

Marginally restricted sequential D-optimal designs

Jesús LÓPEZ-FIDALGO, Raul MARTÍN-MARTÍN and Douglas P. WIENS

Key words and phrases: Conditional design; D-optimality; linear regression; logistic model; marginal design; nonadaptive design; optimal design; randomization.

MSC 2000: Primary 62K05; secondary 62J05.

Abstract: In many experiments, not all explanatory variables can be controlled. When the units arise sequentially, different approaches may be used. The authors study a natural sequential procedure for “marginally restricted” D-optimal designs. They assume that one set of explanatory variables (\mathbf{x}_1) is observed sequentially, and that the experimenter responds by choosing an appropriate value of the explanatory variable x_2 . In order to solve the sequential problem a priori, the authors consider the problem of constructing optimal designs with a prior marginal distribution for \mathbf{x}_1 . This eliminates the influence of units already observed on the next unit to be designed. They give explicit designs for various cases in which the mean response follows a linear regression model; they also consider a case study with a nonlinear logistic response. They find that the optimal strategy often consists of randomizing the assignment of the values of x_2 .

Plans séquentiels D-optimaux sous contraintes marginales

Résumé : Dans bien des expériences, on ne peut pas contrôler toutes les variables explicatives. Lorsque les unités arrivent séquentiellement, diverses approches sont possibles. Les auteurs étudient une procédure séquentielle naturelle pour les plans D-optimaux sous “contraintes marginales.” Ils supposent qu’un ensemble de variables explicatives (\mathbf{x}_1) est observé séquentiellement et que l’expérimentateur réagit en choisissant un niveau approprié pour la variable explicative x_2 . Pour résoudre ce problème séquentiel a priori, les auteurs proposent la construction de plans optimaux étant donné une loi a priori marginale pour \mathbf{x}_1 . Ceci a pour effet d’éliminer l’influence des unités déjà observées sur l’affectation de la prochaine unité. Ils fournissent des plans explicites dans divers cas où la réponse moyenne est modélisée par régression linéaire ; ils traitent aussi un exemple concret dans lequel la réponse est non linéaire et logistique. Ils concluent que la stratégie optimale consiste souvent à randomiser l’affectation des valeurs de x_2 .

1. INTRODUCTION

In optimal experimental design theory a common assumption is that all of the explanatory variables can be controlled completely. This assumption may be unrealistic. In many areas of application it is common that some of the known factors are not subject to control. For example, in a medical experiment the assignment of treatments to the patients is controlled, but age, blood pressure, etc., are not. We assume that the values of the uncontrolled variables are known before those of the controlled variables are assigned. Such designs are called marginally restricted (MR) designs. Cook & Thibodeau (1980) introduced the idea of MR D-optimal designs and Nachtsheim (1989) provided equivalence theorems and algorithms for D-optimality and D_s -optimality. Huang & Hsu (1993) generalized these results for linear criteria. López-Fidalgo & Garcet-Rodríguez (2004) proved the convergence of the algorithms for a class of criteria and extended these ideas and results to the case of uncontrolled variables whose values are known only after the experiment is realized. Martín-Martín, Torsney & López-Fidalgo (2007) used multiplicative algorithms for a more efficient construction of MR optimal designs.

Consider a linear model defined by regressors

$$\mathbf{f}^t(\mathbf{x}_1, x_2) = (f_1(\mathbf{x}_1, x_2), \dots, f_m(\mathbf{x}_1, x_2)),$$

with m linearly independent functions on some compact space χ , where $\chi = \chi_1 \times \chi_2$ is the design space for the explanatory variables in the model. A single observation involves selecting

$(\mathbf{x}_1, x_2) \in \chi_1 \times \chi_2$ and observing a random variable Y with regression function

$$E[Y | x] = \mathbf{f}^t(\mathbf{x}_1, x_2)\boldsymbol{\gamma}$$

and constant variance. The variable \mathbf{x}_1 , possibly vector-valued, is known before the experiment is realized and cannot be controlled by the experimenter. The variable x_2 is under control and must be assigned for each given value of \mathbf{x}_1 . The function \mathbf{f} is assumed known and $\boldsymbol{\gamma}$ is unknown. Any probability measure ξ on χ is referred to as an experimental design, and the set of all designs is denoted by Ξ .

A joint design is completely determined by a marginal design ξ_1 on χ_1 and conditional designs $\xi_{2|1}(\cdot | \mathbf{x}_1)$ on χ_2 given $\mathbf{x}_1 \in \chi_1$. Its corresponding information matrix $\mathbf{M}(\xi)$ is defined as

$$\mathbf{M}(\xi) = \int_{\chi_1} \int_{\chi_2} \mathbf{f}(\mathbf{x}_1, x_2) \mathbf{f}^t(\mathbf{x}_1, x_2) \xi_{2|1}(dx_2 | \mathbf{x}_1) \xi_1(d\mathbf{x}_1). \quad (1)$$

Let \mathcal{M} be the compact set of information matrices. Suppose that Φ is a criterion function on \mathcal{M} , such that it is convex and nonincreasing with respect to the Loewner ordering (i.e., if $\mathbf{M} \leq \mathbf{N}$ in the sense of positive semidefiniteness, then $\Phi(\mathbf{M}) \geq \Phi(\mathbf{N})$). A design ξ^* is called Φ -optimal if it achieves $\min\{\Phi[\mathbf{M}(\xi)] : \xi \in \Xi\}$. To measure the goodness of a specific design for a particular criterion an additional hypothesis is usually made in the definition of a criterion function, namely that $\Phi(\lambda\mathbf{M}) = \frac{1}{\lambda} \Phi(\mathbf{M})$, $\lambda > 0$. Then, the Φ -efficiency of a design ξ , relative to ξ^* , is defined as

$$\text{eff}_\Phi[\mathbf{M}(\xi)] = \Phi[\mathbf{M}(\xi^*)]/\Phi[\mathbf{M}(\xi)].$$

The most popular criterion, and that to which we now restrict the discussion, is D-optimality. The goal of this criterion is to maximize the determinant of the information matrix; specifically

$$\Phi[\mathbf{M}(\xi)] = |\mathbf{M}|^{-1/m}. \quad (2)$$

Minimizing (2) is equivalent to minimizing the volume of a confidence ellipsoid on the parameters.

Since the values of the variable \mathbf{x}_1 are known in advance, they determine a design ξ_1 . Thus, we should consider the space Ξ_R restricted to joint designs with ξ_1 as the marginal design. Then the design problem is that of constructing an optimal conditional design $\xi_{2|1}^*(\cdot | \mathbf{x}_1)$. We note that Ξ_R and \mathcal{M}_R are convex sets.

Since the marginal design comes from observed values of \mathbf{x}_1 , ξ_1 has always been considered in the literature to have finite support. In this article we also allow (continuous or discrete) marginal distributions with infinite support. In such cases it may happen that the joint design will not have finite support. Nevertheless, for each $\mathbf{x}_1 \in \chi_1$ there is a guarantee of the existence of a conditional design $\xi_{2|1}(\cdot | \mathbf{x}_1)$ to be carried out in practice, with at most $1 + m(m+1)/2$ points in its support (López-Fidalgo & Garcet-Rodríguez 2004).

López-Fidalgo & Garcet-Rodríguez (2004) presented a lower bound for the efficiency for marginally restricted designs:

$$\text{eff}_\Phi[\mathbf{M}(\xi)] \geq 1 + \frac{\int_{\chi_1} \inf_{x_2 \in \chi_2} \partial\Phi[\mathbf{M}(\xi), \mathbf{M}_{(\mathbf{x}_1, x_2)}] \xi_1(d\mathbf{x}_1)}{\Phi[\mathbf{M}(\xi)]}, \quad (3)$$

where

$$\partial\Phi[\mathbf{M}(\xi), \mathbf{M}_{(\mathbf{x}_1, x_2)}] = \lim_{\varepsilon \downarrow 0} \frac{\Phi[(1-\varepsilon)\mathbf{M}(\xi) + \varepsilon\mathbf{f}(\mathbf{x}_1, x_2)\mathbf{f}^t(\mathbf{x}_1, x_2)]}{\varepsilon}.$$

The relevant equivalence theorem states that this bound is attained only by the Φ -optimal design ξ^* which, in order that it have unit efficiency, must necessarily satisfy

$$\int_{\chi_1} \inf_{x_2 \in \chi_2} \partial\Phi[\mathbf{M}(\xi^*), \mathbf{M}_{(\mathbf{x}_1, x_2)}] \xi_1(d\mathbf{x}_1) = 0.$$

For the D-optimality criterion (2) we have

$$\partial\Phi[\mathbf{M}(\xi), \mathbf{M}_{(\mathbf{x}_1, x_2)}] = \frac{1 - d(\mathbf{x}_1, x_2; \xi)/m}{|\mathbf{M}(\xi)|^{1/m}},$$

where $d(\mathbf{x}_1, x_2; \xi)$ is the dispersion function defined by

$$d(\mathbf{x}_1, x_2; \xi) = \mathbf{f}^t(\mathbf{x}_1, x_2) \mathbf{M}^{-1}(\xi) \mathbf{f}(\mathbf{x}_1, x_2). \quad (4)$$

A necessary and sufficient condition for the design ξ^* to be MR D-optimal is thus

$$\int_{\chi_1} \max_{x_2} d(\mathbf{x}_1, x_2; \xi^*) \xi_1(d\mathbf{x}_1) = m. \quad (5)$$

After some algebra the bound (3), applied in the case of D -efficiency, is found to be

$$\text{eff}_{D-\text{MR}}(\xi) \geq 2 - \frac{1}{m} \int_{\chi_1} \max_{x_2 \in \chi_2} d(\mathbf{x}_1, x_2; \xi) \xi_1(d\mathbf{x}_1). \quad (6)$$

A result of some independent interest, proven in the Appendix, gives a sharper bound. It generalizes a bound for the D -efficiency proposed by Atwood (1969), and may also be viewed as a consequence of a general formula for the efficiency bound given by Dette (1996).

THEOREM 1. *If $\mathbf{M}(\xi)$ is a nonsingular matrix, then*

$$\text{eff}_{D-\text{MR}}(\xi) \geq \frac{m}{\int_{\chi_1} \max_{x_2} d(\mathbf{x}_1, x_2; \xi) \xi_1(d\mathbf{x}_1)}. \quad (7)$$

The bound (7) is sharper than (6), i.e.,

$$2 - \frac{1}{m} \int_{\chi_1} \max_{x_2 \in \chi_2} d(\mathbf{x}_1, x_2; \xi) \xi_1(d\mathbf{x}_1) \leq \frac{m}{\int_{\chi_1} \max_{x_2} d(\mathbf{x}_1, x_2; \xi) \xi_1(d\mathbf{x}_1)},$$

since this inequality can be re-arranged as

$$0 \leq \left(m - \int_{\chi_1} \max_{x_2} d(\mathbf{x}_1, x_2; \xi) \xi_1(d\mathbf{x}_1) \right)^2.$$

Both bounds are attained only by the MR D-optimal design ξ^* ; this follows from (5).

The main goal of this article is to design optimally for the arrival of a new experimental unit with a particular value \mathbf{x}_1 . Two methods are considered:

Method 1. Here we first design optimally for an initial set of observations. Then each time a new experimental unit appears, its value \mathbf{x}_1 is noted and a value x_2 is assigned in an optimal manner. Thus that part of the design already performed is preserved.

Method 2. In this case we assume a priori a particular distribution for the variable \mathbf{x}_1 , e.g., a normal or a multinomial distribution, and design for the whole process. This results in a conditional design $\xi_{2|1}(x_2 | \mathbf{x}_1)$ which is applied upon observing a new unit with value \mathbf{x}_1 .

We first note the more “traditional” approach, in which one has an entire sample of size n at the outset. The value of \mathbf{x}_1 is measured in each unit. These values form a discrete, empirical marginal design, and one can go on to compute the MR optimal design using the algorithms

furnished by Cook & Thibodeau (1980), López-Fidalgo & Garcet-Rodríguez (2004) or Martín-Martín, Torsney & López-Fidalgo (2007).

In Section 2 a sequential procedure is provided according to Method 1. This procedure starts with an empirical marginal design defined by the observed values of this variable at the moment. Subsequently new observations will arise. Thus when a new observation appears, the marginal design changes, and we calculate the best MR D-optimal design taking into account the design already performed. This means searching for a new optimal design point x_2 each time a new unit appears, according to the observed value x_1 for this unit.

Sometimes this procedure is not adequate in practice and a complete experimental design should be given at the very beginning. Thus, a plausible marginal distribution ξ_1 for x_1 is assumed a priori. Then a restricted optimal joint design is computed for this theoretical marginal distribution. In Section 3 we consider this approach to the problem for a two-treatment design problem and a linear regression model. This problem, with a finite, discrete marginal distribution, may be solved in the same way as for the empirical marginal distribution—i.e., by applying the traditional approach. The problem of a marginal design with infinite support cannot be solved under this framework—even large supports lead to a high computational cost—and so new methods are introduced in Section 3 to accommodate this case of marginals ξ_1 with infinite support. In many cases this leads to the, not unexpected, strategy of randomizing the assignment of the values of x_2 . In Section 4 we apply these techniques to a case study involving a nonlinear regression response.

We remark that ‘sequential’ is interpreted here in terms of the order in which values of x_1 are observed. This is to be contrasted with ‘adaptive’ design, in which the choice of design points depends on previously observed values of the response variable and for which the resulting inferential problems can be considerably more complicated—see, e.g., the various contributions in Flournoy & Rosenberger (1995).

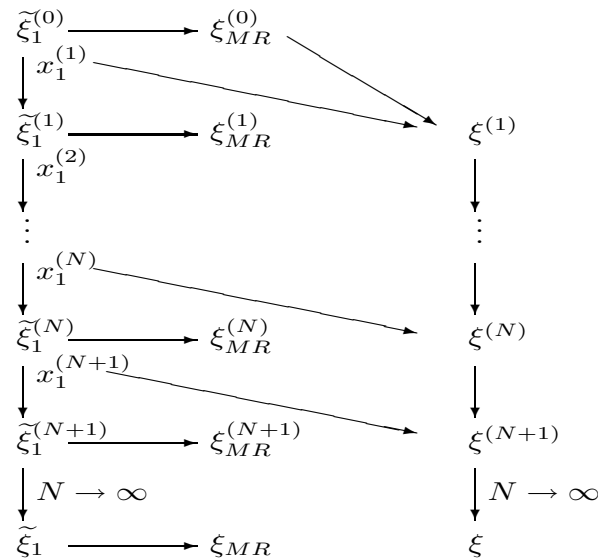


FIGURE 1: Construction of sequential designs.

2. SEQUENTIAL CONSTRUCTION OF marginally RESTRICTED D-OPTIMAL DESIGNS

An initial empirical marginal design $\xi_1^{(0)}$ is set up with the measurements of the variable x_1 for the experimental units in the study at an initial time. From this, the MR D-optimal design $\xi_{MR}^{(0)}$

is determined, minimizing (2). When a new observation $\mathbf{x}_1^{(1)}$ arrives, the marginal design $\xi_1^{(0)}$ is modified to $\xi_1^{(1)}$. In order to preserve $\xi_{\text{MR}}^{(0)}$, this is done by choosing

$$x_2^{(1)} = \arg \max_{x_2} d(\mathbf{x}_1^{(1)}, x_2; \xi_{\text{MR}}^{(0)}) = \arg \max_{x_2} |\mathbf{M}(\xi^{(1)})|,$$

where $\xi^{(1)}$ is the joint design with the point $(\mathbf{x}_1^{(1)}, x_2^{(1)})$ added. The second equality is a direct consequence of Fedorov (1972, p. 177) applied to this restricted case. Continuing with this process we realize a sequential procedure to design optimally the experiment in consecutive steps. We note that the designs obtained in this way are at each step the best designs preserving the joint design that has already been performed at that particular moment and taking into account the new point $\mathbf{x}_1^{(N)}$.

Given a marginal design $\xi_1^{(N)}$, the proper MR D-optimal design $\xi_{\text{MR}}^{(N)}$ may be obtained as well. This can generally not be implemented in a manner consistent with the design $\xi^{(N-1)}$ already applied. Nonetheless, one naturally wonders if the two sequences of designs so obtained agree asymptotically. This is an open question of some interest. This procedure is illustrated in Figure 1.

Two simple examples are considered to illustrate the procedure and to shed some light on this question.

2.1. Spring balance weighing model.

This popular model consists of $m = 2$ binary factors indicating the presence (1) or absence (0) of weights in a weighing operation. The parameters γ_1, γ_2 represent the unknown weights. It has been frequently considered in the optimal design literature. Thus Cheng (1987) determined the Φ -optimal approximate designs. Dette & Studden (1993) presented a geometric solution for the E -optimal spring balance weighing design. Harman (2004) proposed minimal efficiency ratios of the D-optimal designs for these models.

Consider a linear regression model with two variables without intercept,

$$E[Y | x_1, x_2] = \gamma_1 x_1 + \gamma_2 x_2, \quad \chi_1 = \{0, 1\}, \quad \chi_2 = \{0, 1\}.$$

The variable x_2 is completely under the control of the experimenter, and x_1 is a variable not subject to control with its values known before the test is performed.

The experiment is carried out for some initial experimental units and then the response variable is observed. Assume now that new experimental units arise sequentially. For each new experimental unit the variable x_1 may be observed. Then, a decision has to be made about what particular value of x_2 should be selected for this specific experimental unit to optimize the design.

Sequential optimal designs according to the marginal restriction will be constructed. Depending on the value of x_1 , we have two possibilities for each case. The MR D-optimal sequential procedure described at the beginning of this section may be summarized for this particular example as follows:

1. Let $\xi_1^{(0)}$ be the initial marginal design.
2. We calculate $\xi_{\text{MR}}^{(0)}$ under the restriction of the marginal design $\xi_1^{(0)}$.
3. At step $N + 1$ there are two possibilities for the new experimental unit coming to the study,

(a) If $x_1^{(N+1)} = 1$, then

$$x_2^{(N+1)} = \begin{cases} 1 & \text{if } \xi_{2|1}^{(N)}(1|1) < \frac{1}{2}, \\ 0 & \text{if } \xi_{2|1}^{(N)}(1|1) > \frac{1}{2}, \\ \text{randomize } \{0, 1\} & \text{if } \xi_{2|1}^{(N)}(1|1) = \frac{1}{2}. \end{cases}$$

(b) If $x_1^{(N+1)} = 0$, then $x_2^{(N+1)} = 1$.

We apply the algorithm to a particular case in order to clarify these ideas. Take the initial marginal design to be

$$\xi_1^{(0)} \equiv \begin{Bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{Bmatrix}.$$

After some algebra the MR D-optimal design is found to be

$$\xi_{\text{MR}}^{(0)} \equiv \begin{Bmatrix} (0,0) & (0,1) & (1,0) & (1,1) \\ 0 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{Bmatrix}.$$

Suppose we know, from retrospective studies, that in a long term the variable x_1 will be distributed according to a Bernoulli distribution with parameter $\theta = 1/5$. We simulate 8 observations for x_1 : $x_1 = 0, 0, 0, 1, 0, 0, 0, 0$.

TABLE 1: Sequential design weights at points $x_1 \in \{0, 1\}$ (second column) and at points $(x_1, x_2) \in \{(0,0), (0,1), (1,0), (1,1)\}$ (third and fourth columns). The last column shows the L^∞ norm of the difference between the sequential design $\xi^{(N)}$, preserving $\xi^{(N-1)}$, and the corresponding theoretical MR D-optimal design $\xi_{\text{MR}}^{(N)}$ at each step.

N	$\xi_1^{(N)}$	$\xi_{\text{MR}}^{(N)}$	$\xi^{(N)}$	$\max_x \{ \xi_{\text{MR}}^{(N)}(x) - \xi^{(N)}(x) \}$
1	2/3, 1/3	0, 2/3, 1/6, 1/6	0, 3/5, 1/5, 1/5	0.200
2	3/4, 1/4	0, 3/4, 1/8, 1/8	0, 4/6, 1/6, 1/6	0.133
3	4/5, 1/5	0, 4/5, 1/10, 1/10	0, 5/7, 1/7, 1/7	0.086
4	4/6, 2/6	0, 4/6, 1/6, 1/6	0, 5/8, 1/8, 2/8	0.175
5	5/7, 2/7	0, 5/7, 1/7, 1/7	0, 6/9, 1/9, 2/9	0.133
6	6/8, 2/8	0, 6/8, 1/8, 1/8	0, 7/10, 1/10, 2/10	0.100
7	7/9, 2/9	0, 7/9, 1/9, 1/9	0, 8/11, 1/11, 2/11	0.081
8	8/10, 2/10	0, 8/10, 1/10, 1/10	0, 9/12, 1/12, 2/12	0.050

Applying the procedure described above, we obtain the results given in Table 1. There seems to be some evidence of convergence, although it is nonmonotonic.

2.2. Example with a continuous marginal distribution generator.

The same model—linear regression through the origin—is considered here with $\chi_1 = [0, 1]$. The optimal designs are computed in the same manner as in the previous section, but the values of x_1 are now simulated with a continuous distribution instead of the Bernoulli distribution. As an example, a simulation with a Beta(1, 2) marginal distribution was performed, yielding the initial marginal design $\xi_1^{(0)} = \{0.0840, 0.3419, 0.4370\}$. After some algebra the MR D-optimal design computed assigned $x_2 = 1$ for the three values: $\xi_{\text{MR}}^{(0)} \equiv \{(0.0840, 1), (0.3419, 1), (0.4370, 1)\}$. New simulated values of Beta(1, 2) were obtained and using the sequential procedure of Section 2.1, the following points entered the design sequentially:

(0.3058, 1), (0.8083, 0), (0.0170, 1), (0.0306, 1), (0.7886, 0), (0.1619, 1), (0.6865, 0), (0.2712, 1).

TABLE 2: Norm L^∞ of the difference between the sequential design and the corresponding theoretical MR D-optimal design at each step N .

N	1	2	3	4	5	6	7	8
$\max_x \{ \xi_{\text{MR}}^{(N)}(x) - \xi^{(N)}(x) \}$	0.070	0.123	0.167	0.143	0.125	0.111	0.100	0.090

Table 2 shows the norm L^∞ of the difference between the sequential design and the corresponding theoretical MR D-optimal design at each step. Since the marginal, in the limit, is now continuous the joint MR D-optimal is continuous in the first variable. Therefore a normalization of the sequential weights according to the optimal design had to be done in order to compute the norm. Again, although there is an absence of monotonicity there is a decreasing trend in the norms, supporting convergence of the designs.

3. marginally restricted D-optimal design with a general marginal distribution

As was the case in Section 2.1 and Section 2.2, a theoretical distribution may be assumed for the uncontrolled variable \mathbf{x}_1 . Thus, a general MR optimal design may be computed a priori and applied sequentially as soon as the new units arrive, without additional computations. If the marginal distribution is discrete with only a few points in its support, then the general algorithms (e.g., Cook & Thibodeau 1980; López-Fidalgo & Garcet-Rodríguez 2004; Martín-Martín, Torsney & López-Fidalgo 2007) are still applicable. In fact in each step the number of (conditional) designs computed is the number of support points of the marginal design. However, if the distribution has infinite support—for instance, in the case of a continuous distribution—these procedures can no longer be applied. Even for a large support, the computations become very demanding.

In earlier work, not shown here, we took $\chi_2 = \{0, 1\}$. The (univariate) cumulative distribution function G_1 corresponding to ξ_1 , viz. $G_1(x) = \int_{-\infty}^x \xi_1(dt)$, was taken to be a particular beta distribution. This cumulative distribution function was approximated by a step function with jumps at the midpoints of adjacent quantiles of G_1 . These quantiles were equally spaced along the vertical axis, except for adjustments to the first and last of them with a half length in each such subinterval. A similar discretization was studied, with equal spacings along the horizontal axis. In either case, the solution to this approximating problem was that the optimal function $p(x_1) = P_{\xi_{2|1}}(x_2 = 1 | x_1)$ was of the form $p(x_1) = I(x_1 \in S)$ for a certain interval S . This motivated a general solution.

We begin the development with the following convenient characterization of condition (5), in the case of a K -treatment problem (more generally, the case in which χ_2 is discrete and of cardinality K).

THEOREM 2. *Suppose that $\chi_2 = \{t_0, t_1, \dots, t_{K-1}\}$ and that the design ξ^* has conditional $\xi_{2|1}^*$ satisfying*

$$p_k(\mathbf{x}_1) \stackrel{\text{def}}{=} \xi_{2|1}^*(t_k | \mathbf{x}_1) = \begin{cases} \frac{1}{j}, & \text{if } \max_{x_2 \in \chi_2} d(\mathbf{x}_1, x_2; \xi^*) \text{ is attained} \\ & \text{at } j \text{ values of } x_2, \text{ including } t_k, \\ 0, & \text{otherwise.} \end{cases}$$

Then (5) holds, so that ξ^ is marginally restricted D-optimal.*

The proofs of the theorems of this section are in the Appendix. We will apply Theorem 2 to several two-treatment problems. For these cases we define

$$p(\mathbf{x}_1) = \xi_{2|1}(t_1 | \mathbf{x}_1) = P(x_2 = t_1 | \mathbf{x}_1);$$

by Theorem 2 we are then to construct $p(\cdot)$ in such a way that

$$p(\mathbf{x}_1) = 1 \quad \Leftrightarrow \quad \mathbf{f}^t(\mathbf{x}_1, t_1) \mathbf{M}^{-1}(\xi^*) \mathbf{f}(\mathbf{x}_1, t_1) > \mathbf{f}^t(\mathbf{x}_1, t_0) \mathbf{M}^{-1}(\xi^*) \mathbf{f}(\mathbf{x}_1, t_0). \quad (8)$$

3.1. Linear regression through the origin.

In this first application x_1 is a univariate variable. We denote by μ_k the noncentral moments of x_1 , under the marginal G_1 :

$$\mu_k = E_{G_1}[X^k] = \int_{-\infty}^{\infty} x_1^k dG_1(x_1), \quad k = 1, 2,$$

and by σ^2 the variance $\mu_2 - \mu_1^2$. We take $t_k = k$, so that $\chi_2 = \{0, 1\}$ and

$$\mathbf{f}^t(x_1, x_2) = (x_1, x_2), \quad x_2 \in \chi_2.$$

From (1) the information matrix is

$$\begin{aligned} \mathbf{M}(\xi) &= \int_{\chi_1} \{ \mathbf{f}(x_1, 0) \mathbf{f}^t(x_1, 0) [1 - p(x_1)] + \mathbf{f}(x_1, 1) \mathbf{f}^t(x_1, 1) p(x_1) \} \xi_1(dx_1) \\ &= \begin{pmatrix} \mu_2 & E_{G_1}[Xp(X)] \\ E_{G_1}[Xp(X)] & E_{G_1}[p(X)] \end{pmatrix}. \end{aligned} \quad (9)$$

Then (8) becomes

$$p(x) = 1 \Leftrightarrow \mu_2 - 2x_1 E_{G_1}[Xp(X)] > 0. \quad (10)$$

We will show that the solution is one of

$$p_z(x_1) = I_{(-\infty, z) \cap \chi_1}(x_1) = \begin{cases} 1, & \text{if } x_1 < z \text{ and } x_1 \in \chi_1, \\ 0, & \text{otherwise,} \end{cases}$$

or

$$p^z(x_1) = I_{(z, \infty) \cap \chi_1}(x_1) = \begin{cases} 1, & \text{if } x_1 > z \text{ and } x_1 \in \chi_1, \\ 0, & \text{otherwise,} \end{cases}$$

for an appropriate value z^* of z . Note that we are not asserting that $z^* \in \chi_1$. For instance if $\chi_1 = (a, b)$ and $z^* > b$, then $p_{z^*}(x_1) = I_{\chi_1}(x_1)$.

THEOREM 3. Assume that $0 < \mu_2 < \infty$ and that $G_1(\cdot)$ is continuous. Define

$$\psi_+(z) = \mu_2 - 2z \int_{-\infty}^z x dG_1(x), \quad \psi_-(z) = \mu_2 - 2z \int_z^{\infty} x dG_1(x). \quad (11)$$

- (a) If $\mu_1 > 0$ define z^* to be the unique positive solution to $\psi_+(z) = 0$. Then the design with conditional $\xi_{2|1}^*(1|x_1) = p_{z^*}(x_1)$ is marginally restricted D-optimal.
- (b) If $\mu_1 < 0$ define z^* to be the unique negative solution to $\psi_-(z) = 0$. Then the design with conditional $\xi_{2|1}^*(1|x_1) = p^{z^*}(x_1)$ is marginally restricted D-optimal.
- (c) If $\mu_1 = 0$ then the design with conditional $\xi_{2|1}^*(1|x_1) = p_{\infty}(x_1) = p^{\infty}(x_1) = I_{\chi_1}(x_1)$ is marginally restricted D-optimal.

Example 3.1. The marginally restricted D-optimal design is invariant with respect to linear transformations of the regressors in the same way as are D-optimal designs (Fedorov 1972, Th. 2.2.4). The example considered here is regression through the origin. This means the optimal design is not invariant with respect to a translation of the design space. In fact the optimal design depends strongly on the position of the design space with respect to the origin. A simple example will illustrate this. Suppose that subjects with ages x_1 arrive for assignment to one of two treatment

groups ($x_2 \in \{0, 1\}$). If, say, the ages are uniformly distributed over $[20, 60]$, then $\mu_2 = 5200/3$ and (11) becomes

$$120 \cdot \psi_+(z) = 3z^3 - 1200z - 208000 = 0;$$

hence $z^* \approx 44.321$. Then all subjects of age $x_1 < z^*$ go to treatment 1; the others to treatment 0. On the other hand, if the ages are linearly transformed so as to range uniformly over $[0, 1]$, then $\mu_2 = 1/3$, $\psi_+(z) = \mu_2 - z^3$ and $z^* = 3^{-1/3}$. In the original units this gives $x_1 = 40 \cdot 3^{-1/3} + 20 = 47.734$ as the dividing age between the treatment groups. And, if the ages are coded to range uniformly over $[-1, 1]$, then $\mu_1 = 0$ and all assignments are to treatment 1. This, rather counterintuitive, feature of the design disappears if an intercept is included in the model—see Section 3.2, Section 3.3 below. Of course for this example, the optimal design is invariant with respect to changes in scale only.

3.2. Linear regression with an intercept.

Example 3.1 highlights the advisability of fitting a model with an intercept in problems which can be expected to be invariant under translations of the design space. Thus, in this section we take $\chi_2 = \{0, 1\}$ and

$$\mathbf{f}^t(\mathbf{x}_1, x_2) = (1, \mathbf{x}_1, x_2), \quad x_2 \in \chi_2.$$

We drop the restriction that \mathbf{x}_1 be univariate; it is now an arbitrary random vector. Let $\mathbf{z}_1 = (1, \mathbf{x}_1^t)^t$ and define

$$\begin{aligned} \mathbf{M}_0 &= \int_{\chi_1} \mathbf{z}_1 \mathbf{z}_1^t \xi_1(d\mathbf{x}_1) \stackrel{\text{def}}{=} \begin{pmatrix} 1 & \boldsymbol{\mu}_{\mathbf{x}}^t \\ \boldsymbol{\mu}_{\mathbf{x}} & \boldsymbol{\mu}_{\mathbf{x}\mathbf{x}} \end{pmatrix}, \\ \mathbf{m}_p &= \int_{\chi_1} \mathbf{z}_1 p(\mathbf{x}_1) \xi_1(d\mathbf{x}_1) \stackrel{\text{def}}{=} \begin{pmatrix} \mu_p \\ \boldsymbol{\mu}_{\mathbf{x}p} \end{pmatrix}. \end{aligned}$$

Then the information matrix is

$$\mathbf{M}(\xi) = \begin{pmatrix} \mathbf{M}_0 & \mathbf{m}_p \\ \mathbf{m}_p^t & \mu_p \end{pmatrix},$$

with

$$|\mathbf{M}(\xi)| = |\mathbf{M}_0| \{ \mu_p(1 - \mu_p) - (\boldsymbol{\mu}_{\mathbf{x}p} - \mu_p \boldsymbol{\mu}_{\mathbf{x}})^t (\boldsymbol{\mu}_{\mathbf{x}\mathbf{x}} - \boldsymbol{\mu}_{\mathbf{x}} \boldsymbol{\mu}_{\mathbf{x}}^t)^{-1} (\boldsymbol{\mu}_{\mathbf{x}p} - \mu_p \boldsymbol{\mu}_{\mathbf{x}}) \}.$$

The term $|\mathbf{M}_0|$ does not depend upon $p(\cdot)$, and that in braces remains the same if $p(\cdot)$ is replaced by $1 - p(\cdot)$. The choice $p^*(\mathbf{x}_1) \equiv 1/2$ both maximizes the first term in braces, and minimizes the second, which is nonnegative but vanishes under p^* . Thus the MR D-optimal design randomizes the assignment x_2 , for any \mathbf{x}_1 . Given our emphasis on nonadaptive designs, this is perhaps to be expected.

3.3. Regression with an intercept and interaction.

We again take $\chi_2 = \{0, 1\}$, but now

$$\mathbf{f}^t(\mathbf{x}_1, x_2) = (1, \mathbf{x}_1^t, x_2, x_2 \mathbf{x}_1^t) = (\mathbf{z}_1^t, x_2 \mathbf{z}_1^t), \quad x_2 \in \chi_2.$$

The information matrix is

$$\mathbf{M}(\xi) = \begin{pmatrix} \mathbf{M}_0 & \mathbf{M}_p \\ \mathbf{M}_p & \mathbf{M}_p \end{pmatrix},$$

where \mathbf{M}_0 is as in Section 3.2 and

$$\mathbf{M}_p = \int_{\chi_1} \mathbf{z}_1 \mathbf{z}_1^t p(\mathbf{x}_1) \xi_1(d\mathbf{x}_1).$$

Then by a standard formula for the determinant of a partitioned matrix,

$$\begin{aligned} |\mathbf{M}(\xi)| &= |\mathbf{M}_p| |\mathbf{M}_0 - \mathbf{M}_p \mathbf{M}_p^{-1} \mathbf{M}_p| = |\mathbf{M}_p| |\mathbf{M}_0 - \mathbf{M}_p| \\ &= \det \int_{\mathcal{X}_1} \mathbf{z}_1 \mathbf{z}_1^t p(\mathbf{x}_1) \xi_1(d\mathbf{x}_1) \cdot \det \int_{\mathcal{X}_1} \mathbf{z}_1 \mathbf{z}_1^t (1 - p(\mathbf{x}_1)) \xi_1(d\mathbf{x}_1). \end{aligned}$$

Thus both $p(\cdot)$ and $(1 - p(\cdot))$ result in the same value of $|\mathbf{M}(\xi)|$. Since $\log |\mathbf{M}(\xi)|$ is a concave functional of $p(\cdot)$ it is increased by averaging $p(\cdot)$ and $(1 - p(\cdot))$; hence the maximizer is necessarily $p^*(\mathbf{x}_1) \equiv 1/2$. Again, the optimal procedure is to randomize the assignments x_2 , for any \mathbf{x}_1 .

4. CASE STUDY

This article was motivated by an example in López-Fidalgo & Garcet-Rodríguez (2004). They proposed an optimal experimental design for prediction of morbidity after lung resection. Here we reconsider this problem, using the techniques developed in this article. We take “percentage of maximum volume of expired air” (EA) and “decrease in blood oxygen concentration during exercise” (BC) as the components of a continuous variable \mathbf{x}_1 and “type of surgery” as the variable x_2 with two possible values, i.e. $\mathcal{X}_2 = \{0, 1\}$. Then the response variable Y is the morbidity of the patient after the lung resection. The design will consist of assigning a type of surgery according to the value of \mathbf{x}_1 for a particular patient.

The logistic model $E[Y | \mathbf{x}_1, x_2] = F(\mathbf{z}^t(\mathbf{x}_1, x_2)\boldsymbol{\gamma})$, where $F(t) = (1 + e^{-t})^{-1}$ is the logistic cumulative distribution function and $\mathbf{z}(\mathbf{x}_1, x_2)$ is a vector of regressors, is frequently used in this context. The Fisher information matrix may be computed as follows. The log-likelihood for a particular observation is

$$L = y \log F(\mathbf{z}^t(\mathbf{x}_1, x_2)\boldsymbol{\gamma}) + (1 - y) \log [1 - F(\mathbf{z}^t(\mathbf{x}_1, x_2)\boldsymbol{\gamma})].$$

Then with $w(t) = F(t)[1 - F(t)] = \text{sech}^2(t)/4$, the information matrix for a particular point is

$$I(\mathbf{x}_1, x_2, \boldsymbol{\gamma}) = E_Y \left[\frac{\partial L}{\partial \boldsymbol{\gamma}} \frac{\partial L}{\partial \boldsymbol{\gamma}^t} \right] = w(\mathbf{z}^t(\mathbf{x}_1, x_2)\boldsymbol{\gamma}) \mathbf{z}(\mathbf{x}_1, x_2) \mathbf{z}^t(\mathbf{x}_1, x_2).$$

Therefore the optimal design problem may be considered as for a typical linear model, but with

$$\mathbf{f}(\mathbf{x}_1, x_2, \boldsymbol{\gamma}) = \sqrt{w(\mathbf{z}^t(\mathbf{x}_1, x_2)\boldsymbol{\gamma})} \mathbf{z}(\mathbf{x}_1, x_2)$$

depending on the unknown parameters.

In López-Fidalgo & Garcet-Rodríguez (2004), the response contained main effects only, and so we first consider the regression function $\mathbf{z}^t(x_1, x_2)\boldsymbol{\gamma} = x_1\gamma_1 + x_2\gamma_2$ with x_1 denoting EA. Then (4) becomes

$$d(x_1, x_2, \boldsymbol{\gamma}) = w(x_1\gamma_1 + \gamma_2 x_2) \frac{Cx_1^2 - 2Bx_1x_2 + Ax_2^2}{AC - B^2},$$

where

$$\begin{aligned} A &= E_{\xi_1} [x_1^2 \{w(x_1\gamma_1)(1 - p(x_1)) + w(x_1\gamma_1 + \gamma_2)p(x_1)\}], \\ B &= E_{\xi_1} [x_1 w(x_1\gamma_1 + \gamma_2)p(x_1)], \\ C &= E_{\xi_1} [w(x_1\gamma_1 + \gamma_2)p(x_1)]. \end{aligned}$$

Thus

$$d(x_1, 1; \xi) > d(x_1, 0; \xi) \quad \Leftrightarrow \quad w(x_1\gamma_1 + \gamma_2)(Cx_1^2 - 2Bx_1 + A) > w(x_1\gamma_1)Cx_1^2. \quad (12)$$

Taking into account retrospective analysis (Varela, Cordovilla, Jiménez & Novoa 2001) we assumed for x_1 a normal distribution with mean $\mu = 70$, standard deviation $\sigma = 7$ and nominal values of the parameters $\gamma_1 = 0.01$ and $\gamma_2 = 1$. Motivated by Theorem 2 we were led to conjecture that a design with $\xi_{2|1}^*(1|x_1) = p_{z^*}(x_1)$ for some z^* would satisfy (12). We then computed $z^* = 70.4794$ and checked (5): with $\phi(\cdot; \mu, \sigma^2 r)$ denoting the normal density we obtained

$$\int_{-\infty}^{z^*} d(x_1, 1, \gamma) \phi(x_1; \mu, \sigma^2) dx_1 + \int_{z^*}^{\infty} d(x_1, 0, \gamma) \phi(x_1; \mu, \sigma^2) dx_1 = 2.$$

Thus, the optimal design consists of applying the first treatment to the patients with EA under 70.5% and the second to the rest of the patients.

Comments similar to those in Example 3.1 apply here as well—this design may introduce a confounding between treatment effect and EA. Thus we now consider the regression function

$$\mathbf{z}^t(\mathbf{x}_1, x_2)\gamma = (1, \mathbf{x}_1^t, x_2, x_2\mathbf{x}_1^t)\gamma = \mathbf{z}_1^t\gamma_1 + x_2\mathbf{z}_1^t\gamma_2,$$

where $\mathbf{z}_1^t = (1, \mathbf{x}_1^t)$ as in Section 3.3 and $\mathbf{x}_1^t = (\text{EA}, \text{BC})$. We include interactions and an intercept as well as main effects. As at (9),

$$\mathbf{M}(\xi) = \begin{pmatrix} \mathbf{L}_p + \mathbf{M}_p & \mathbf{M}_p \\ \mathbf{M}_p & \mathbf{M}_p \end{pmatrix},$$

where

$$\begin{aligned} \mathbf{L}_p &= \int_{\chi_1} \mathbf{z}_1 \mathbf{z}_1^t w(\mathbf{z}_1^t \gamma_1) (1 - p(\mathbf{x}_1)) \xi_1(d\mathbf{x}_1), \\ \mathbf{M}_p &= \int_{\chi_1} \mathbf{z}_1 \mathbf{z}_1^t w(\mathbf{z}_1^t \gamma_1 + \mathbf{z}_1^t \gamma_2) p(\mathbf{x}_1) \xi_1(d\mathbf{x}_1). \end{aligned}$$

We find that

$$\begin{aligned} |\mathbf{M}(\xi)| &= |\mathbf{M}_p| |\mathbf{L}_p|, \\ \mathbf{M}^{-1}(\xi) &= \begin{pmatrix} \mathbf{L}_p^{-1} & -\mathbf{L}_p^{-1} \\ -\mathbf{L}_p^{-1} & \mathbf{L}_p^{-1} + \mathbf{M}_p^{-1} \end{pmatrix}, \\ d(\mathbf{x}_1, x_2; \xi) &= w(\mathbf{z}_1^t \gamma_1 + x_2 \mathbf{z}_1^t \gamma_2) \{ (1 - x_2)^2 \mathbf{z}_1^t \mathbf{L}_p^{-1} \mathbf{z}_1 + x_2^2 \mathbf{z}_1^t \mathbf{M}_p^{-1} \mathbf{z}_1 \}. \end{aligned}$$

Then with

$$\psi(\mathbf{x}_1) = \mathbf{z}_1^t [w(\mathbf{z}_1^t \gamma_1 + \mathbf{z}_1^t \gamma_2) \mathbf{M}_p^{-1} - w(\mathbf{z}_1^t \gamma_1) \mathbf{L}_p^{-1}] \mathbf{z}_1,$$

we have that

$$d(\mathbf{x}_1, 1; \xi) > d(\mathbf{x}_1, 0; \xi) \Leftrightarrow \psi(\mathbf{x}_1) > 0.$$

Any exact, but necessarily local, solution will now depend on assumed values of γ_1 and γ_2 in a very complicated manner. Such a solution requires the construction of a set $S(\gamma_1, \gamma_2; p)$ satisfying

$$S(\gamma_1, \gamma_2; p) = \{\mathbf{x}_1 \mid \psi(\mathbf{x}_1) > 0\} \quad \text{and} \quad p(\mathbf{x}_1) = I(\mathbf{x}_1 \in S(\gamma_1, \gamma_2; p)).$$

But there is a simple solution, which is optimal when $\gamma_2 = \mathbf{0}$. The interpretation of this condition is that the model allows for a main effect and interactions but in fact these effects are absent. Under this condition $\mathbf{M}_p = \mathbf{L}_{1-p}$, and it follows as in Section 3.3 that then $p^*(\mathbf{x}_1) \equiv 1/2$, i.e. randomization of the treatment assignments is locally optimal.

5. SUMMARY

In this article D-optimal designs have been provided when some of the factors cannot be controlled but their values are known before the experiment is carried out, while the remaining factor is under the control of the experimenter. We have called these designs marginally restricted (MR) designs after Cook & Thibodeau (1980).

The sequential procedure proposed in Section 2 of this article was motivated by an example from López-Fidalgo & Garcet-Rodríguez (2004). Assuming that we have calculated an optimal design for a group of patients and then a new patient arrives, the marginal design changes. We want to assign a treatment in such a way that the corresponding joint design in each step will be optimal. In this sequential procedure the previous design is used in order to determine how the next run should be conducted. Instead, in Section 3 and Section 4, we have proposed an a priori locally optimal strategy, according to which the assignments are randomized.

In general, a theoretical distribution may be assumed from the beginning instead of changing the marginal empirical distribution when a new unit appears in the study. Under this framework, there is no need for computations every time a new unit joins the sample. If the marginal distribution is discrete, with only a few points in its support, then conditional optimal designs may be computed for each of these points in the usual way. If the distribution is continuous, this procedure can no longer be applied. For the K -treatment design problem considered in this article, the objective is to find a function defined on a continuous space, typically an interval, optimizing the criterion (2). Solutions to this problem have been provided to this particular case with $K = 2$ and several linear regression models.

APPENDIX

Proof of Theorem 1. Let ξ be a particular design. Since the information matrix is symmetric and positive definite, there exists a nonsingular matrix \mathbf{U} such that $\mathbf{U}\mathbf{M}(\xi)\mathbf{U}^t = \mathbf{I}$. Let ξ^{MR} be an MR D-optimal design.

Let us consider a new model whose regression functions are $\varphi(\mathbf{x}) = \mathbf{U}f(\mathbf{x})$, $\forall \mathbf{x} \in \chi$. The information matrix of any design η for this new model is $\mathbf{M}_\varphi(\eta) = \mathbf{U}\mathbf{M}(\eta)\mathbf{U}^t$ and $\det \mathbf{M}_\varphi(\eta) = (\det \mathbf{U})^2 \det \mathbf{M}(\eta)$. The dispersion function and the efficiency are invariant to linear transformations:

$$\begin{aligned} d_\varphi(\mathbf{x}, \eta) &= \varphi^t(\mathbf{x})\mathbf{M}^{-1}(\eta)\varphi(\mathbf{x}) = d(\mathbf{x}, \eta), \\ \frac{\det \mathbf{M}(\eta)}{\det \mathbf{M}(\xi^{\text{MR}})} &= \frac{\det \mathbf{M}_\varphi(\eta)}{\det \mathbf{M}_\varphi(\xi^{\text{MR}})}. \end{aligned}$$

Since $\mathbf{M}_\varphi(\xi) = \mathbf{I}$, it can be assumed that $\mathbf{M}(\xi) = \mathbf{I}$. Then

$$\begin{aligned} \int_{\chi_1} \max_{x_2} d(\mathbf{x}_1, x_2; \xi) \xi_1(d\mathbf{x}_1) &\geq \int_{\chi} d(\mathbf{x}_1, x_2; \xi) \xi_{2|1}^{\text{MR}}(dx_2 | \mathbf{x}_1) \xi_1(d\mathbf{x}_1) \\ &= \int_{\chi} d(\mathbf{x}_1, x_2; \xi) \xi^{\text{MR}}(dx) \\ &= \text{tr} \left[\int_{\chi} \mathbf{f}(\mathbf{x}_1, x_2) \mathbf{f}^t(\mathbf{x}_1, x_2) \xi^{\text{MR}}(d\mathbf{x}) \right] \\ &= \text{tr}[\mathbf{M}(\xi^{\text{MR}})]. \end{aligned}$$

Thus

$$\begin{aligned} \text{Eff}_{D-\text{MR}}(\xi) &= \left(\frac{\det \mathbf{M}(\xi)}{\det \mathbf{M}(\xi^{\text{MR}})} \right)^{1/m} \\ &= \left(\frac{1}{\det \mathbf{M}(\xi^{\text{MR}})} \right)^{1/m} \end{aligned}$$

$$\begin{aligned}
&\geq \frac{m}{\text{tr}[\mathbf{M}(\xi^{\text{MR}})]} \\
&\geq \frac{m}{\int_{\chi_1} \max_{x_2} d(\mathbf{x}_1, x_2; \xi) \xi_1^{\text{MR}}(d\mathbf{x}_1)}. \quad \square
\end{aligned}$$

Proof of Theorem 2. First assume that there are no multiple maxima. Define a partition $\{\chi_{1,k}\}_{k=0}^{K-1}$ of χ_1 by $\chi_{1,k} = \{\mathbf{x}_1 \in \chi_1 \mid p_k(\mathbf{x}_1) = 1\}$. Note that

$$\begin{aligned}
\mathbf{M}(\xi^*) &= \int_{\chi_1} \int_{\chi_2} \mathbf{f}(\mathbf{x}_1, x_2) \mathbf{f}^t(\mathbf{x}_1, x_2) \xi_{2|1}^*(x_2 \mid \mathbf{x}_1) \xi_1(d\mathbf{x}_1) \\
&= \sum_{k=0}^{K-1} \int_{\chi_1} \mathbf{f}(\mathbf{x}_1, t_k) \mathbf{f}^t(\mathbf{x}_1, t_k) p_k(\mathbf{x}_1) \xi_1(d\mathbf{x}_1) \quad (13)
\end{aligned}$$

$$= \sum_{k=0}^{K-1} \int_{\chi_{1,k}} \mathbf{f}(\mathbf{x}_1, t_k) \mathbf{f}^t(\mathbf{x}_1, t_k) \xi_1(d\mathbf{x}_1). \quad (14)$$

Now

$$\begin{aligned}
\int_{\chi_1} \max_{x_2} d(\mathbf{x}_1, x_2; \xi^*) \xi_1(d\mathbf{x}_1) &= \sum_{k=0}^{K-1} \int_{\chi_{1,k}} d(\mathbf{x}_1, t_k; \xi^*) \xi_1(d\mathbf{x}_1) \\
&= \sum_{k=0}^{K-1} \int_{\chi_{1,k}} \mathbf{f}^t(\mathbf{x}_1, t_k) \mathbf{M}^{-1}(\xi^*) \mathbf{f}(\mathbf{x}_1, t_k) \xi_1(d\mathbf{x}_1) \quad (15) \\
&= \text{tr} \left[\mathbf{M}^{-1}(\xi^*) \sum_{k=0}^{K-1} \int_{\chi_{1,k}} \mathbf{f}(\mathbf{x}_1, t_k) \mathbf{f}^t(\mathbf{x}_1, t_k) \xi_1(d\mathbf{x}_1) \right] \\
&= m,
\end{aligned}$$

by (14).

In the case of ties, i.e., $d(\mathbf{x}_1, t_{k_1}; \xi^*) = \dots = d(\mathbf{x}_1, t_{k_j}; \xi^*) = \max_{x_2 \in \chi_2} d(\mathbf{x}_1, x_2; \xi^*)$ with $k_1 < \dots < k_j$, then the assignment of values j^{-1} to each $p_{k_i}(\mathbf{x}_1)$ has the same effect as if one were to adopt instead the convention of setting $p_{k_1}(\mathbf{x}_1) = 1$ and $p_{k_2}(\mathbf{x}_1) = \dots = p_{k_j}(\mathbf{x}_1) = 0$. But this leaves both (13) and (15) fixed, so that again (5) holds. \square

Proof of Theorem 3. Assertion c) is immediate: if $\mu_1 = 0$, then $\det[\mathbf{M}(\xi_\infty)] = \mu_2 \geq \det[\mathbf{M}(\xi)]$ for any design ξ . Assume now that $\mu_1 > 0$. Consider the existence and uniqueness of z^* defined by $\psi_+(z^*) = 0$. We have $\psi_+(0) = \mu_2 > 0$, and $\psi_+(\infty) = -\infty$ (here we use $\mu_1 > 0$). Thus, there exists a positive root z^* . At such a point

$$E_{G_1}[X p_{z^*}(X)] = \int_{-\infty}^{z^*} x dG_1(x) = \mu_2 / (2z^*) > 0,$$

so that $\psi_+(z)$ is strictly decreasing in a neighbourhood of any such root, which is therefore unique.

Now note that for $x_1 \in \chi_1$,

$$p_{z^*}(x_1) = 1 \Leftrightarrow x_1 < z^* = \frac{\mu_2}{2E_{G_1}[X p_{z^*}(X)]} \Leftrightarrow \mu_2 - 2x_1 E_{G_1}[X p_{z^*}(X)] > 0;$$

thus (10) holds and assertion (a) follows from Theorem 2. The case $\mu_1 < 0$ is completely symmetric. \square

ACKNOWLEDGEMENTS

This work has been supported by Junta de Comunidades de Castilla-La Mancha and Ministerio de Educación y Ciencia. The work of Douglas P. Wiens is supported by the Natural Sciences and Engineering Research Council of Canada. The helpful comments of an anonymous referee are gratefully acknowledged.

REFERENCES

- C. L. Atwood (1969). Optimal and efficient designs of experiments. *The Annals of Mathematical Statistics*, 40, 1570–1602.
- C. S. Cheng (1987). An application of the Kiefer–Wolfowitz equivalence theorem to a problem in Hadamard transform optics. *The Annals of Statistics*, 15, 1593–1603.
- R. D. Cook & L. A. Thibodeau (1980). Marginally restricted D -optimal designs. *Journal of the American Statistical Association*, 75, 366–371.
- H. Dette (1996). Lower bound for efficiencies with applications. In *Research Developments in Probability and Statistics: Festschrift in Honor of Madan L. Puri on the Occasion of his 65th Birthday* (E. Brunner & M. Denker, eds.), VSP, Utrecht, pp. 111–124.
- H. Dette & W. J. Studden (1993). Geometry of E -optimality. *The Annals of Statistics*, 21, 416–433.
- V. V. Fedorov (1972). *Theory of Optimal Experiments*. Academic Press, New York.
- N. Flournoy & W. F. Rosenberger, eds. (1995). *Adaptive Designs: Selected Proceedings of a 1992 Joint AMS–IMS–SIAM Summer Conference*. IMS Lecture Notes–Monograph Series 25, Institute of Mathematical Statistics, Hayward, California.
- R. Harman (2004). Lower bound on efficiency ratios based on ϕ_p -optimal designs. In *MODa7, Advances in Model-Oriented Design and Analysis: Proceedings of the 7th International Workshop on Model-Oriented Design and Analysis held in Heeze, The Netherlands, June 14–18, 2004* (A. Di Bucchianico, H. Läuter & H. P. Wynn, eds.), Physica-Verlag, Heidelberg, pp. 89–96.
- M. N. L. Huang & M. C. Hsu (1993). Marginally restricted linear-optimal designs. *Journal of Statistical Planning and Inference*, 35, 251–266.
- J. López-Fidalgo & S. Garcet-Rodríguez (2004). Optimal experimental designs when some independent variables are not subject to control. *Journal of the American Statistical Association*, 99, 1190–1199.
- R. Martín-Martín, B. Torsney & J. López-Fidalgo (2007). Construction of marginally and conditionally restricted designs using multiplicative algorithms. *Computational Statistics & Data Analysis*, 51, 5547–5561.
- C. J. Nachtsheim (1989). On the design of experiments in the presence of fixed covariates. *Journal of Statistical Planning and Inference*, 22, 203–212.
- G. Varela, R. Cordovilla, M. F. Jiménez & N. Novoa (2001). Utility of standardized exercise oximetry to predict cardiopulmonary morbidity after lung resection. *European Journal of Cardio-Thoracic Surgery*, 19, 351–354.

Received 10 November 2007

Accepted 3 June 2008

Jesús LÓPEZ-FIDALGO: jesus.lopezfidalgo@uclm.es
 Raul MARTÍN-MARTÍN: Raul.MMartin@uclm.es
Instituto de Matemática Aplicada a la Ciencia y la Ingeniería
Universidad de Castilla-La Mancha
ES-13071 Ciudad Real, Spain

Douglas P. WIENS: doug.wiens@ualberta.ca
Department of Mathematical and Statistical Sciences
University of Alberta, Edmonton, Alberta
Canada T6G 2G1