

The proof of Theorem 5.1 is similar to that of the D-optimality of generalized Youden designs for $v \geqslant 6$ in Kiefer (1973).

For the general Youden hyperrectangles, if we imitate the above proof by defining $\mathfrak{N} = \{k : 0 \leqslant k \leqslant mt,\ b_i | k \text{ for some } i\}$, then $\Delta(r)$ is not always nondecreasing on $\mathfrak{N}$, and the proof breaks down. Finding new tools for the general setting seems very difficult.

The setting considered in Kiefer (1975) was $E(v; b_1, b_2; 1)$. One can also define generalized Youden designs in the more general setting $E(v; b_1, b_2; t)$. Then Kiefer's argument gives us the following

THEOREM 5.2. *If there exists a* GYD$(v; b_1, b_2; t)$, *then it is A-optimal, and except for $v = 4$, it is D-optimal.*

Therefore, by Theorem 3.1, we have

THEOREM 5.3. *If $E(v; b_1, b_2, \cdots, b_n; t)$ is regular relative to $n - 2$ factors, then a* YHR$(v; b_1, \cdots, b_n; t)$ *is A-optimal, and except for $v = 4$, it is D-optimal.*

And finally,

THEOREM 5.4. *If $E(v; b_1, b_2, \cdots, b_n; t)$ is regular relative to $\{k + 1, \cdots, n\}$, $k \geq 3$, and $b_1 = b_2 = \cdots = b_k$, then a* YHR$(v; b_1, \cdots, b_n; t)$ *is A- and D-optimal.*

Acknowledgment. I would like to express my sincere thanks to Professor Jack Kiefer for his encouragement, guidance and many helpful discussions. I also thank the referee for a careful reading and many valuable suggestions.

REFERENCES

[1] CHENG, C. S. (1978). Optimality of certain asymmetrical experimental designs. *Ann. Statist.* **6** 1239–1261.

[2] CHENG, C. S. (1977). Construction of Youden hyperrectangles. Unpublished manuscript.

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

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

[5] KIEFER, J. (1971). The role of symmetry and approximation in exact design optimality. In *Statistical Decision Theory and Related Topics* (S. S. Gupta and J. Yackel, eds.). Academic Press, New York, 109–118.

[6] KIEFER, J. (1973). D-optimality of the GYD for $v \geq 6$. Unpublished manuscript.

[7] KIEFER, J. (1975). Construction and optimality of generalized Youden designs. In *A Survey of Statistical Design and Linear Models* (J. N. Srivastava, ed.). North-Holland, Amsterdam, 333–353.

[8] KISHEN, K. (1949). On the construction of Latin and hypergraeco-Latin cubes and hypercubes. *J. Indian Soc. Agri. Stat.* **2** 20–48.

[9] SEARLE, S. R. (1971). *Linear Models*. Wiley, New York.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720