The Annals of Statistics
1981, Vol. 9, No. 1, 201–205

OPTIMALITY AND CONSTRUCTION OF PSEUDO-YOUDEN DESIGNS[1]

BY CHING-SHUI CHENG

University of California, Berkeley

In the two-way heterogeneity setting, the optimality of a generalized Youden design (GYD) has been proved by Kiefer (1975a). A GYD is a design which is a balanced block design (BBD) when each of {rows} and {columns} is considered as blocks. It is observed in the present paper that when the number of rows is equal to the number of columns, a design is optimal as long as the rows and columns together form a BBD. Such a design is termed a pseudo-Youden design (PYD). A square GYD is also a PYD, but the converse is not true. Thus, the stringent conditions imposed on the definition of a GYD are substantially relaxed. A PYD is easier to construct and has the same efficiency as a GYD if they exist simultaneously. Patchwork and geometric methods are combined to construct a family of PYD's. It is also indicated when the construction of a GYD is impossible. A 6×6 PYD with 9 varieties is constructed. This design has the property that the number of rows is less than the number of varieties, which is never achieved by a square GYD. There is also an analogous theory for higher-dimensional designs.

1. Introduction and optimality. Kiefer (1975a) proved the optimality of generalized Youden designs (GYD) for the elimination of two-way heterogeneity. The main purpose of this paper is to show that when the number of rows is equal to the number of columns, a more flexible design called pseudo-Youden design (PYD) is also optimal. The advantage of using a PYD is that its construction is much more flexible than a GYD. Sometimes a PYD can be constructed while a GYD does not exist.

To save space, we refer the readers to Kiefer (1975a) and Cheng (1978) for relevant definitions, setups, etc.

In the usual additive and homoscedastic setting for comparing v varieties via b_1b_2 experimental units arranged into b_1 rows and b_2 columns, the coefficient matrix of the reduced normal equation for estimating the variety effects (C-matrix) under a design d is

$$C_d = \mathrm{diag}(r_{d1}, \cdots, r_{dv}) - b_2^{-1}N_{d1}N'_{d1} - b_1^{-1}N_{d2}N'_{d2} + b_1^{-1}b_2^{-1}[r_{di}, r_{dj}], \tag{1.1}$$

where r_{di} is the number of appearances of variety i, N_{d1} is the variety-row incidence matrix, N_{d2} is the variety-column incidence matrix, and $[r_{di}r_{dj}]$ is the $v \times v$ matrix with (i, j)th entry $r_{di}r_{dj}$.

Kiefer (1975a) defined a *generalized Youden design* (GYD) and proved that a GYD is A- and E-optimal, and is D-optimal if $v \neq 4$. Basically, a GYD is a design which is balanced in both directions, i.e., it is a balanced block design (BBD) when each of {rows} and {columns} is considered as blocks.

The motivation for considering a GYD is that it is the most symmetric design. For a GYD, the C-matrix is completely symmetric in the sense that all the diagonal elements are the same and all the off-diagonals are the same. This is a crucial property in proving the optimality.

The present note is mainly concerned with the case $b_1 = b_2 = b$. In this situation, the C-matrix of a design d can be written as

$$C_d = \mathrm{diag}(r_{d1}, \cdots, r_{dv}) - b^{-1}(N_{d1}N'_{d1} + N_{d2}N'_{d2}) + b^{-2}[r_{di}r_{dj}]. \tag{1.2}$$

Received December 1978; revised October 1979.

[1] This research was prepared with the support of National Science Foundation Grant No. MCS 78-01422.

AMS 1970 *subject classifications.* Primary 62K05; secondary 62K10, 05B05.

Key words and phrases. Generalized Youden design, pseudo-Youden design, balanced block design, A-optimality, D-optimality, E-optimality.

Let $\tilde{N}_d = [N_{d1} \mid N_{d2}]$. Then

$$C_d = \operatorname{diag}(r_{d1}, \cdots, r_{dv}) - b^{-1}\tilde{N}_d\tilde{N}_d' + b^{-2}[r_{di}r_{dj}]. \tag{1.3}$$

It is clear that this matrix is completely symmetric if $\tilde{N}_d$ is the incidence matrix of a BBD. On the other hand, $\tilde{N}_d$ is in fact the incidence matrix of the block design obtained by combining the b rows and b columns of d together as blocks. So, *C_d is completely symmetric as long as this "combined design" is a* BBD; *d does not have to be a* GYD! This kind of design is termed a *pseudo-Youden design* (PYD). When $b_1 = b_2$, a GYD is also a PYD, but the converse is not true. Certainly, the construction of a PYD is much more flexible than that of a GYD. Thus, the stringent conditions imposed on the definition of a GYD are substantially relaxed. Furthermore, using the technique of Kiefer (1975a) one can show that a PYD is A- and E-optimal and is D-optimal if $v \neq 4$.

There is an analogous theory for higher-dimensional designs. 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 factor i. There are $b_1 b_2 \cdots b_n$ cells in this hyperrectangle coordinatized by the n-tuples of integers $(i_1, \cdots, i_n)$ with $1 \le i_j \le b_j$, $1 \le j \le n$. For each i with $1 \le i \le n$, the union of all the cells with the same ith coordinate is called a *hyperplane* (in *direction i*). Similar to the two-way setting, when $b_1 = b_2 = \cdots = b_n = b$, the C-matrix of a design d is completely symmetric as long as the block design obtained by considering all the nb hyperplanes of d as blocks is a BBD. Such a design is called an *n-dimensional pseudo-Youden design.* Similar to Cheng (1978), one can show that an n-dimensional PYD is E-optimal, and is A- and D-optimal if $b \ge 2n + 2$.

It is interesting to note that by the well-known Fisher's inequality on BIBD, a $b_1 \times b_2$ GYD exists only if at least one $b_i \ge v$. This is no longer true for PYD. There is a 6×6 PYD with 9 varieties. This is a big reduction in size from a 9×9 Latin square or any GYD accommodating 9 varieties. This example is presented in Section 3.

For convenience, the block design obtained by considering all the hyperplanes of a PYD d as blocks is called the *combined design* of d. A $b \times b \times \cdots \times b$ n-dimensional PYD with v varieties is abbreviated as PYD(n; v, b). A Youden hypercube defined in Cheng (1978) is abbreviated as YHC, and we denote a balanced block design with v varieties and b blocks of size k by BBD(v, b, k).

2. Construction of pseudo-Youden designs. The construction presented in this section is a patchwork method which goes back to Kiefer (1975b). Based on the fundamental patchwork result (Theorem 2.1), a family of PYD's are constructed via the use of finite geometries (Theorem 2.2 and Theorem 2.4).

We can establish the following analogue of Theorem 3.3 of Cheng (1979).

THEOREM 2.1. *Assume $b = t + c$, $v \mid c^n$ and v divides $\prod_{j=1; j\neq i}^{n} \alpha_j$ $(1 \le i \le n)$ where $\alpha_j = t$ or c but not all $\alpha_j = c$. Suppose that there exist a* BBD(v, nb, c^{n-1}) *d_1 and an n-dimensional $c \times c \times \cdots \times c$ array d_2 with v varieties such that all the hyperplanes of d_2 constitute nc blocks of d_1 and the remaining nt blocks of d_1 can be divided into n groups $B_1, B_2, \cdots, B_n$ each containing t blocks such that all the varieties are equally replicated within each group. Then there is a* PYD(n; v, b).

The proof is similar to that of Theorem 3.3 of Cheng (1979), and hence is omitted.

As an application of Theorem 2.1, we have the following

THEOREM 2.2. *Let s be a prime power and m be a positive integer such that $s \equiv 1 \pmod{n}$ and $m \equiv n \pmod{s}$. Then there is a* PYD (n; v, b) *with $v = s^n$ and $b = mn^{-1}[s^n + s^{n-1} + \cdots + s]$.*

PROOF. Let $c = s$, and $t = b - c = mn^{-1}[s^n + s^{n-1} + \cdots + s] - s$. The assumption "$s \equiv 1$

(mod n)" implies that

$$n \mid s^{n-1} + s^{n-2} + \cdots + 1. \tag{2.1}$$

So b is an integer and $s \mid b$. Write $m = es + n$. Then

$$\begin{aligned} t &= n^{-1}(es + n)[s^n + s^{n-1} + \cdots + s] - s \\ &= n^{-1}e[s^{n+1} + s^n + \cdots + s^2] + s^n + s^{n-1} + \cdots + s^2 \\ &= es^2 n^{-1}[s^{n-1} + s^{n-2} + \cdots + 1] + s^n + s^{n-1} + \cdots + s^2. \end{aligned} \tag{2.2}$$

It follows from (2.1) and (2.2) that

$$s^2 \mid t, \tag{2.3}$$

and therefore v divides $\prod_{j=1; j\neq i}^{n} \alpha_j$ $(1 \le i \le n)$ where $\alpha_j = t$ or c but not all $\alpha_j = \text{c}$.

Let EG(n; s) be the n-dimensional Euclidean geometry based on the Galois field GF(s). Then each $(n-1)$-flat contains s^{n-1} points, and there are $s^n + s^{n-1} + \cdots + s$ $(n-1)$-flats. If we consider each $(n-1)$-flat as a block, and the s^n points in EG(n; s) as the varieties, then we get a BIBD(s^n, $s^n + s^{n-1} + \cdots + s$, s^{n-1}). Let d_1 be m copies of this BIBD, then d_1 is a BIBD(v, nb, c^{n-1}). It is clear that we can arrange the s^n points in EG(n; s) into a hyperrectangle of size $s \times s \times \cdots \times s$ such that each of the hyperplane of this hyperrectangle is an $(n-1)$-flat of EG(n; s). This gives the design d_2 in Theorem 2.1. The remaining conditions in Theorem 2.1 follow from (2.3) and the fact that each pencil of $(n-1)$-flats in EG(n; s) contains s $(n-1)$-flats. □

Under the conditions of Theorem 2.2, suppose that there exists a YHC with the same parameters as the PYD constructed there. Then there must exist a BIBD(v, b, k) with $v = s^n$, $b = mn^{-1}(s^n + s^{n-1} + \cdots + s)$, and $k = s^{n-1}$, and the number of replications of each variety is

$$r = mn^{-1}[s^{n-1} + s^{n-2} + \cdots + 1].$$

Then

$$\begin{aligned} \lambda &= r(k-1)/(v-1) \\ &= mn^{-1}[s^{n-1} + s^{n-2} + \cdots + 1](s^{n-1} - 1)/(s^n - 1) \\ &= mn^{-1}(s^{n-2} + \cdots + 1) \text{ must be an integer.} \end{aligned}$$

Write $m = es + n$, then

$$\lambda = s^{n-2} + \cdots + 1 + n^{-1}e(s^{n-1} + \cdots + s).$$

So $n^{-1}e(s^{n-1} + \cdots + s)$ is an integer, i.e., $n \mid e(s^{n-1} + \cdots + s)$.

Now, the assumption "$s \equiv 1 \pmod{n}$" implies that $s^{n-1} + \cdots + s \equiv n - 1 \pmod{n}$. Accordingly, $n \mid e(n-1)$. But since n and $n-1$ are relatively prime, we must have $n \mid e$. Thus, $n \mid m$.

It is clear that if $n \mid m$, then a GYD is constructible.

In summary, we have the following

THEOREM 2.3. *Under the same conditions as in Theorem 2.2, a Youden hypercube (or a generalized Youden design if $n = 2$) is constructible if and only if $n \mid m$.*

So there are lots of parameter values for which a PYD can be constructed while a YHC (or a GYD) does not exist. The conditions put on a YHC are really too stringent. Even if there exists a YHC, there is a lot of freedom to rearrange the varieties so that the structure of a YHC is destroyed, but the resulting design is still optimal. A GYD and a PYD with the same size have the same C-matrix and hence have the same performance for the elimination of two-way heterogeneity. However, if in fact there is only one direction of heterogeneity, then certainly a GYD is a better design.

The following two 6 × 6 designs are a GYD and a PYD which is not a GYD:

1	2	3	4	1	4
2	3	4	1	4	2
3	4	1	2	2	3
4	1	2	3	3	1
1	3	2	4	1	3
2	4	3	1	2	4

a non-GYD PYD

1	2	3	4	1	3
2	3	4	1	4	2
3	4	1	2	3	4
4	1	2	3	2	1
1	2	3	4	1	4
2	3	4	1	3	2

a GYD

Ash (1977) constructed GYD's for essentially all practical parameter values. There is one case she could not resolve, i.e., it is unknown whether there is a 40 × 40 GYD with $v = 25$. However, later on, using the idea of pseudo-Youden designs, she was able to construct a 40 × 40 PYD with $v = 25$. This provides us with an optimal design which is as good as the missing (or nonexisting) GYD.

Similar to Theorem 3.4 of Cheng (1979), we have the following

THEOREM 2.4. *Let s be a prime power and m, l, d be positive integers such that $d \equiv 1 \pmod{n}$, $s \equiv 1 \pmod{n}$, $m \equiv n \pmod{s}$, and $l - 1 = (n - 1)d$. Then there is a* PYD$(n; v, b)$ *with $v = s^l$, and $b = mn^{-1}[s^d(s^l - 1)/(s - 1)]$.*

PROOF. Let $c = s^d$ and $t = b - c = s^d[mn^{-1}(s^l - 1)/(s - 1) - 1]$. Then similar to the proof of Theorem 2.2, it can be shown that $n \mid (s^{l-1} + \cdots + s + 1)$ and $s \mid [mn^{-1}(s^l - 1)/(s - 1) - 1]$. Thus b is an integer and v divides $\prod_{j=1; j \neq i}^{n} \alpha_j$ $(1 \leq i \leq n)$ where $\alpha_j = t$ or c but not all $\alpha_j = c$. So it suffices to show the existence of the two designs d_1 and d_2 in Theorem 2.2.

Let $\bar{d}$ be the BIBD of all $(l - 1)$-flats in the l-dimensional Euclidean geometry with s points per line. Then $\bar{d}$ is a BIBD$(s^l, s(s^l - 1)/(s - 1), s^{l-1})$, i.e., a BIBD$(s^l, (ms^{d-1})^{-1}nb, c^{n-1})$. We can take d_1 to be ms^{d-1} copies of $\bar{d}$.

Let $P_1, P_2, \cdots, P_n$ be n independent pencils of $(l - 1)$-flats in EG(l, s). By the construction in the proof of Theorem 3.4 of Cheng (1979), there exists a $c \times c \cdots \times c$ n-dimensional hypercube in which each of the $(l - 1)$-flats in P_i appears s^{d-1} times as hyperplanes in direction i, $\forall i$ with $1 \leq i \leq n$. This provides an eligible candidate for d_2. □

Note that Theorem 2.2 is a special case of Theorem 2.4 with $d = 1$. Again, under the same conditions as in Theorem 2.4, a Youden hypercube (or a generalized Youden design if $n = 2$) is constructible if and only if $n \mid m$.

3. A 6 × 6 PYD with 9 varieties. All the PYD's constructed in the last section have $b > v$. Thus, the corresponding combined designs are BBD, not BIBD. The following is a 6 × 6 PYD with 9 varieties. Thus $b < v$, and the combined design is indeed a BIBD.

(3.1)

4	7	8	6	9	5
3	1	2	8	7	9
2	5	1	3	6	4
9	3	6	2	5	8
7	6	9	4	1	3
5	8	4	7	2	1

This PYD is D-, A-, and E-optimal. If we stick to generalized Youden designs, we are unable to construct such a small design which has completely symmetric C-matrix and is optimal! A similar design was also reported by Shah (1977) and Kshirsagar (1957).

REFERENCES

[1] ASH, A. S. (1977). Construction of generalized Youden designs. Thesis, Univ. Illinois at Chicago Circle.

[2] CHENG, C. S. (1978). Optimal designs for the elimination of multi-way heterogeneity. *Ann. Statist.* **6** 1262–1272.

[3] CHENG, C. S. (1979). Construction of Youden hyperrectangles. *J. Statist. Plann. Inference* **3** 109–118.

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

[5] KIEFER, J. (1975b). Balanced block designs and generalized Youden designs. I. Construction (patchwork). *Ann. Statist.* **3** 109–118.

[6] KSHIRSAGAR, A. M. (1957). On balancing in designs in which heterogeneity is eliminated in two directions. *Calcutta Statist. Assoc. Bull.* **7** 469–476.

[7] SHAH, K. R. (1977). Analysis of designs with two-way elimination of heterogeneity. *J. Statist. Plann. Inference* **1** 207–216.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720