Chapter 1

A NOTE ON OPTIMAL BOUNDED DESIGNS

Michael Sahm
University of Heidelberg, Institute for Applied Mathematics,
Im Neuenheimer Feld 294, D-69 120 Heidelberg, Germany
msahm@statlab.uni-heidelberg.de

Rainer Schwabe
University of Tübingen, Institute of Medical Biometry,
Westbahnhofstr. 55, D-72 070 Tübingen, Germany
rainer.schwabe@uni-tuebingen.de

Abstract An equivalence theorem is formulated to characterize optimal designs with general prespecified direct constraints on the design intensity. As an application locally optimal designs are obtained with bounded intensity for the logistic regression model. Moreover, it is shown that for additive models optimal marginally bounded designs can be generated from their optimal counterparts in the corresponding marginal models.

Keywords: constrained design, equivalence theorem, optimal design, logistic regression, bounded design intensity, additive model, marginal design

Dedicated to Professor Valerii V. Fedorov
on occasion of his 60th birthday

Introduction

In many applications constraints are given on the possible settings for a design variable either by the availabilty of experimental units with given properties or by the need to include certain preselected subjects or sites in the study. The first constraint of availability provides an upper bound for the designs under consideration, whereas the second constraint of prespecified settings results in a lower bound for the designs.

For the continuous case where the design intensity is bounded by an atom free measure from above equivalence theorems have been propagated by Wynn (1982) and Fedorov (1989). More recent references on this topic are Cook and Fedorov (1995) and Fedorov and Hackl (1997). For particular covariate models Wierich (1985) obtained characterizations for the finite sample optimal design when both upper and lower bounds are given for the competing designs.

In the present note we formulate an equivalence theorem for the situation of general constraints given by upper and lower bounds for the design intensity. These constraints can be specified as arbitrary measures and, hence, both the continuous case of Wynn and Fedorov and the discrete case of Wierich are simultaneously covered. As a nonstandard application we obtain locally optimal designs with bounded intensity for the logistic regression model. Finally, optimal marginally bounded designs are constructed for additive models as a product of the optimal counterparts in the corresponding marginal models.

1. GENERAL SETTING

We consider the general linear model

$$Y(x) = f(x)^\top \beta + Z(x),$$

where f is a set of known regression functions and β is the vector of unknown parameters. Repeated experiments result in uncorrelated and homoscedastic errors Z.

Due to the generality of the problem we restrict our attention to generalized designs ξ which are defined as probability measures on the design region $\mathcal{X}$ of possible settings for the design variable x. The set of all generalized designs on $\mathcal{X}$ will be denoted by Ξ. For the definition and properties of generalized (approximate) designs we refer to the seminal paper by Kiefer (1959) and the monograph by Fedorov (1972).

The quality of a design ξ is usually measured by its information matrix $M(\xi) = \int_{\mathcal{X}} f(x)f(x)^\top \xi(dx)$ which is essentially proportional to the inverse of the covariance matrix for the best linear unbiased estimator of the parameter vector β. For the comparison of different designs we make use of so–called design criteria Φ which are convex or pseudo–convex functions of ξ based on the corresponding information matrices $M(\xi)$. The most popular criterion is that of D–optimality which aims at minimizing the determinant of the inverse $M(\xi)^{-1}$ of the information matrix. This approach minimizes the volume of the confidence ellipsoid in case of normal errors. Other convex criteria include those based on the variance function $f(x)^\top M(\xi)^{-1} f(x)$ of the prediction $\widehat{Y}(x)$ of the response $f(x)^\top \beta$ for various x (IMSE and minimax criteria).

A powerful tool for investigating convex criteria Φ is the concept of directional derivatives. The quantity $F_\Phi(\xi, \eta) = \lim_{\alpha \searrow 0} \frac{1}{\alpha}(\Phi(\alpha\eta + (1-\alpha)\xi) -$

$\Phi(\xi))$ is called the directional derivative of Φ at ξ in the direction of η. Denote by $F_\Phi(\xi; x) = F_\Phi(\xi, \delta_x)$ the directional derivative in the direction of the one-point (Dirac) measure δ_x. Then Φ is called differentiable if $F_\Phi(\xi, \eta) = \int_\mathcal{X} F_\Phi(\xi; x)\, \eta(dx)$.

Let Ξ' be a convex set of competing designs, $\Xi' \subset \Xi$. A by now well–known consequence of the theory of convex analysis is the following theorem by Whittle (1973):

Theorem A.

(i) *ξ^* minimizes $\Phi(\xi)$ in Ξ' if and only if $F_\Phi(\xi^*, \eta) \geq 0$ for every design $\eta \in \Xi'$.*

(ii) *If Φ is differentiable, then ξ^* minimizes $\Phi(\xi)$ in Ξ if and only if $F_\Phi(\xi^*; x) \geq 0$ for every setting $x \in \mathcal{X}$. Moreover, if ξ^* minimizes $\Phi(\xi)$, then $F_\Phi(\xi^*; x) = 0$ for ξ^*–almost all x.*

2. AN EQUIVALENCE THEOREM FOR BOUNDED DESIGNS

We are interested in the general situation of explicitly specified upper and lower bounds μ and ν for the competing designs, $\nu \leq \xi \leq \mu$, where μ and ν are arbitrary measures on the design region $\mathcal{X}$ satisfying the natural conditions $\nu \leq \mu$ and $\nu(\mathcal{X}) \leq 1 \leq \mu(\mathcal{X})$. Then the set $\Xi_\nu^\mu = \{\xi;\, \nu \leq \xi \leq \mu\}$ of competing designs is non–empty and convex. In the following we will only consider differentiable design criteria.

Theorem 1. *The following three statements are equivalent:*

(a) ξ^* minimizes $\Phi(\xi)$ in Ξ_ν^μ.

(b) There exists a number s such that $F_\Phi(\xi^*; x) \leq s$ for $(\xi^* - \nu)$–almost all x and, simultaneously, $F_\Phi(\xi^*; x) \geq s$ for $(\mu - \xi^*)$–almost all x.

(c) There are subsets $\mathcal{X}_1, \mathcal{X}_2 \subset \mathcal{X}$ and a number s such that

(i) $\xi^* = \nu$ *on* $\mathcal{X}_1$ *and* $\xi^* = \mu$ *on* $\mathcal{X}_2$,

(ii) $\sup_{x\in\mathcal{X}_2} F_\Phi(\xi^*; x) \leq s \leq \inf_{x\in\mathcal{X}_1} F_\Phi(\xi^*; x)$,

(iii) $F_\Phi(\xi^*; x) = s$ *on* $\mathcal{X} \setminus (\mathcal{X}_1 \cup \mathcal{X}_2)$.

The proof of the theorem will be given in an appendix. Statement (b) of Theorem 1 may be rewritten in terms of essential suprema and essential infima as $\operatorname{ess\,sup}_{\xi^*-\nu} \mathrm{F}_\Phi(\xi^*; \mathrm{x}) \leq \operatorname{ess\,inf}_{\mu-\xi^*} \mathrm{F}_\Phi(\xi^*; \mathrm{x})$. If we let $\nu = 0$, $\Xi^\mu = \{\xi;\, \xi \leq \mu\}$, we obtain a generalization of the theorem by Wynn (1982) and Fedorov (1989) for optimal designs bounded from above:

Corollary 1. *The following three statements are equivalent:*

(a) ξ^* minimizes $\Phi(\xi)$ in Ξ^μ.

(b) There exists a number s such that $F_\Phi(\xi^*; x) \leq s$ for ξ^*–almost all x

and, simultaneously, $F_\Phi(\xi^*; x) \geq s$ *for* $(\mu - \xi^*)$*–almost all* x.

(c) There are subsets $\mathcal{X}_1, \mathcal{X}_2 \subset \mathcal{X}$ *and a number* s *such that*

(i) $\xi^* = 0$ *on* $\mathcal{X}_1$ *and* $\xi^* = \mu$ *on* $\mathcal{X}_2$,

(ii) $\sup_{x\in\mathcal{X}_2} F_\Phi(\xi^*; x) \leq s \leq \inf_{x\in\mathcal{X}_1} F_\Phi(\xi^*; x)$,

(iii) $F_\Phi(\xi^*; x) = s$ *on* $\mathcal{X} \setminus (\mathcal{X}_1 \cup \mathcal{X}_2)$.

For optimal designs bounded from below, $\xi^* \in \Xi_\nu = \{\xi;\, \xi \geq \nu\}$, a slightly more delicate consideration yields the following result which combines features of Theorem 1 and part (ii) of Theorem A.

Corollary 2. *The following three statements are equivalent:*

(a) ξ^* *minimizes* $\Phi(\xi)$ *in* Ξ_ν.

(b) There exists a number s *such that* $F_\Phi(\xi^*; x) = s$ *for* $(\xi^* - \nu)$*–almost all* x *and, simultaneously,* $F_\Phi(\xi^*; x) \geq s$ *for all* $x \in \mathcal{X}$.

(c) There is a subset $\mathcal{X}_1 \subset \mathcal{X}$ *and a number* s *such that*

(i) $\xi^* = \nu$ *on* $\mathcal{X}_1$,

(ii) $\inf_{x\in\mathcal{X}_1} F_\Phi(\xi^*; x) \geq s$,

(iii) $F_\Phi(\xi^*; x) = s$ *on* $\mathcal{X} \setminus \mathcal{X}_1$.

In fact, letting $\nu = 0$ in Corollary 2 we get a reformulation of Whittle's Theorem A (ii) where $s = 0$. The main feature of the above results is that the design region can be split up into two subsets $\mathcal{X}_1$ and $\mathcal{X}_2$ for which the optimal design ξ^* coincides either with the lower bound or with the upper bound and for which the directional derivative at ξ^* is constant on the remaining set $\mathcal{X} \setminus (\mathcal{X}_1 \cup \mathcal{X}_2)$. In particular, in the continuous case this remaining set is typically a null set with respect to μ and we obtain a zero–one law in the sense that the optimal design ξ^* coincides with either the lower or upper bound.

3. APPLICATION: BOUNDED DESIGNS FOR LOGISTIC REGRESSION

As an illustration of the equivalence theorem we present an application to a logistic regression model where the intensity is bounded. As in all nonlinear regression models the information matrix depends on the unknown parameter β, and we consider locally optimal designs which solve the optimization problem for specified values of β. Note that the concept of local optimality can be understood as a local linearization of the model. Hence, linear methods can be applied.

In particular, we consider the binary response model

$$P_x(Y = 1) \;=\; F(b(x-a)) \;=\; \frac{1}{1 + e^{-b(x-a)}}$$

where the probability of a positive response at dose x is modeled by a logistic regression. Here the dose x may be given either on a linear or a logarithmic scale. The parameter $\beta = (a, b)$ consists of the effective mean dose (ED50) a and the slope $b > 0$, and F is the standard logistic distribution function $F(x) = (1 + e^{-x})^{-1}$. For a design ξ the asymptotic variance of the maximum likelihood estimator for β is proportional to the inverse of the Fisher information matrix

$$M(\xi;\beta) \;=\; \int_{I\!R} \begin{pmatrix} b^2 & -b\,(x-a) \\ -b\,(x-a) & (x-a)^2 \end{pmatrix} h(b(x-a))\,\xi(dx)$$

where $h(x) = F'(x)^2/(F(x)(1-F(x)) = F(x)(1-F(x))$.

For the D–criterion the derivative of the criterion function is given by

$$F_{\Phi,\beta}(\xi;x) \;=\; 2 - h(b(x-a))\,((-b,(x-a))M(\xi;\beta)^{-1}(-b,(x-a))^\top\,.$$

We consider the situation where the competing designs have a bounded density. To be more specific, we want to determine the D–optimal designs in Ξ^μ, where the upper bound μ is given by the density $g(x) = c \cdot 1_{[D_1,D_2]}(x)$ with respect to the Lebesgue measure, i. e. $g(x) = c$ for doses x which lie between a minimal possible dose D_1 and the maximal applicable dose D_2, $c(D_2 - D_1) > 1$, and $g(x) = 0$ otherwise. Usually, $D_1 = 0$ if the dose x is measured on a linear scale and $D_1 = -\infty$ if x is measured on a logarithmic scale. In total we have a compound constraint: neither too high nor too low doses are possible and, additionally, neither are replications of 'similar' doses.

By Corollary 1 there are disjoint sets $\mathcal{X}_1$ and $\mathcal{X}_2$ such that the density g^* of the optimal design ξ^* is equal to 0 on $\mathcal{X}_1$ and is equal to c on $\mathcal{X}_2$. Furthermore, there is a number s such that $\sup_{x\in\mathcal{X}_2} F_{\Phi,\beta}(\xi^*;x) \le s \le \inf_{x\in\mathcal{X}_1} F_{\Phi,\beta}(\xi^*;x)$ and $F_{\Phi,\beta}(\xi^*;x) = s$ on $\mathcal{X} \setminus (\mathcal{X}_1 \cup \mathcal{X}_2)$.

As for unrestricted optimal designs in regression models, a powerful tool is the investigation of the number of possible modes of the directional derivative $F_{\Phi,\beta}(\xi^*;x)$.

Proposition. *In the present setting, for every design ξ the directional derivative $F_{\Phi,\beta}(\xi^*;x)$ has ,at most, three local extrema as a function of x.*

This assertion is proved essentially by showing that the derivative of the function $u(x) = F_\Phi(\xi^*;x)$ has at most three changes of sign which, in turn, is achieved by expressing $u'(x)$ as a power series and bounding the number of sign changes of the coefficients. A similar result for the double exponential model can be found in Dette and Sahm (1997). The proposition is also related to the geometric condition in Sitter and Wu (1993).

The proposition, together with the fact that $\lim_{x\to\pm\infty} F_{\Phi,\beta}(\xi;x) = 2$, implies that $\mathcal{X} \setminus (\mathcal{X}_1 \cup \mathcal{X}_2)$ consists of not more than four points and, hence, the density g^* of the optimal design is constant to c on, at most, two intervals with total length $1/c$.

Result. *If the constraint μ is given by the density $g(x) = c \cdot 1_{[D_1,D_2]}(x)$ then, in a logistic regression model, the locally D–optimal design ξ^* is given by the density g^* which is either of the form*

(i) $g^*(x) = c \cdot 1_{[d_1,d_2]}(x)$, *where* $D_1 \leq d_1 < d_2 \leq D_2$, $d_2 - d_1 = 1/c$ *and* $F_{\Phi,\beta}(\xi^*; x) \leq \max(F_{\Phi,\beta}(\xi^*; d_1), F_{\Phi,\beta}(\xi^*; d_2))$ *for* $d_1 \leq x \leq d_2$. *Moreover,* $F_{\Phi,\beta}(\xi^*; d_1) = F_{\Phi,\beta}(\xi^*; d_2)$ *in case* $D_1 < d_1$ *and* $d_2 < D_2$.

(ii) $g^*(x) = c \cdot 1_{[d_1,d_2]}(x) + c \cdot 1_{[d_3,d_4]}(x)$ *where* $D_1 \leq d_1 < d_2 < d_3 < d_4 \leq D_2$, $d_2 - d_1 + d_4 - d_3 = 1/c$ *and* $F_\Phi(\xi^*, d_2) = F_\Phi(\xi^*, d_3) \geq \max(F_{\Phi,\beta}(\xi^*; d_1), F_{\Phi,\beta}(\xi^*; d_4))$, *with equality if either* $D_1 < d_1$ *or* $d_4 < D_2$.

From these two cases the optimal designs can be obtained numerically. In Table 1 we exhibit a few examples of locally optimal designs for different values of the slope parameter b and the mean effective dose (ED50) a in the standardized situation, $c = 1$, $D_1 = -\infty$ and $D_2 = 0$, when the dose x is reported on a logarithmic scale.

b	a	d_1	d_2	d_3	d_4
0.5	-2	-5.98	-5.47	-0.49	0
0.5	0	-5.23	-4.70	-0.47	0
0.5	2	-4.81	-4.27	-0.46	0
1	-2	-3.80	-3.30	-0.70	-0.20
1	0	-2.83	-2.28	-0.45	0
1	2	-2.55	-1.97	-0.42	0
5	-2	-2.57	-2.07	-1.93	-1.43
5	0	-1	0	—	—
5	2	-1	0	—	—

Table 1. Locally D–optimal designs for a logistic regression model; bounded intensity

Note that for small values of a the constraint of a maximal dose is of no importance and the resulting locally optimal designs are only constrained by the design density bounded by one. For smaller values of the slope parameter b, the optimal design consists of two separate intervals whereas for larger values these intervals are merged to a single one.

Remarks. 1. The result derived above does not depend on the particular shape of the bounding density, i.e. the constant density. It also holds in the general setting $g(z) = \varrho(z)\, 1_{[D_1,D_2]}(z)$ with $\varrho(z) > 0$.

2. By translation and rescaling any restriction of the form $\xi \le \mu$, μ given by the density $g(z) = c\,1_{[D_1,D_2]}$, $c > 0$, can be transformed to one of the form $\widetilde{g}(z) = 1_{[D,0]}$. More precisely, let $\xi^*_{a,b,g}$ denote the locally optimal design for parameters a and b under the constraint that the design density is bounded by $g(x) = c\,1_{[D_1,D_2]}$. Then, if $D_2 < \infty$, $\xi^*_{a,b,g}(x) = \xi^*_{\widetilde{a},\widetilde{b},\widetilde{g}}(c(x - D_2))$, where $\widetilde{a} = c(a - D_2)$, $\widetilde{b} = b/c$ and $\widetilde{g}(z) = 1_{[c(D_1-D_2),0]}$.
If $D_2 = \infty$ and $D_1 > -\infty$ then by an additional reflection, we have $\xi^*_{a,b,g}(x) = \xi^*_{\widetilde{a},\widetilde{b},\widetilde{g}}(c\,(D_1 - x))$, whith $\widetilde{a} = c(D_1 - a)$, $\widetilde{b} = b/c$ and $\widetilde{g}(z) = 1_{(\infty,0]}$.
Alternatively one could only consider the design problem in the standard form $a = 0$ and $b = 1$, since, again owed to the transformation equivariance, $\xi^*_{a,b,g}(x) = \xi^*_{0,1,\widetilde{g}}(b(x - a))$ where $\widetilde{g}(z) = c/b\ 1_{[b(D_1-a),bD_2]}(z)$.

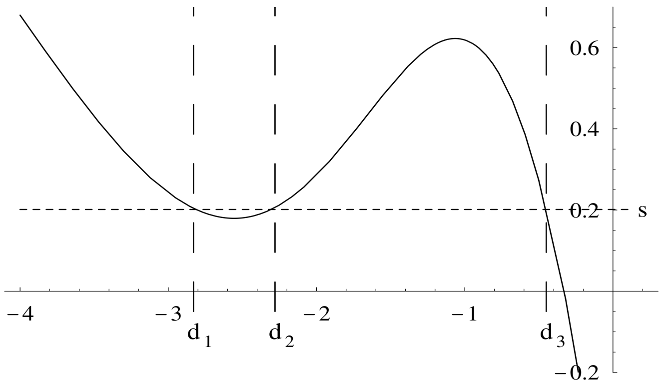

Figure 1. Directional derivative $F_{\Phi,\beta}(\xi; x)$ of the locally optimal design

Figure 1 exhibits the directional derivative $F_{\Phi,\beta}(\xi; x)$ of the locally optimal design for the parameter values $b = 1$ and $a = 0$. The reference lines indicate the support points $d_1, \ldots, d_4$ of the design (cf Table 1) and the corresponding value s for the bound occuring in the equivalence theorem.

4. MARGINALLY BOUNDED DESIGNS

In this section we consider additive two–factor models

$$Y(x_1, x_2) = \beta_0 + f_1(x_1)^\top \beta_1 + f_2(x_2)^\top \beta_2 + Z(x_1, x_2)$$

with explicit constant term β_0, where the components x_1 and x_2 of the design variable $x = (x_1, x_2)$ may be adjusted independently, $x \in \mathcal{X}_1 \times \mathcal{X}_2$. Due to separate constraints on both variables we are looking for marginally bounded designs ξ where their one–dimensional projections ξ_1 and ξ_2 are bounded by

$\nu_1 \leq \xi_1 \leq \mu_1$ and $\nu_2 \leq \xi_2 \leq \mu_2$, respectively. Note that the projections are defined by $\xi_1(B_1) = \xi(B_1 \times \mathcal{X}_2)$ for every (measurable) subset B_1 of $\mathcal{X}_1$, and for ξ_2 accordingly. Denote by $\Xi_{\nu_1,\nu_2}^{\mu_1,\mu_2} = \{\xi;\, \nu_1 \leq \xi_1 \leq \mu_1,\, \nu_2 \leq \xi_2 \leq \mu_2\}$ the set of competing marginally bounded designs. Then we can characterize the optimal design for the D–criterion as the measure theoretic product of the optimal designs in the corresponding marginal models. This expands the decomposition results by Schwabe and Wierich (1995) and Schwabe (1996) to the situation of marginally bounded designs.

Theorem 2. *If the designs ξ_i^* are D–optimal in $\Xi_{\nu_i}^{\mu_i}$ for the marginal models given by the response functions $E(Y^{(i)}(x_i)) = \beta_0 + f_i(x_i)^\top \beta_i$, $x_i \in \mathcal{X}_i$, $i = 1,\, 2$, then the product design $\xi_1^* \otimes \xi_2^*$ is D–optimal in $\Xi_{\nu_1,\nu_2}^{\mu_1,\mu_2}$.*

Proof. As has been shown in Schwabe and Wierich (1995) every design ξ is dominated by the product $\xi_1 \otimes \xi_2$ of its marginals with respect to the D–criterion, $\det M(\xi) \leq \det M(\xi_1 \otimes \xi_2)$, and for product designs the determinant can be decomposed according to $\det M(\xi_1 \otimes \xi_2) = \det M_1(\xi_1) \det M_2(\xi_2)$, where $M_1(\xi_1)$ and $M_2(\xi_2)$ are the information matrices in the marginal models. Moreover, with $\xi \in \Xi_{\nu_1,\nu_2}^{\mu_1,\mu_2}$ its marginals satisfy the marginal bounds $\nu_i \leq \xi_i \leq \mu_i$ and, hence, the product $\xi_1 \otimes \xi_2$ is in $\Xi_{\nu_1,\nu_2}^{\mu_1,\mu_2}$. Thus an optimal design can be found in the class of product designs and the best product design is constructed as the product of the D–optimal marginal designs. □

In case of marginally restricted designs, where the first marginal is fixed, $\xi_1 = \mu_1$, and no bounds are imposed on the second marginal ξ_2, we recover the result by Cook and Thibodeau (1980) that the product design $\mu_1 \otimes \xi_2^*$ is D–optimal for the additive model when a D–optimal marginal design ξ_2^* is used.

Obviously, Theorem 2 can be extended directly to additive models with more than two factors. In contrast to the unbounded situation a statement similar to Theorem 2 is not true for Kronecker product type models with complete interactions. This illustrates that there are structural differences between Kronecker product type and additive models.

Example. For multiple linear regression $Y(x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + Z(x_1, x_2)$ on a 20×20 point grid with uniformly bounded marginals, $\xi_i(x_i) \leq \mu_i(x_i) = \frac{1}{10}, x_i = 1, ..., 20$, the D–optimal design is given by equal weights $\frac{1}{100}$ on each point of the four 5×5 grids located at the corners of the design region as exhibited in Figure 2. Note that this design differs from the design obtained by Müller (1998, p. 90) which is optimal under the constraint $\xi \leq \mu_1 \otimes \mu_2$.

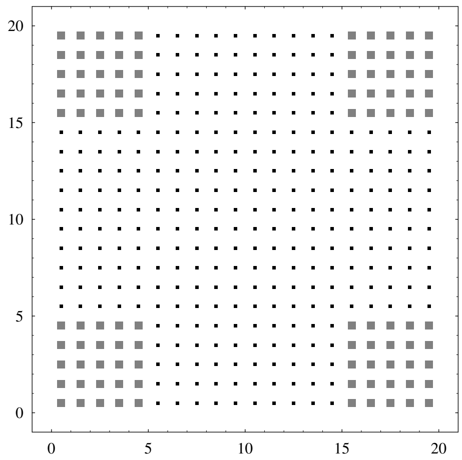

Figure 2. D–optimal design for multiple linear regression on a 20×20 grid with uniformly bounded marginals

5. APPENDIX: PROOF OF THEOREM 1

We first show that (a) implies (b) by contradiction: Assume that (b) does not hold. Then there are subsets $\mathcal{A}$ and $\mathcal{B}$ of the design region $\mathcal{X}$ such that $a := \xi^*(\mathcal{A}) - \nu(\mathcal{A}) > 0$, $b := \mu(\mathcal{B}) - \xi^*(\mathcal{B}) > 0$ and $\inf_{x\in\mathcal{A}} F_\Phi(\xi^*; x) > \sup_{x\in\mathcal{B}} F_\Phi(\xi^*; x)$. Define a competing design ξ' by $\xi' = \xi^* - \frac{\min(a,b)}{a}(\xi^* - \nu)$ on $\mathcal{A}$, $\xi' = \xi^* + \frac{\min(a,b)}{b}(\mu - \xi^*)$ on $\mathcal{B}$ and $\xi' = \xi^*$ on the complimentary set $\mathcal{X} \setminus (\mathcal{A} \cup \mathcal{B})$. Then $\nu \leq \xi' \leq \mu$ and $F_\Phi(\xi^*, \xi') < 0$. Hence, in view of Theorem A (i) the design ξ^* cannot be optimal in contradiction to (a).

To see that (b) implies (c) we only have to set $\mathcal{X}_1 = \{x;\, F_\Phi(\xi^*; x) > s\}$ and $\mathcal{X}_2 = \{x;\, F_\Phi(\xi^*; x) < s\}$ which have the required properties.

Finally, to show that (c) implies (a) we investigate the directional derivative $F_\Phi(\xi^*, \eta)$ for every competing design η. In view of condition (i) in (c) we have $\eta \geq \xi^*$ on $\mathcal{X}_1$ and $\eta \leq \xi^*$ on $\mathcal{X}_2$ for every design $\eta \in \Xi_\nu^\mu$. Hence,

$$\int_{\mathcal{X}_1} F_\Phi(\xi^*; x)\eta(dx) \geq \int_{\mathcal{X}_1} F_\Phi(\xi^*; x)\xi^*(dx) + \inf_{x\in\mathcal{X}_1} F_\Phi(\xi^*; x)(\eta(\mathcal{X}_1) - \xi^*(\mathcal{X}_1))$$

and

$$\int_{\mathcal{X}_2} F_\Phi(\xi^*; x)\eta(dx) \geq \int_{\mathcal{X}_2} F_\Phi(\xi^*; x)\xi^*(dx) - \sup_{x\in\mathcal{X}_2} F_\Phi(\xi^*; x)(\xi^*(\mathcal{X}_2) - \eta(\mathcal{X}_2)).$$

Employing conditions (ii) and (iii) of (c) we, thus, obtain

$$F_\Phi(\xi^*, \eta) = \int_{\mathcal{X}} F_\Phi(\xi^*; x)\eta(dx) \geq \int_{\mathcal{X}} F_\Phi(\xi^*; x)\xi^*(dx) + s(\eta(\mathcal{X}) - \xi^*(\mathcal{X})) = 0$$

which establishes the optimality of ξ^* in Ξ_ν^μ with respect to Φ in view of Theorem A (i). $\square$

Acknowlegdement. The work of the second author was partly supported by the research grant Ku 719/2 of the Deutsche Forschungsgemeinschaft at the Free University, Berlin.

References

Cook R.D. and Fedorov V.V., (1995), Constrained optimization of experimental design. *Statistics* **28**, 129-178

Cook, R.D. and Thibodeau L.A., (1980), Marginally restricted D-optimal designs. *J. Amer. Statist. Assoc.* **75**, 366-371

Dette H. and Sahm M., (1997), E-optimal designs for the double exponential model. In *MODA 5 – Advances in Model–Oriented Data Analysis and Experimental Design* (Atkinson A.C., Pronzato L. and Wynn H.P., eds.). Heidelberg, Physica, 11-20

Fedorov V.V., (1972), *Theory of Optimal Experiments.* New York, Academic Press.

Fedorov V.V., (1989), Optimal design with bounded density: Optimization algorithms of the exchange type. *J. Statist. Plann. Inference* **22**, 1-13

Fedorov V.V. and Hackl P., (1997), *Model–Oriented Design of Experiments. Lecture Notes in Statistics* **125**. New York, Springer.

Kiefer J.C., (1959), Optimal experimental designs. *J. R. Statist. Soc., Ser. B* **21**, 272-304

Müller W.G., (1998), *Collecting Spatial Data.* Heidelberg, Physica.

Schwabe R., (1996), *Optimum Designs for Multi–Factor Models. Lecture Notes in Statistics* **113**. New York, Springer.

Schwabe R. and Wierich W., (1995), D–optimal designs of experiments with non–interacting factors. *J. Statist. Plann. Inference* **44**, 371-384

Sitter R.R. and Wu C.F.J., (1993), Optimal designs for binary response experiments: Fieller–, D– and A–criteria. *Scand. J. Statist.* **20**, 329-341

Whittle P., (1973), Some general points in the theory of optimal experimental design. *J. R. Statist. Soc., Ser. B* **35** 916-921

Wierich W., (1985), Optimum designs under experimental constraints for a covariate model and an intra–class regression model. *J. Statist. Plann. Inference* **12**, 27-40

Wynn H.P., (1982), Optimum submeasures with applications to finite population sampling. In *Statistical Decision Theory and Related Topics III. Proc. Third Purdue Symp., Vol. 2.* (Gupta S.S. and Berger J.O., eds.). New York, Academic Press, 485-495