Jour. Ind. Soc. Ag. Statistics
49 (Golden Jub. No.), 1996-97 : 167–176

D-Optimal Design Measures for Parallel Line Assays with Application to Exact Designs

Rahul Mukerjee
Indian Institute of Management, Calcutta, India

SUMMARY

D-optimal design measures are derived for symmetric parallel line assays. These lead to exact designs, with very high D-efficiency, both in the absence and presence of blocks and entail gain in D- efficiency compared to the more conventional equireplicate designs.

Key words :. Approximate theory, Efficiency, Non-equireplicate design.

1. *Introduction*

The efficiency designing of parallel line assays, taking due care of the contrasts of interest has received considerable attention in the literature; see Finney [2] for early results and an elegant discussion and Gupta and Mukerjee [5] for an account of the more recent developments. Interesting results on efficient equireplicate designs were reported among others by Das and Kulkarni [1], Kyi Win and Dey [7], Nigam and Boopathy [9], Gupta, Nigam and Puri [6] and Gupta [3]. Gupta and Mukerjee [4] attempted to unify these results to some extent.

A review of the literature shows that designs proposed so far for parallel line assays are mostly equireplicate. Such designs cannot be used when the number, n, of experimental units is not an integral multiple of the number, v, of treatments. Moreover, as noted in Mukerjee and Gupta [8], even when n/v is an integer, a non-equireplicate design can be more efficient than the best equireplicate design for estimating the contrasts of interest. These authors worked with the A-criterion and posed an open problem concerning the development of efficient designs under other criteria.

In the present paper, consider the D-criterion and, as in Mukerjee and Gupta [8], allow the competing class of designs to include equireplicate as well as non-equireplicate designs. Note that the D-criterion is invariant of the scaling of the treatment contrasts of interest and, therefore, in a sense more appealing than the A-criterion. However, in the present context, the exact D-optimal design problem is far more intractable than the corresponding A-optimal design problem. This is because, as one can demonstrate via examples, a counterpart of the crucial Lemma 2.1 of Mukerjee and Gupta is not available under the D-criterion. As such, unlike what happens with the A-criterion, even in the absence of blocks, the non-linear integer programming problem associated with the D-criterion (see 2.3 below) cannot in general be reduced to one involving fewer variables. In Section 2, one can observe that considerable simplification is possible taking recourse to the approximate theory. This yields D-optimal design measures which can, in turn, lead to exact designs with very high D-efficiency both in the absence and presence of blocks. A method for constructing D-efficient block designs via the approximate theory has been presented in Section 3. Just as with the A-criterion, our D-efficient designs are often non- equireplicate. It may also be noted with satisfaction that they tend to perform nicely under the A-criterion as well.

2. *The Approximate Theory*

Consider a symmetric parallel line bioassay experiment involving two preparations, standard and test, each at m (≥ 2) equispaced doses. Let these $v\ (= 2m)$ treatments be coded as 1, 2, ..., 2m and let $\tau = (\tau_1, ..., \tau_v)'$ be the vector of treatment effects where τ_i and τ_{m+i} denote the effects of the ith dose of the standard and test preparation respectively. As happens in most practical situations, suppose interest lies in the preparation contrast, the combined regression contrast and the parallelism contrast given respectively by $\pi_1' \tau, \pi_2' \tau$ and $\pi_3' \tau$, where

$$\pi_1' = (1_m' \ -1_m'),\ \pi_2' = (e' \ \ e'),\ \pi_3' = (e' \ -e') \tag{2.1}$$

with $e = f - \frac{1}{2}(m+1)\,1_m$, $f = (1, 2, ..., m)'$ and 1_m representing the m x 1 vector with each element unity. Let P be a 3xv matrix defined as

$$P = \begin{pmatrix} \pi'_1 \\ \pi'_2 \\ \pi'_3 \end{pmatrix} = \begin{pmatrix} 1'_m & -1'_m \\ e' & e' \\ e' & -e' \end{pmatrix} \tag{2.2}$$

By (2.1) and (2.2), $P\tau$ represents the contrasts of interest. Let $p_1, \ldots, p_v$ denote the columns of P.

First consider an unblocked set-up with n experimental units. For $1 \leq i \leq v$, let r_i be the replication number of the ith treatment, where $r_1, \ldots, r_v$ are positive integers satisfying $r_1 + \ldots + r_v = n$. We require all the r_i's to be positive to ensure the estimability of $P\tau$. As usual, the errors are assumed to be uncorrelated and homoscedastic with common variance σ^2. Then it is easy to see that $\text{Disp}(P\hat{\tau}) = \sigma^2 PR^{-1}P'$, where $R = \text{diag}(r_1, ..., r_v)$ and $P\hat{\tau}$ is the best linear unbiased estimator of $P\tau$ in the design under consideration.

The exact D-optimal design problem in the unblocked case can, therefore, be formulated as

$$\text{minimize } |PR^{-1}P'| \text{ subject to } r_1 + \ldots + r_v = n \tag{2.3}$$

where $r_1, \ldots, r_v$ are positive integers.

As noted earlier, this exact problem, which requires non-linear integer programming, is intractable even for moderately large n and v. Considerable simplification is achieved if one takes recourse to the approximate theory, writes $x_i = r_i / n$ $(1 \le i \le v)$ and allows $x = (x_1, ..., x_v)'$ to vary continuously in the set $T = \{x : x_1 + \ldots + x_v = 1,\ x_1 > 0, ..., x_v > 0\}$ (cf. Silvey [10]). This leads to a continuous version of the problem (2.3) where, writing $X = \text{diag}(x_1, ..., x_v)$, one has to minimize $|PX^{-1}P'|$, or equivalently $f(x) = \log |PX^{-1}P'|$, over $x \varepsilon T$, to get a D-optimal design measure.

Now, f(x) is strictly convex over T and tends to $+\infty$ along any x-sequence converging to a point in $\bar{T} - T$ where $\bar{T}$ is the closure of T. Hence there exists a unique point, say $x^* = (x_1^*, \ldots, x_v^*)'$, in T where the minimum of f(x), over $x \varepsilon T$, is attained. Using the Lagrangian method, x^* is given by the unique stationary point of the strictly convex function $f^*(x) = f(x) + \lambda(x_1 + \ldots + x_v)$, where the constant λ is so chosen that $x_1^* + \ldots + x_v^* = 1$. Since

$$\partial f^*(x) / \partial x_i = -x_i^{-2}\ p'_i\,(PX^{-1}P')^{-1}\,p_i + \lambda, \quad 1 \le i \le v$$

it follows that the point x^*, representing the D-optimal design measure, is given by the unique solution in T of the system of equations

$$x_i^2 = \frac{1}{3}\,p'_i\ (PX^{-1}P')^{-1}p_i, \quad 1 \le i \le v \tag{2.4}$$

In particular, if m=2 then by (2.2) and (2.4), $x^* = \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)'$ which is in agreement with the findings in Mukerjee and Gupta [8] under the A-criterion. Hereafter, consider $m \ge 3$ and, separately for even and odd m, explore how (2.4) can be simplified via a reduction in the number of variables.

2.1 *Case of even m*

Let $m = 2t$, $t \ge 2$. Let $R_t = \{ y = (y_1, ..., y_t)' : y_1 > 0, \ldots, y_t > 0 \}$ and for $y \varepsilon R_t$, define

$$g_1(y) = \Sigma_{j=1}^{t}\ y_j^{-1}, \quad g_2(y) = \Sigma_{j=1}^{t}\ y_j^{-1}\ \{j - \frac{1}{2}(m+1)\}^2 \tag{2.5}$$

Lemma 1. The system of equations

$$y_j^2 = \frac{1}{12}\left[\{ g_1 (y) \}^{-1} + 2 \{ g_2 (y) \}^{-1} \{ j - \frac{1}{2}(m+1) \}^2 \right], \ 1 \le j \le t \tag{2.6}$$

has a unique solution in R_t.

Proof. Let $g(y) = \log g_1 (y) + 2 \log g_2 (y) + 12 \Sigma_{j=1}^{t} y_j$, $y \varepsilon R_t$. Then g(y) is strictly convex over R_t and there is a unique point in R_t where the minimum of g(y), over $y \varepsilon R_t$, is attained. Consequently, g(y) has a unique stationary point in R_t. The result now follows considering the partial derivatives of g(y).

Theorem 1. Let $y^* = (y_1^*, \ldots, y_t^*)'$ be the unique solution of (2.6) in R_t. Then the vector

$$x = (y_1^*, \ldots, y_t^*, y_t^*, \ldots, y_1^*, y_1^*, \ldots, y_t^*, y_t^*, \ldots, y_1^*)' \tag{2.7}$$

with v (= 4t) elements, belongs to T and represents the D-optimal design measure for the estimation of P τ.

Proof. Since y^* solves (2.6), replacing y by y^* in (2.6),

$$y_j^* = \frac{1}{12} (y_j^*)^{-1} [\{ g_1 (y^*)\}^{-1} + 2 \{ g_2 (y^*)\}^{-1} \{ j - \frac{1}{2}(m+1) \}^2]$$

$$1 \le j \le t$$

and, in view of (2.5), summing the above over j, $y_1^* + \ldots + y_t^* = \frac{1}{4}$. Hence by (2.7), $x^* \varepsilon T$. It remains to show that x^* satisfies (2.4).

To that effect, write $x^* = (x_1^*, \ldots, x_v^*)'$, $p_i = (p_{1i}, p_{2i}, p_{3i})'$, $1 \le i \le v$, and define the sets $S_j = \{ j, m+1-j, m+j, 2m+1-j\}$, $1 \le j \le t$, which provide a disjoint partitioning of $\{ 1, \ldots, v \}$. Then by (2.2) and (2.7), for $1 \le i \le v$,

$$p_{1i}^2 = 1, \quad p_{2i}^2 = p_{3i}^2 = \{ j - \frac{1}{2}(m+1) \}^2 \quad \text{if } i \varepsilon S_j \tag{2.8}$$

$$x_i^* = y_j^* \quad \text{if } i \varepsilon S_j \tag{2.9}$$

Let X^* be a vxv diagonal matrix with diagonal entries given by the elements of x^*. Then by (2.2), (2.5) and (2.7), after some algebra,

$$P(X^*)^{-1}P' = 4\,\text{diag}\,[g_1(y^*), g_2(y^*), g_2(y^*)]$$

Hence by (2.8), if $i \varepsilon S_j$ then

$$\frac{1}{3}p'_i\{P(X^*)^{-1}P'\}^{-1}p_i = \frac{1}{12}[\{g_1(y^*)\}^{-1} + 2\{g_2(y^*)\}^{-1}\{j - \frac{1}{2}(m+1)\}^2] \tag{2.10}$$

Since y^* solves (2.6), it follows from (2.9) and (2.10) that x^* solves (2.4).

Remark 1. For even m, Theorem 1 considerably simplifies the derivation of the D-optimal design measure. Instead of (2.4), one needs to consider the system of equations (2.6) involving only $t = \frac{1}{4}v$ variables. We solved (2.6) iteratively for $t = 2, 3, 4, 5$ (i.e., v = 8, 12, 16, 20) starting with the initial choice $y_j = (4t)^{-1}$, $1 \le j \le t$. The solution $y^* = (y_1^*, \ldots, y_t^*)'$, yielding the D-optimal design measure, has been shown in Table 1 for $t = 2, 3, 4, 5$ (i.e., v = 8, 12, 16, 20).

2.2 *Case of odd m*

Let m = 2t+1, t ≥ 1. Let

$$R_{t+1} = \{y=(y_0, y_1, \ldots, y_t)' : y_0 > 0, y_1 > 0, \ldots, y_t > 0\}$$

and for $y \varepsilon R_{t+1}$, define

$$h_1(y) = \frac{1}{2}y_0^{-1} + \Sigma_{j=1}^{t} y_j^{-1}, \quad h_2(y) = \Sigma_{j=1}^{t} y_j^{-1}\{j - \frac{1}{2}(m+1)\}^2$$

Lemma 2. The system of equations

$$y_0^2 = \frac{1}{12}\{h_1(y)\}^{-1}$$

$$y_j^2 = \frac{1}{12}[\{h_1(y)\}^{-1} + 2\{h_2(y)\}^{-1}\{j - \frac{1}{2}(m+1)\}^2], \ 1 \le j \le t \tag{2.11}$$

has a unique solution in R_{t+1}.

Proof. Follows along the line of proof of Lemma 1 now considering the function $h(y) = \log h_1(y) + 2\log h_2(y) + 6y_0 + 12\,\Sigma_{j=1}^{t} y_j$, $y \varepsilon R_{t+1}$

Table 1. D-optimal design measures

v	y_0^*	y_1^*	y_2^*	y_3^*	y_4^*	y_5^*
6	0.0892	0.2054	-	-	-	-
8	-	0.1652	0.0848	-	-	-
10	0.0542	0.1390	0.0839	-	-	-
12	-	0.1194	0.0802	0.0504	-	-
14	0.0388	0.1046	0.0755	0.0505	-	-
16	-	0.0930	0.0706	0.0503	0.0361	-
18	0.0302	0.0838	0.0659	0.0493	0.0359	-
20	-	0.0762	0.0616	0.0479	0.0360	0.0283

Proceeding along the line of proof of Theorem 1, one can get the following result.

Theorem 2. Let $y^* = (y_0^*, y_1^*, ..., y_t^*)'$ be the unique solution of (2.11) in R_{t+1}. Then the vector

$$x^* = (y_1^*, ..., y_t^*, y_0^*, y_t^*, ..., y_1^*, y_1^*, ..., y_t^*, y_0^*, y_t^*, ..., y_1^*)' \qquad (2.12)$$

with v (= 4t+2) elements, belongs to T and represents the D-optimal design measure for the estimation of $P\tau$.

Remark 2. For odd m, Theorem 2 considerably simplifies the derivation of the D-optimal design measure. Instead of (2.4), one needs to consider the system of equations (2.11) involving $t+1 = \frac{1}{4}(v+2)$ variables. We solved (2.11) interatively for t = 1, 2, 3, 4 starting with the initial choice $y_j = (4t+2)^{-1}$, $0 \le j \le t$. The solution, yielding the D-optimal design measure, are shown in Table 1 for t=1, 2, 3, 4 (i.e., for v=6, 10, 14, 18).

2.3 *Examples in the unblocked case*

As the following examples reveal, the D-optimal design measures can yield exact designs with very high D-efficiency. These exact D-efficient designs are non-equireplicate and the gain, compared to corresponding equireplicate designs, is not ignorable.

Example 1. Let n = 24, v = 8. From Theorem 1 and Table 1, the D-optimal design measure is given by $x^* = (x_1^*, ..., x_8^*)'$, where $x_1^* = x_4^* = x_5^* = x_8^* = 0.1652$, $x_2^* = x_3^* = x_6^* = x_7^* = 0.0848$. For $1 \le i \le 8$, rounding off nx_i^* to the nearest integer, x^* yields the exact design d_0 given by the vector of replication numbers $r_0 = (4, 2, 2, 4, 4, 2, 2, 4)'$. As before, let X^* and R_0 be diagonal matrices with diagonal entries given by the elements of x^* and r_0 respectively. Then the D-efficiency of d_0, measured as

$$\text{eff} = \{ |P(X^*)^{-1}P'| / |P(n^{-1}R_0)^{-1}P'| \}^{1/3} \tag{2.13}$$

is as high as 0.9999. Also, the D-efficiency of the equireplicate design, relative to d_0 and defined along the line of (2.13), equals 0.9149.

Example 2. Let n = 50, v = 10. The D-optimal design measure x*, obtained from Theorem 2 and Table 1, yields the exact design d_0 given by the vector of replication numbers $r_0 = (7, 4, 3, 4, 7, 7, 4, 3, 4, 7)'$. By (2.13), the D-efficiency of d_0 equals 0.9982 while the D-efficiency of the equireplicate design, relative to d_0, turns out to be 0.9106.

Example 3. Let n = 10, v = 6. Then no equireplicate design exists and, as before, x* yields the exact design d_0 with $r_0 = (2, 1, 2, 2, 1, 2)'$ and having D-efficiency 0.9977.

Remark 3. In each of the above examples, the D-efficiency of d_0 is defined with respect to a D-optimal design measure which is not actually attainable in the corresponding exact setting. As such, even the very high figures for the D- efficiency of d_0, as reported above, are rather conservative and in some or all of the above examples d_0 may actually be D-optimal within the relevant classes of exact designs.

Remark 4. Following Mukerjee and Gupta [8] one can check that in each of these examples the D-efficient designs d_0 is also A-optimal for the estimation of appropriately normed version of $P\tau$ within the class of the designs involving the same number of observations.

3. *D-efficient Block Designs*

Suppose now it is intended to conduct the experiment in b blocks each of size k. As before, there are v=2m treatments. We work under the usual fixed effects additive linear model with uncorrelated and homoscedastic errors and denote the error variance by σ^2. The approximate theory developed above can be utilized to obtain highly D-efficient block designs when the block size is an integral multiple of 4. The approach is reminiscent of that in Mukerjee and Gupta [8] in the context of A-optimality. First, given v, b, k, the replication numbers are chosen as dictated by the D-optimal design measure and then blocking is done in such a manner that $P\tau$ is estimated orthogonally to block effects. Assuming that k is an integral multiple of 4, the constructional steps are described below separately for even and odd m.

First let $m = 2t$, $t \geq 2$.

Step 1. From Theorem 1 and (2.6), or using Table 1, obtain the D- optimal design measure $x^* = (x_1^*, ..., x_v^*)'$.

Step 2. Find positive integers $r_1, ..., r_v$ such that for $1 \leq i \leq v$, r_i is as close to x_i^* bk as possible subject to the conditions

$$r_j = r_{m+1-j} = r_{m+j} = r_{2m+1-j}, \quad (1 \leq j \leq t)$$

$$\Sigma_{i=1}^{v} \; r_i = bk$$

Let u_j be the common value of r_j, r_{m+1-j}, r_{m+j} and r_{2m+1-j} $(1 \leq j \leq t)$ and note that $\Sigma_{j=1}^{t} \; u_j = \frac{1}{4}bk$.

Step 3. Construct a design d involving t symbols, say $\phi_1, ..., \phi_t$, and b blocks each of size $\frac{1}{4}k$ such that ϕ_j is replicated u_j times in d, $1 \leq j \leq t$.

Step 4. Finally obtain a design d_0 from d replacing the symbol ϕ_j in d by the four treatments $j, m+1-j, m+j$ and $2m+1-j$, $(1 \leq j \leq t)$.

Next let $m = 2t+1$, $t \geq 1$.

Step 1. From Theorem 2 and (2.11), or using Table 1, obtain the D-optimal design measure $x^* = (x_1^*, ..., x_v^*)'$.

Step 2. Find positive integers $r_1, ..., r_v$ such that for $1 \leq i \leq v$, r_i is as close to x_i^* bk as possible subject to the conditions

$$r_j = r_{m+1-j} = r_{m+j} = r_{2m+1-j}, \quad (1 \leq j \leq t)$$

$$r_{t+1} = r_{m+t+1}, \text{ the common value being an even number}$$

$$\Sigma_{i=1}^{v} \; r_i = bk$$

Let u_j be the common value of r_j, r_{m+1-j}, r_{m+j} and r_{2m+1-j}, $(1 \leq j \leq t)$ and write $r_{t+1} = r_{m+t+1} = 2u_0$. Then $\Sigma_{j=0}^{t} \; u_j = \frac{1}{4}bk$.

Step 3. Construct a design d involving t + 1 symbols, say $\phi_0, \phi_1, \cdots, \phi_t$, and b blocks each of size $\frac{1}{4}k$ such that ϕ_j is replicated u_j times in d, $0 \le j \le t$.

Step 4. Obtain a design d_0 from d replacing the symbol ϕ_0 in d by the two treatments $t+1$ and $m+t+1$, each repeated twice, and for $1 \le j \le t$, the symbol ϕ_j in d by the four treatments $j, m+1-j$, $m+j$ and $2m+1-j$, each repeated once.

For m either even or odd, let N denote the incidence matrix of d_0. Then, with $R = \text{diag}(r_1, ..., r_v)$, in view of (2.2), our construction ensures that $PR^{-1} N = 0$, which is precisely the condition for estimation of $P\tau$ orthogonally to block effects; cf. Gupta and Mukerjee [5]. Hence under d_0, $\text{Disp}(P\hat{\tau}) = \sigma^2 PR^{-1} P'$. Thus, with reference to the estimation of $P\tau$, the D-efficiency of d_0 is given by (2.13) with n there replaced by bk. Unless bk is too small, the proximity of r_i to $x_i^* bk$ $(1 \le i \le v)$, as stipulated in Step 2, now ensures a high value for this D-efficiency.

In Step 3 of our construction, for m either even or odd, there is some flexibility in the choice of the design d. We recommend that as far as practicable d should be chosen as a connected design. Then d_0 will also be connected - see Mukerjee and Gupta [8] for a related discussion which is meaningful also in the present context.

Example 4. Let v=12, b=3, k=8. After obtaining x^* from Table 1 and Theorem 1, Step 2 yields $(r_1, \dots, r_v)' = (3, 2, 1, 1, 2, 3, 3, 2, 1, 1, 2, 3)'$. Hence $u_1 = 3$, $u_2 = 2$, $u_3 = 1$ and the blocks of d can be taken as $\{\phi_1, \phi_2\}$, $\{\phi_1, \phi_2\}$, $\{\phi_1, \phi_3\}$ following Step 3. Finally by Step 4, d_0 is given by the blocks $\{1, 6, 7, 12, 2, 5, 8, 11\}$, $\{1, 6, 7, 12, 2, 5, 8, 11\}$ and $\{1, 6, 7, 12, 3, 4, 9, 10\}$. By (2.13), the D-efficiency of d_0 is 0.9926. Using a similar formula, it can also be seen that the D-efficiency of any equireplicate design, relative to d_0, is at most 0.9141 in so far as the estimation of $P\tau$ is concerned.

Example 5. Let v = 10, b = 5, k = 8. After obtaining x^*, from Table 1 and Theorem 2, Step 2 yields $(r_1, ..., r_v)' = (6, 3, 2, 3, 6, 6, 3, 2, 3, 6)'$. Hence $u_0 = 1$, $u_1 = 6$, $u_2 = 3$ and starting from the design d with blocks $\{\phi_1, \phi_1\}$, $\{\phi_1, \phi_0\}$, $\{\phi_1, \phi_2\}$, $\{\phi_1, \phi_2\}$, $\{\phi_1, \phi_2\}$, it is easy to construct the design d_0 as in Step 4. The D-efficiency of d_0 is seen to equal 0.9925. One can also check that the D-efficiency of no equireplicate design, relative to d_0, can exceed 0.9158.

Remark 5. In each of the above examples, d_0 is connected and, following Mukerjee and Gupta [8], A-optimal for P_T within the class of designs with the same v, b, k. Note that Remark 3, holds in the context of block designs as well.

ACKNOWLEDGEMENT

This work was supported by a grant from the Centre for Management and Development Studies, Indian Institute of Management, Calcutta.

REFERENCES

[1] Das, M. N. and Kulkarni, G. A., 1966. Incomplete block designs for bio-assays. *Biometrics,* **22,** 706-729.

[2] Finney, D. J., 1978. *Statistical Methods in Biological Assay,* 3rd ed. Charles Griffin, London.

[3] Gupta, S., 1988. Designs for symmetrical parallel line assays obtainable through group divisible designs. *Comm. Statist. - Theory Methods*, **17,** 3865-3868.

[4] Gupta, S. and Mukerjee, R., 1990. On incomplete block designs for symmetric parallel line assays. *Austr. J. Statist.,* **32,** 337-344.

[5] Gupta, S. and Mukerjee, R., 1996. Developments in incomplete block designs for parallel line bioassays. In : *Handbook of Statistics,* (S. Ghosh and C.R. Rao eds.), **13,** North-Holland, Amsterdam (to appear).

[6] Gupta, V.K., Nigam, A.K. and Puri, P.D., 1987. Characterization and construction of incomplete block designs for symmetrical parallel line assays. *J. Indian Soc. Agric. Statist.,* **39,** 161-166.

[7] Kyi, Win and Dey, A., 1980. Incomplete block designs for parallel line assays. *Biometrics,* **36,** 487-492.

[8] Mukerjee, R. and Gupta, S., 1995. A-efficient designs for bioassays. *J. Statist. Plann. Inf.,* **48,** 247-259.

[9] Nigam, A. K. and Boopathy, G.M., 1985. Incomplete block designs for symmetrical parallel line assays. *J. Statist. Plann. Inf.*, **11,** 111-117.

[10] Silvey, S.D., 1980. *Optimal Design.* Chapman and Hall, London.