The Annals of Mathematical Statistic 5
1970, Vol. 41, No. 5, 1655–1664

THE SEQUENTIAL GENERATION OF *D*-OPTIMUM EXPERIMENTAL DESIGNS[1]

BY HENRY P. WYNN

Imperial College, London

0. Summary. It is possible to obtain convergence to a D-optimum measure, as defined by Kiefer and Wolfowitz, by successively adding points to a given initial experimental design. The points added correspond to points of maximum variance of the usual least squares estimate of the response mean for the particular regression model at each stage. A new bound is given for the generalized variances involved and an example is worked out.

1. Introduction. Difficulties may arise in choosing an experimental design which is optimum in some sense in situations in which a linear regression model is assumed. A procedure is described which may help in overcoming at least two of these difficulties. The first difficulty arises when the model, or the design space, is sufficiently complicated to prevent the immediate calculation of an optimum design. The second difficulty occurs when some observations, not necessarily forming part of an optimum design, have already been taken and it is required to take further observations in an optimum way, possibly including a change of model. Both difficulties may of course arise together.

The procedure prescribes the sequential addition of points to a given initial design in such a way that an optimum design is approached. This enables one either to compute an optimum given an arbitrary initial design or to choose extra points for a partially completed experiment. The extra set of observations, though not necessarily giving an optimum improvement itself, usually brings the design closer to optimality. It is also possible to calculate bounds which indicate how close one is to an optimum design at any stage without knowing the true optimum design but only the improvements up to that stage. The procedure and bounds are valid for general linear regression models.

The basic criterion of design optimality which we shall use here is that of D-optimality developed largely by Kiefer (1959, 1961a, 1961b, 1962a, 1962b) and Kiefer and Wolfowitz (1959, 1960). More recently some of the results in these papers were proved again and extended by Karlin and Studden (1966a, Chapter X; 1966b). D-optimum designs have been calculated for many different models; see also Hoel (1958, 1965). However, difficulties like those mentioned above still arise. Furthermore, the general design measure introduced by Kiefer and Wolfowitz which enables analytical results to be obtained does not always give optimum discrete designs which can be used in experimental work. But good approximations for large designs are possible. Thus, while relying heavily on previous work in this

Received October 13, 1969.
[1] This research was supported by the Science Research Council (U.K.).

field, we shall demonstrate further connections between the theoretical results and procedures which can be used in practice.

Since this paper was first submitted similar and extensive work in the field by V. V. Federov (Preprint No. 7, Department of Statistical Methods, Moscow State University) has come to the author's notice.

2. Definitions and background. Following Kiefer and Wolfowitz, we suppose that $f_1, \cdots, f_k$ are k given linearly independent functions on a space $\mathscr{X}$ which are continuous in a topology in which $\mathscr{X}$ is compact. The space $\mathscr{X}$ will usually be a closed compact set in a Euclidean space of a particular dimension. We assume that at each point x in $\mathscr{X}$ a random variable Y_x is defined and is such that

$$E(Y_x) = \theta' f(x),$$

where $f(x)$ is the $k \times 1$ column vector of functions f_i evaluated at x and θ is a $k \times 1$ column vector of unknown real parameters. We assume also that $\text{Var}(Y_x) = \sigma^2$, $\text{Cov}(Y_{x_1}, Y_{x_2}) = 0$ for x, x_1, x_2 in $\mathscr{X}(x_1 \neq x_2)$.

In defining experimental designs, it is important to distinguish carefully between *discrete* designs and design *measures*. A discrete design with n points is a set of n points $x_1, \cdots, x_n$, in $\mathscr{X}$, not necessarily distinct. A discrete design will be denoted by D_n, where the subscript always denotes the number of points in the design. A design measure, referred to in future merely as a measure, is a probability measure, denoted by ξ, on $\mathscr{X}$. Specifically, ξ is a member of the set, $\mathscr{D}$, of all measures defined on the Borel Field, $\mathscr{B}$, generated by the open sets of $\mathscr{X}$ and such that

$$\int_{\mathscr{X}} \xi(dx) = 1.$$

It is assumed that $\mathscr{B}$ contains all one-point sets.

For the discrete design D_n in $\mathscr{X}$ we have the corresponding $n \times k$ design matrix X_n, whose ith row is the vector $f(x_i)$. The design matrix corresponding to a discrete design will always be given the same subscript. For a measure ξ on $\mathscr{X}$ we write

$$m_{ij}(\xi) = \int_{\mathscr{X}} f_i(x) f_j(x) \xi(dx).$$

Let $M(\xi)$ be the $k \times k$ matrix whose i, jth entry is $m_{ij}(\xi)$.

Now from any discrete design D_n, we can form a measure, ξ_n, by attaching a mass of $1/n$ to each point of D_n. We call ξ_n the associated measure of D_n. Again, a discrete design and its associated measure will take the same subscripts. We have, therefore, from these definitions

$$X_n' X_n = nM(\xi_n). \tag{2.1}$$

A measure ξ^* is called D-optimum if

$$\det\{M(\xi^*)\} = \sup_{\xi \in \mathscr{D}} \det\{M(\xi)\}. \tag{2.2}$$

Kiefer and Wolfowitz have investigated the equivalence of several types of optimality. Defining

$$d(x, \xi) = f(x)' M^{-1}(\xi) f(x), \tag{2.3}$$

they show (Kiefer and Wolfowitz, 1960) that D-optimality of ξ^* as defined by (2.2) is equivalent to

$$\inf_{\xi\in\mathscr{D}} \sup_{x\in\mathscr{X}} d(x,\xi) = \sup_{x\in\mathscr{X}} d(x,\xi^*) \tag{2.4}$$

and also to

$$\sup_{x\in\mathscr{X}} d(x,\xi^*) = k. \tag{2.5}$$

Recall that k is the dimension of f.

A discussion of the statistical meaning of D-optimality is given by Kiefer (1959). First, $\det(X_n'X_n)^{-1}$ is proportional to the generalized variance of the parameters forming θ. Also, suppose that the Y_x are normally distributed and therefore independent. Maximizing $\det(X_n'X_n)$ is then equivalent to maximizing the Gaussian curvature of the power surface of the usual F test, or χ^2 test if σ^2 is known, at the null hypothesis $\theta = \theta_0$, where θ_0 is some fixed value of θ.

From (2.1), we have

$$d(x,\xi_n) = nf(x)'(X_n'X_n)^{-1}f(x) \tag{2.6}$$

and $\sigma^2 f(x)'(X_n'X_n)^{-1}f(x)$ is the variance of the usual least squares estimate of the expected response, $E(Y_x)$, at x. Thus the equivalence of (2.2) and (2.4) means that if D_n is such that ξ_n is D-optimum then the design minimizes the maximum over $\mathscr{X}$ of the variance function derived from using that design.

It should be mentioned (see, for proof, Karlin and Studden 1966a, page 323) that given any measure ξ we can find another measure ξ' which has a finite support comprising fewer than $\frac{1}{2}k(k+1)+2$ points such that $M(\xi) = M(\xi')$. If ξ^* has a finite support $x_1, \cdots, x_n$, then another condition which can easily be seen to be equivalent to (2.5), and therefore to (2.2) and (2.4) is that $\sup d(x, \xi^*)$ is achieved at $x_1, \cdots, x_n$. Note that the fact that ξ^* possesses finite support does not necessarily mean that it is the associated measure of some discrete design since the masses at the support points may well be irrational. It is the coincidence of support points and points of maximum variance for D-optimum measures that leads intuitively to our procedure.

3. The procedure. Let D_{n_0} be a discrete design with n_0 points, $x_1, \cdots, x_{n_0}$ which is admissible in the sense that $X_{n_0}'X_{n_0}$ is non-singular. From D_{n_0}, by successive addition of points, we shall generate a sequence of designs such that in the limit the associated measures become D-optimum. Thus, we first find a point x_{n_0+1} in $\mathscr{X}$ which maximizes the variance function obtained by using D_{n_0}; that is choose x_{n_0+1} such that

$$\sup_{x\in\mathscr{X}} d(x,\xi_{n_0}) = d(x_{n_0+1},\xi_{n_0}). \tag{3.1}$$

Then form a new design D_{n_0+1}, with n_0+1 points by adding X_{n_0+1} to D_{n_0} and continue this process to obtain a sequence of designs, $D_{n_0} \subset D_{n_0+1} \subset \cdots \subset D_n \subset \cdots$, where D_n is obtained from D_{n-1} by adding a point of maximum variance, over $\mathscr{X}$, of the estimated response mean obtained from using D_{n-1}. The basic result of this paper concerns the sequence of associated measures $\{\xi_n\}_{n_0}^{\infty}$ and is contained in

THEOREM 1. *As* $n \to \infty$, $\lim \det\{M(\xi_n)\} = \det\{M(\xi^*)\}$, *where* ξ^* *is a D-optimum measure.*

The following simple identity in matrix algebra which has been referred to by several authors, for example Scheffé (1959, page 417), is important in what follows. Its proof is simple and will be omitted.

LEMMA. *If* W *is a symmetric positive definite* $k \times k$ *matrix and* $\underset{\sim}{a}$ *a* $k \times 1$ *column vector, then*

$$\frac{\det(W+\underset{\sim}{a}\underset{\sim}{a}')}{\det(W)} - 1 = \underset{\sim}{a}'W^{-1}\underset{\sim}{a}. \tag{3.2}$$

Now consider the sequence of X_n corresponding to our constructed sequence of D_n. If we add any point x in $\mathscr{X}$ to D_n to get D_{n+1}, we must add a corresponding row, $f(x)$, to X_n. Thus

$$X'_{n+1}X_{n+1} = X_n'X_n + f(x)f(x)'.$$

Putting $W = X_n'X_n$ and $\underset{\sim}{a} = f(x)$ in (3.2) gives

$$\frac{\det(X'_{n+1}X_{n+1})}{\det(X_n'X_n)} - 1 = f(x)'(X_n'X_n)^{-1}f(x). \tag{3.3}$$

This important relationship implies that if we choose x to maximize $\det(X'_{n+1}X_{n+1})$, we also maximize the variance function, which is proportional to the right-hand side of (3.3). Theorem 1, therefore, implies that we can approach a D-optimum measure by maximizing $\det(X_n'X_n)$ one point at a time, or if $\det(X_n^{+\prime}X_n^{+})$ is the maximum possible over all discrete designs with n points, then

$$\frac{\det(X_n'X_n)}{\det(X_n^{+\prime}X_n^{+})} \to 1, \quad \text{as} \quad n \to \infty.$$

Using (2.1) and putting $x = x_{n+1}$ in (3.3), we have

$$n\left[\left(\frac{n+1}{n}\right)^k \cdot \frac{\det\{M(\xi_{n+1})\}}{\det\{M(\xi_n)\}} - 1\right] = \bar{d}(\xi_n), \tag{3.4}$$

where $\bar{d}(\xi) = \sup_{x \in \mathscr{X}} d(x, \xi)$. Now Kiefer (1961b) gives an inequality connecting the nearness of $\bar{d}(\xi)$ to k with the nearness of $\det\{M(\xi)\}$ to $\det\{M(\xi^*)\}$, the optimum. The inequality is

$$\frac{\det\{M(\xi)\}}{\det\{M(\xi^*)\}} \geqq \exp\{k - \bar{d}(\xi)\}. \tag{3.5}$$

Putting $\xi_n = \xi$ in (3.5) and substituting for $\bar{d}(\xi_n)$ from (3.4), we have, after rearrangement,

$$\frac{\det\{M(\xi_{n+1})\}}{\det\{M(\xi_n)\}} \geqq \left(\frac{n}{n+1}\right)^k \left\{1 + \frac{1}{n}\left(k + \log\left[\frac{\det\{M(\xi^*)\}}{\det\{M(\xi_n)\}}\right]\right)\right\}. \tag{3.6}$$

It is this basic inequality which we shall use to prove Theorem 1.

Since $0 < \det\{M(\xi_n)\} \leqq \det\{M(\xi^*)\}$ for all n, to show that $\det\{M(\xi_n)\} \to \det\{M(\xi^*)\}$, we must show that given an $\varepsilon > 0$ there is an n^* such that

$$\det\{M(\xi_n)\} > (1-\varepsilon)\det\{M(\xi^*)\} \quad \text{for all} \quad n \geqq n^*.$$

Divide the sequence $\{\xi_n\}_{n_0}^{\infty}$ into two disjoint subsequences Ξ_1 and Ξ_2 such that

$\xi_n \in \Xi_1$ if and only if $\det\{M(\xi_n)\} \geqq (1-\frac{1}{2}\varepsilon)\det\{M(\xi^*)\}$,

$\xi_n \in \Xi_2$ if and only if $\det\{M(\xi_n)\} < (1-\frac{1}{2}\varepsilon)\det\{M(\xi^*)\}$.

We show first that Ξ_1 is non-empty; that is that we can, at least, find $\det\{M(\xi_n)\}$ as near as we please to $\det\{M(\xi^*)\}$.

Suppose Ξ_1 were empty. Put $\delta = \log(1/(1-\frac{1}{2}\varepsilon))$ and use (3.6) to obtain

$$\frac{\det\{M(\xi_{n+1})\}}{\det\{M(\xi_n)\}} > \left(\frac{n}{n+1}\right)^k\left\{1+\frac{1}{n}(k+\delta)\right\}, \tag{3.7}$$

since then

$$\frac{\det\{M(\xi^*)\}}{\det\{M(\xi_n)\}} > \frac{1}{1-\frac{1}{2}\varepsilon}$$

for all $n(\geqq n_0)$. Expanding the right-hand side of (3.7) in powers of $1/n$ we have

$$\frac{\det\{M(\xi_{n+1})\}}{\det\{M(\xi_n)\}} > 1+\frac{\delta}{n}+\frac{a_2}{n^2}+\frac{a_3}{n^3}+\ \cdots,$$

where the coefficients $a_2, a_3 \cdots$ depend only on k and δ. Thus we can find an $\eta > 0$ and an n^+ such that

$$\frac{\det\{M(\xi_{n+1})\}}{\det\{M(\xi_n)\}} \geqq 1+\frac{\eta}{n} \quad \text{for all} \quad n > n^+.$$

Choose an $n > n^+$, say $n = m$. Then, for any integer $r > 0$, we must have

$$\det\{M(\xi_{m+r})\} \geqq \det\{M(\xi_m)\} \times \textstyle\prod_{n=m}^{r-1}(1+\eta/n).$$

Thus, since $\prod_{n=m}^{r-1}(1+\eta/n) > 1+\sum_{n=m}^{r-1}(\eta/n)$, $\det\{M(\xi_{m+r})\} \to \infty$ as $r \to \infty$. But $\det\{M(\xi_m)\} > 0$ and we know that $\det\{M(\xi^*)\}$ is finite under our assumptions of a compact $\mathscr{X}$ and continuous f_i. Thus we have a contradiction and Ξ_1 must be non-empty. Moreover, since the choice of ε was arbitrary, Ξ_1 must contain an infinity of ξ_n.

For $\xi_{n_1} \in \Xi_1$ we have, from (3.6)

$$\frac{\det\{M(\xi_{n_1+1})\}}{\det\{M(\xi_{n_1})\}} \geqq \left(\frac{n_1}{n_1+1}\right)^k\left(1+\frac{k}{n_1}\right), \tag{3.8}$$

since certainly

$$\log\left[\frac{\det\{M(\xi^*)\}}{\det\{M(\xi_1)\}}\right] \geqq 0.$$

Expanding the right-hand side of (3.8) in powers of $1/n_1$, we obtain

$$\frac{\det\{M(\xi_{n_1+1})\}}{\det\{M(\xi_{n_1})\}} \geqq 1+\frac{b_2}{n_1{}^2}+\frac{b_3}{n_1{}^3}+\ldots,$$

where $b_2, b_3, \cdots$ depend only on k. Therefore, we can find an n' such that for all $n_1 > n'$ and $\xi_{n_1} \in \Xi_1$

$$\det\{M(\xi_{n_1+1})\} \geqq (1-\tfrac{1}{2}\varepsilon)\det\{M(\xi_{n_1})\}. \tag{3.9}$$

This implies that for $\xi_{n_1} \in \Xi_1$ and $n_1 > n'$, even if ξ_{n_1+1} is not in Ξ_1, we must still have

$$\det\{M(\xi_{n_1+1})\} \geqq (1-\tfrac{1}{2}\varepsilon)(1-\tfrac{1}{2}\varepsilon)\det\{M(\xi^*)\} > (1-\varepsilon)\det\{M(\xi^*)\}. \tag{3.10}$$

Furthermore, from (3.7), we can find an n'' such that for all $\xi_{n_2} \in \Xi_2$ with $n_2 > n''$

$$\frac{\det\{M(\xi_{n_2+1})\}}{\det\{M(\xi_{n_2})\}} > 1. \tag{3.11}$$

Now select n^* so that $\xi_{n^*} \in \Xi_1$, $n^* > n'$ and $n^* > n''$, then the two inequalities (3.10) and (3.11) compel any ξ_n with $n \geqq n^*$ to satisfy

$$\det\{M(\xi_n)\} > (1-\varepsilon)\det\{M(\xi^*)\},$$

where $\xi_n \in \Xi_1$ or Ξ_2. Thus, the theorem is proved.

4. Bounds. Kiefer (1961b, Theorem 2.34) proves, in addition to (3.5), that

$$\frac{\det\{M(\xi^*)\}}{\det\{M(\xi)\}} \geqq \exp\frac{\{\bar{d}(\xi)-k\}^2}{2k(k+1)}$$

provided $\bar{d}(\xi)-k \leqq 1$.

We now derive an inequality, which seems to be better than the above and is valid for all $\bar{d}(\xi)$.

If ξ is a measure for which $M(\xi)$ is non-singular, x a point in $\mathscr{X}$ and α a number $(0 < \alpha < 1)$, we can use the lemma to obtain

$$\frac{\det[M(\xi)+\{(1-\alpha)/\alpha\}f(x)f(x)']}{\det\{M(\xi)\}}-1 = \left(\frac{1-\alpha}{\alpha}\right)d(x,\xi). \tag{4.1}$$

But $\alpha M(\xi)+(1-\alpha)f(x)f(x)' = M\{\alpha\xi+(1-\alpha)\xi'\}$, where ξ' is the measure attaching mass unity to the point x in $\mathscr{X}$. Put $\alpha\xi+(1-\alpha)\xi' = \xi''$. Then (4.1) becomes

$$\frac{\det\{M(\xi'')\}}{\det\{\alpha M(\xi)\}}-1 = \left(\frac{1-\alpha}{\alpha}\right)d(x,\xi).$$

Choose x so that $d(x, \xi) = \bar{d}(\xi)$. We must have

$$\left(\frac{1-\alpha}{\alpha}\right)\bar{d}(\xi) = \frac{\det\{M(\xi'')\}}{\det\{\alpha M(\xi)\}}-1$$

$$\leqq \frac{\det\{M(\xi^*)\}}{\alpha^k\det\{M(\xi)\}}-1,$$

since $\det\{M(\xi'')\} \leqq \det\{M(\xi^*)\}$, the optimum. We can make this inequality as strong as possible by noticing that since α is arbitrary in (0, 1)

$$\bar{d}(\xi) \leqq \inf_{0 \leqq \alpha \leqq 1} \left(\frac{\alpha}{1-\alpha}\right)\left(\frac{\beta}{\alpha^k}-1\right), \tag{4.2}$$

where

$$\beta = \frac{\det\{M(\xi^*)\}}{\det\{M(\xi)\}}.$$

To find the value of the right-hand side of (4.2) we differentiate $(1-\alpha)^{-1}\alpha(\beta/\alpha^k-1)$ with respect to α and put the result equal to zero to obtain

$$\alpha^k - k\beta\alpha + \beta(k-1) = 0. \tag{4.3}$$

Since $\beta \geqq 1$, this has a unique solution in (0, 1) and gives the required minimum. Rather than solve this kth order polynomial equation for α, we let

$$\inf_{0 \leqq \alpha \leqq 1} \left(\frac{\alpha}{1-\alpha}\right)\left(\frac{\beta}{\alpha^k}-1\right) = d.$$

Let the solution of (4.3) be α^*. Then $\alpha^{*k} = \beta(k\alpha^* - k + 1)$ and

$$d = \frac{\alpha^*}{1-\alpha^*}\left(\frac{1}{k\alpha^*-k+1}-1\right)$$

eliminating β. Thus,

$$d = \frac{k\alpha^*}{k\alpha^*-k+1}$$ giving

$$\alpha^* = \frac{d(k-1)}{k(d-1)}. \tag{4.4}$$ But

$$\beta = \frac{\alpha^{*k}}{k\alpha^*-k+1};$$

thus, substituting for α^* from (4.4) we have

$$\beta = \left(\frac{d}{k}\right)^k \left(\frac{k-1}{d-1}\right)^{k-1}.$$

This gives an inequality explicit for β, rather than $(\bar{d}\xi)$, which takes the form,

$$\beta = \frac{\det\{M(\xi^*)\}}{\det\{M(\xi)\}} \geqq \left\{\frac{\bar{d}(\xi)}{k}\right\}^k \left\{\frac{k-1}{\bar{d}(\xi)-1}\right\}^{k-1}. \tag{4.5}$$

The bound (4.5) is valid for all $\bar{d}(\xi)$ whereas Kiefer's holds for $\bar{d}(\xi) \leqq k+1$. Furthermore, a simple calculation involving the derivatives of the two bounds shows that (4.5) is strictly better at least in the neighbourhood of $\bar{d}(\xi) = k$.

When $\beta = 1$ in (4.2) we have $\bar{d}(\xi) \leqq k$. It is easy to show that $\bar{d}(\xi) \geqq k$. Thus, in this case $\bar{d}(\xi) = k$ and we must have equality in (4.5). This provides a proof that

(2.2) implies (2.5) which is an alternative to the proof given by Kiefer and Wolfowitz (1960). The inequality (4.5) is interesting in its own right, especially as it holds for any linear model for which we can find non-singular $M(\xi)$. In order to get an inequality explicit for $\bar{d}(\xi)$, we would need to plot the value of the right-hand side of (4.5) against $\bar{d}(\xi)$ and then read off the correct upper bound for $\bar{d}(\xi)$ given β. Our application of (4.5) is to the procedure described in Section 3.

Together with (3.5), we use (4.5) to give bounds for $\det\{M(\xi^*)\}$ when we know both $\det\{M(\xi_n)\}$ and $\det\{M(\xi_{n+1})\}$. These take the form $A_n \leqq \det\{M(\xi^*)\} \leqq B_n$, where

$$A_n = \det\{M(\xi_n)\}\left\{\frac{\bar{d}(\xi_n)}{k}\right\}^k \left\{\frac{k-1}{\bar{d}(\xi_n)-1}\right\}^{k-1},$$

$$B_n = \det\{M(\xi_n)\}\exp\{\bar{d}(\xi_n)-k\}.$$

Here $\bar{d}(\xi_n)$ is given by (3.4). We can combine the bounds for r steps of the procedure and state that

$$\max_{n_0 \leqq n \leqq n_0+r}\{A_n\} \leqq \det\{M(\xi^*)\} \leqq \min_{n_0 \leqq n \leqq n_0+r}\{B_n\}. \tag{4.6}$$

5. Applications. In considering any specific example some general points should be borne in mind. First, instead of dealing with the design space, $\mathscr{X}$, and the model separately it is natural to consider the space $f(\mathscr{X})$ of all k-dimensional vectors which, given f and $\mathscr{X}$, can be chosen as a row of a design matrix X. Also we can always find a D-optimum measure such that the image in $f(\mathscr{X})$ of its support in $\mathscr{X}$ lies on the boundary of $f(\mathscr{X})$. It can easily be shown, furthermore, that D-optimality is preserved under non-singular linear transformations of $f(\mathscr{X})$; that is transformations for which the model can be re-written in the same form by taking a non-singular linear transformation of the parameters.

Another important point is that if we are using the procedure of Section 3, and a D-optimum measure with finite support, or merely the support itself, is known, then maximizing the variance function over the support at each stage will be considerably easier computationally than maximizing it over the whole space $\mathscr{X}$.

We give an example to illustrate some of these ideas. The model chosen is simple but the design space is a little unusual.

Consider the two-dimensional polynomial model

$$E(Y_x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2, \qquad x = (x_1, x_2).$$

The usual regression assumptions outlined in Section 2 are made. Let the design space be the area whose boundary is the quadrilateral with vertices $A(2, 2)$, $B(-1, 1)$, $C(1, -1)$ and $D(-1, -1)$. We can say immediately that it is impossible to transform this by a non-singular linear transformation to, say, a rectangle for which the D-optimum measure is known to consist of masses $\frac{1}{4}$ at each vertex. However, the support points of the D-optimum measure are still A, B, C and D, for the variance function is a maximum at at least one of these points. This follows from the convexity, in $f(\mathscr{X})$, of the variance function and the polygonal nature of the boundary of $f(\mathscr{X})$. Thus, in generating successive points, we restrict ourselves to the points A, B, C and D.

Table 1 gives the results for the first few steps using B, C and D as the initial design. For this initial design $\det\{M(\xi_3)\} = 0.5926$, $\bar{d}(\xi_3) = 25.5$, $A_3 = 2.4252$ and $B_3 \doteqdot 5.9 \times 10^9$. This example is interesting in that the design with 32 points gives

TABLE 1
Points generated up to $n = 12$ *with* $\det\{M(\xi_n)\}$, $\bar{d}(\xi_n)$, A_n *and* B_n

n	Points added to form D_n	$\det\{M(\xi_n)\}$	$\bar{d}(\xi_n)$	A_n	B_n
3	B, C, D	0.5926	25.5000	2.4252	5.9×10^9
4	A	2.3750	3.5790	2.4252	4.2374
5	A	2.3040	3.7500	2.3802	4.8776
6	B	2.3333	4.2857	2.5205	8.4403
7	C	2.5190	3.2407	2.5297	3.2046
8	A	2.4688	3.3165	2.4863	3.3878
9	B	2.4527	3.6846	2.5220	4.8634
10	C	2.5200	3.3429	2.5240	2.9076
11	A	2.4883	3.3478	2.5094	3.5235
12	D	2.5000	3.2000	2.5075	3.0535

an exactly D-optimum measure; $\bar{d}(\xi_{32}) = 3$, $\det\{M(\xi_{32})\} = A_{32} = B_{32} = 2.53125$. The complete sequence of points, starting with BCD from the initial design, up to 32 points is

$$B\,C\,D\,A\,A\,B\,C\,A\,B\,C\,A\,D\,B\,C\,A\,B\,C\,A\,D\,B\,C\,A\,C\,B\,A\,D\,C\,B\,A\,B\,A\,C.$$

The D-optimum measure can be found by counting up the number of times each point is selected in the above sequence. It places masses of 10/32 at A, 9/32 at B and C, and 4/32 at D. Thus the procedure reaches a D-optimum measure in the minimum possible number of points. It also returns, it was found, to the D-optimum measure at multiples of 32 points, though after selecting the points in a somewhat different order. It is interesting also that A_3 and subsequent lower bounds are fairly accurate, more accurate than the corresponding upper bounds. In Table 1 the best bounds for $\det\{M(\xi^*)\}$ up to $n = 12$, as given by (4.6), are $A_7 = 2.5297$ and $B_{10} = 2.9076$.

Since for some steps there was a choice of alternative points which maximized the variance function the particular sequence given above is not unique.

Acknowledgments. The author is grateful to the referee for his useful comments including mentioning the local superiority of (4.5) over Kiefer's bound in Section 2. Also he wishes to express his thanks to Professor D. R. Cox and Dr. Agnes M. Herzberg for their assistance and encouragement.

REFERENCES

[1] HOEL, P. G. (1958). Efficiency problems in polynomial estimation. *Ann. Math. Statist.* **29** 1134–1145.

[2] HOEL, P. G. (1965). Minimax designs in two-dimensional regression. *Ann. Math. Statist.* **36** 1097–1106.

[3] KARLIN, S. and STUDDEN, W. J. (1966a). *Tchebyscheff Systems: with Applications in Analysis and Statistics.* Interscience, New York.

[4] KARLIN, S. and STUDDEN, W. J. (1966b). Optimal experimental designs. *Ann. Math. Statist.* **37** 783–815.

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

[6] KIEFER, J. (1961a). Optimum designs in regression problems, II. *Ann. Math. Statist.* **32** 298–325.

[7] KIEFER, J. (1961b). Optimum experimental designs V, with applications to systematic and rotable designs. *Proc. Fourth Berkeley Symp. Vol.* I, Univ. of California Press, 381–405.

[8] KIEFER, J. (1962a). An extremum result. *Canad. J. Math.* **14** 597–601.

[9] KIEFER, J. (1962b). Two more criteria equivalent to the *D*-optimality of designs. *Ann. Math. Statist.* **33** 792–796.

[10] KIEFER, J. and WOLFOWITZ, J. (1959). Optimum designs in regression problems. *Ann. Math. Statist.* **30** 271–294.

[11] KIEFER, J. and WOLFOWITZ, J. (1960). The equivalent of two extremum problems. *Canad. J. Math.* **12** 363–366.

[12] SCHEFFÉ, H. (1959). *The Analysis of Variance.* Wiley, New York.