

## MINIMAX OPTIMAL DESIGNS IN NONLINEAR REGRESSION MODELS

Holger Dette and Michael Sahn

*Ruhr-Universität Bochum*

*Abstract:* We consider the maximum variance optimality criterion of Elfving (1959) in the context of (nonlinear) response models. Some practical guidelines for the construction of minimax optimal designs are given. In some cases this criterion yields one point designs as a consequence of different scales of the elements in the Fisher information matrix. As an alternative a “standardized” maximum variance criterion is introduced and applied to the calculation of efficient designs. The results are illustrated for binary response models and it is demonstrated that in these models standardized minimax optimality should be preferred to ordinary minimax optimality.

*Key words and phrases:* Binary response models, maximum variance criterion, non-linear regression models, optimal design, standardized variances.

### 1. Introduction

Consider a response variable  $y$  having a distribution from an exponential family  $p(y|x, \theta)$  with Fisher information  $I(\theta, x)$  for  $\theta$  given an observation at  $x$ . Here  $\theta = (\theta_1, \dots, \theta_m)^T$  denotes an  $m$ -dimensional vector of unknown parameters and  $x$  is the explanatory variable which varies in a design space  $\mathcal{X} \subseteq \mathbb{R}^k$ . A design is a probability measure on  $\mathcal{X}$  with finite support and the matrix

$$I(\theta, \xi) = \int_{\mathcal{X}} I(\theta, x) d\xi(x) \in \mathbb{R}^{m \times m}$$

is called the Fisher information matrix of  $\xi$ . A (locally) optimal design maximizes an optimality criterion depending on  $I(\theta, \xi)$  (see Silvey (1980), Atkinson and Donev (1992) or Pukelsheim (1993)). In this paper we consider the problem of constructing optimal designs for binary response models with respect to the partial maximum variance criterion introduced by Elfving (1959)

$$\Phi(\xi) = \max_{i \in \mathcal{M}} e_i^T I^{-1}(\theta, \xi) e_i, \tag{1.1}$$

where  $\xi$  is a design for which the parameters  $\{\theta_i\}_{i \in \mathcal{M}}$  are estimable,  $e_i \in \mathbb{R}^m$  denotes the  $i$ th unit vector ( $i = 1, \dots, m$ ),  $\mathcal{M} \subset \{1, \dots, m\}$  and  $I^{-1}(\theta, \xi)$  denotes

a generalized inverse of  $I(\theta, \xi)$ . In many cases the elements of the Fisher information matrix are of different scale. Therefore we also consider the minimum efficiency criterion

$$\psi(\xi) = \max_{i \in \mathcal{M}} \frac{e_i^T I^{-}(\theta, \xi) e_i}{\min_{\eta} e_i^T I^{-}(\theta, \eta) e_i} = 1 / \min_{i \in \mathcal{M}} \text{eff}_{\theta_i}(\xi), \quad (1.2)$$

which compares the variances relative to their optimal value obtainable by the choice of a design. More precisely, the quantities  $\text{eff}_{\theta_i}(\xi)$  in (1.2) are the efficiencies of the design  $\xi$ , relative to the optimal design for estimating the parameter  $\theta_i$ . Designs minimizing (1.1) are called minimax optimal for the parameters  $\{\theta_i | i \in \mathcal{M}\}$  and have been discussed by many authors (see e.g. Murty (1971), Torsney and López-Fidalgo (1995), Krafft and Schaefer (1995) or Dette and Studen (1994)). Following Dette (1997) we call the designs minimizing (1.2) “standardized” minimax optimal for the parameters  $\{\theta_i | i \in \mathcal{M}\}$ . If  $\mathcal{M} = \{1, \dots, m\}$  we briefly talk about minimax or standardized minimax optimal designs. It is worthwhile to mention that a similar criterion was considered by Müller (1995), who determined designs maximizing the minimum of  $A$ -efficiencies, where the minimum is taken with respect to the parameter  $\theta$ .

Note that in many cases of practical interest the optimality criteria (1.1) and (1.2) produce locally optimal designs which require prior knowledge of the unknown parameter  $\theta$ . Some arguments in favour of local optimality criteria can be found in Ford, Torsney and Wu (1992). For alternative optimality criteria see Wu (1985), Chaloner and Larntz (1989), Sitter (1992).

In Section 2 we provide some theoretical background regarding minimax and standardized minimax optimality which is useful for the construction of optimal designs. We also present a procedure for determining minimax and standardized minimax optimal designs. Some guidelines for the construction of minimax optimal designs in binary response models are given in Section 3.1. It turns out that for some parameter combinations the minimax optimal designs become inefficient in the sense that the number of different dose levels is less than the number of parameters in the model. This motivates the consideration of standardized minimax optimal designs for binary response models which are discussed in Section 3.2. Finally, in Section 4 the results are illustrated for two binary response models, the double exponential and the logistic regression model.

## 2. Constructing Minimax and Standardized Minimax Optimal Designs

The equivalence theory of the minimax optimality criteria can be obtained from Pukelsheim (1993), page 175. More precisely, assume that  $\{I(\theta, x) | x \in \mathcal{X}\}$  is compact, consider a subset  $\mathcal{M} \subset \{1, \dots, m\}$  and a design  $\xi$  for which  $e_j^T \theta$  is

estimable for all  $j \in \mathcal{M}$ . Define  $\rho_k^{-1} = \min_{\eta} e_k^T I^{-}(\theta, \eta) e_k$  (where the minimum is taken over all designs for which  $e_k^T \theta$  is estimable) and

$$\begin{aligned} \mathcal{N}(\xi) &:= \left\{ j \in \mathcal{M} \mid e_j^T I^{-}(\theta, \xi) e_j = \max_{i \in \mathcal{M}} e_i^T I^{-}(\theta, \xi) e_i \right\} \\ \tilde{\mathcal{N}}(\xi) &:= \left\{ j \in \mathcal{M} \mid \rho_j e_j^T I^{-}(\theta, \xi) e_j = \max_{i \in \mathcal{M}} \rho_i e_i^T I^{-}(\theta, \xi) e_i \right\}; \end{aligned}$$

then the design  $\xi$  is minimax optimal for the parameters  $\{\theta_i \mid i \in \mathcal{M}\}$  if and only if there exists a generalized inverse  $G$  of  $I(\theta, \xi)$  and nonnegative numbers  $\alpha_j, j \in \mathcal{N}(\xi)$ , with sum equal to 1 such that

$$\sum_{j \in \mathcal{N}(\xi)} \alpha_j \frac{e_j^T G I(\theta, x) G e_j}{e_j^T I^{-}(\theta, \xi) e_j} \leq 1, \tag{2.1}$$

for all  $x \in \mathcal{X}$ , where equality holds on the support of  $\xi$ . Similarly the design  $\xi$  is standardized minimax optimal for the parameters  $\{\theta_i \mid i \in \mathcal{M}\}$  if and only if there exists a generalized inverse  $G$  of  $I(\theta, \xi)$  and nonnegative numbers  $\tilde{\alpha}_j, j \in \tilde{\mathcal{N}}(\xi)$ , with sum equal to 1 such that

$$\sum_{j \in \tilde{\mathcal{N}}(\xi)} \tilde{\alpha}_j \frac{e_j^T G I(\theta, x) G e_j}{e_j^T I^{-}(\theta, \xi) e_j} \leq 1 \tag{2.2}$$

with equality on the support of  $\xi$ . The following two Lemmata are obvious from these considerations.

**Lemma 2.1.** *A minimax optimal design  $\xi^*$  for the parameters  $\{\theta_i \mid i \in \mathcal{M}\}$  is also minimax optimal for the parameters  $\{\theta_i \mid i \in \mathcal{N}(\xi^*)\}$ . A standardized minimax optimal design  $\tilde{\xi}$  for the parameters  $\{\theta_i \mid i \in \mathcal{M}\}$  is also standardized minimax optimal for the parameters  $\{\theta_i \mid i \in \tilde{\mathcal{N}}(\tilde{\xi})\}$ .*

**Lemma 2.2.** *If for all  $\mathcal{A} \subset \mathcal{M}$  with  $\#\mathcal{A} \leq k < \#\mathcal{M}$  the (standardized) minimax optimal design for the parameters  $\{\theta_i \mid i \in \mathcal{A}\}$  is **not** (standardized) minimax optimal for  $\{\theta_i \mid i \in \mathcal{M}\}$ , then in the generalized inverse of the Fisher information matrix of the (standardized) minimax optimal design for  $\{\theta_i \mid i \in \mathcal{M}\}$  the largest diagonal entry has at least multiplicity  $k + 1$ .*

Lemma 2.2 suggests an iterative procedure for determining minimax optimal designs for the parameters  $\{\theta_i \mid i \in \mathcal{M}\}$ . For  $n = 1, \dots, \#\mathcal{M} - 1$  we proceed as follows.

1. For all subsets  $\mathcal{M}_n \subset \mathcal{M}$  with  $\#\mathcal{M}_n = n$  we calculate the minimax optimal designs  $\xi_n^*$  for the parameters  $\{\theta_i \mid i \in \mathcal{M}_n\}$  under the constraint that the corresponding entries in the diagonal of the inverse fisher information matrix

are all equal and check if  $\xi_n^*$  is already minimax optimal for the parameters  $\{\theta_i | i \in \mathcal{M}\}$ . By (2.1) this is equivalent to showing

$$\max_{j \in \mathcal{M}_n} e_j^T I^-(\theta, \xi_n^*) e_j = \max_{j \in \mathcal{M}} e_j^T I^-(\theta, \xi_n^*) e_j. \quad (2.3)$$

2. If (2.3) is satisfied, then  $\xi_n^*$  is minimax optimal for the parameters  $\{\theta_i | i \in \mathcal{M}\}$  and the procedure stops. Otherwise we put  $n = n + 1$  and repeat step (1).

Although this procedure appears to be cumbersome (because in the worst case one has to go  $\#\mathcal{M}$  steps), the results obtained in the literature (see Murty (1971), Krafft and Schaefer (1995), Dette and Studden (1994)) suggest that in many regression models the algorithm stops after the first or second step.

For the standardized minimax optimality criterion the procedure is similar, replacing (2.3) by

$$\max_{j \in \mathcal{M}_n} \rho_j e_j^T I^-(\theta, \xi_n^*) e_j = \max_{j \in \mathcal{M}} \rho_j e_j^T I^-(\theta, \xi_n^*) e_j.$$

Moreover, the following Lemma shows that for standardized minimax optimality one can always start with  $n = 2$ .

**Lemma 2.3.** *If  $\xi^*$  is a standardized minimax optimal design for the parameters  $\{\theta_i | i \in \mathcal{M}\}$  and  $\#\mathcal{M} \geq 2$ , then  $\#\tilde{\mathcal{N}}(\xi^*) \geq 2$ .*

**Proof.** Let  $\xi^*$  denote the standardized minimax optimal design and assume  $\#\tilde{\mathcal{N}}(\xi^*) = 1$ . Without loss of generality we put  $\tilde{\mathcal{N}}(\xi^*) = \{1\}$ . By Lemma 2.1  $\xi^*$  is also optimal for estimating the parameter  $\theta_1$ ; in other words

$$1 = \rho_1 e_1^T I^-(\theta, \xi^*) e_1 = \max_{j \in \mathcal{M}} \rho_j e_j^T I^-(\theta, \xi^*) e_j \geq \min_{j \in \mathcal{M}} \rho_j e_j^T I^-(\theta, \xi^*) e_j \geq 1.$$

This proves  $\tilde{\mathcal{N}}(\xi^*) = \mathcal{M}$ , contradicting the assumption  $\#\mathcal{M} \geq 2$ .

### 3. Minimax and Standardized Minimax Optimal Designs for Binary Response Models

In this section we demonstrate the calculation of minimax and standardized minimax optimal designs in binary response models. Consider a binary response variable  $y$  with probability of success  $p(x, \theta)$ , where  $\theta$  is a two dimensional parameter. We discuss two different parametrizations

$$p_1(x, \theta) = F(\beta(x - \mu)) \quad \theta = (\mu, \beta), \quad x \in \mathbb{R} \quad (3.1)$$

$$p_2(x, \theta) = F(a + bx) \quad \theta = (a, b), \quad x \in \mathbb{R} \quad (3.2)$$

that are commonly used in practice. Here  $F$  is a known distribution function with symmetric density function  $f$ , which is assumed to be differentiable for  $x \neq 0$ . Define

$$h(x) = \frac{f^2(x)}{F(x)(1 - F(x))}; \quad (3.3)$$

then the Fisher information for the parameter  $\theta$  of a design  $\xi$  is given by

$$I_1(\theta, \xi) = \int h(\beta(x - \mu)) \begin{pmatrix} \beta^2 & -\beta(x - \mu) \\ -\beta(x - \mu) & (x - \mu)^2 \end{pmatrix} d\xi(x) \quad (3.4)$$

if the parametrization (3.1) is used, and by

$$I_2(\theta, \xi) = \int h(a + bx) \begin{pmatrix} 1 & x \\ x & x^2 \end{pmatrix} d\xi(x)$$

if we use parametrization (3.2). Locally optimal designs in these models for various differentiable optimality criteria have been discussed by several authors (see e.g. Wu (1988), Abdelbasit and Plackett (1983), Sitter and Wu (1993), Gaudard, Karson, Lindner and Tse (1993)). Note that the number of support points of (locally) optimal designs depends sensitively on the form of the function  $h$  in (3.3) (see e.g. Sitter and Wu (1993)). We will now give some general guidelines for the determination of minimax and standardized minimax optimal designs in these models; illustrative examples are presented in the following section.

Throughout this paper we will assume that

$$\max_{t \in \mathbb{R}} h(t) = h(0) \quad (3.5)$$

and

$$k^2 = \max_{t \in \mathbb{R}} t^2 h(t) = c^2 h(c), c > 0 \quad (3.6)$$

which is satisfied for most of the commonly used binary response models. Moreover, the point  $c$  where the function  $t^2 h(t)$  attains its maximum is usually unique (up to its sign) and can be obtained from Table 4 in Ford, Torsney and Wu (1992).

### 3.1. Minimax optimal designs

We concentrate on the parametrization (3.1). Standard arguments show that there is a minimax optimal design symmetric with respect to the LD 50. Moreover, by the following lemma it suffices to consider symmetric designs with at most 4 support points.

**Lemma 3.1.** *In a binary response model with parametrization (3.1) there exists a minimax optimal design with at most 4 support points which is symmetric to the LD50  $\mu$ .*

The proof is deferred to the appendix. Note, however, that this result does not depend on the binary response model but holds for any regression model with two parameters having a corresponding symmetry property.

Following the procedure in Section 2 we first determine the optimal designs  $\xi_\mu$  and  $\xi_\beta$  for estimating the LD 50 and the slope. These are easily obtained from Elfving's theorem (Elfving (1952)) and assumptions (3.5), (3.6) observing

$$\rho_1^{-1} = \inf_{\xi} e_1^T I^{-1}(\theta, \xi) e_1 = \frac{1}{\beta^2 h(0)} \quad (3.7)$$

$$\rho_2^{-1} = \inf_{\xi} e_2^T I^{-1}(\theta, \xi) e_2 = \frac{\beta^2}{k^2}. \quad (3.8)$$

This gives for  $\xi_\mu$  a one point design at  $\mu$  and for  $\xi_\beta$  a two point design

$$\xi_\beta = \begin{pmatrix} \mu - \frac{c}{\beta} & \mu + \frac{c}{\beta} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad (3.9)$$

where  $c > 0$  is defined by (3.6). Now  $\xi_\beta$  is minimax optimal if and only if

$$e_1^T I^{-1}(\theta, \xi_\beta) e_1 \leq e_2^T I^{-1}(\theta, \xi_\beta) e_2$$

which is equivalent to  $\beta^2 \geq c$ . In the other case the inverse Fisher information matrix must have equal diagonal elements. We start with a two point design where the equality of the diagonal elements readily yields

$$\xi_0^* = \begin{pmatrix} \mu - \beta & \mu + \beta \\ 1/2 & 1/2 \end{pmatrix}. \quad (3.10)$$

The minimax optimality of  $\xi_0^*$  can be checked via the inequality (2.1) where the  $\alpha_j$  ( $\mathcal{N}(\xi_0^*) = \{1, 2\}$ ) have to be chosen (due to equality in (2.1) at the support points of the minimax optimal design and the differentiability of  $h$ ) as

$$\alpha_2 = 1 - \alpha_1 = -\frac{\beta^2 h'(\beta^2)}{2 h(\beta^2)}.$$

If inequality (2.1) is not fulfilled we increase the number of support points successively. Thus in the next step we consider

$$\xi_w^* = \begin{pmatrix} \mu - \frac{z}{\beta} & \mu & \mu + \frac{z}{\beta} \\ \frac{1-w}{2} & w & \frac{1-w}{2} \end{pmatrix}, \quad (3.11)$$

where  $z > 0$  and

$$w = \frac{h(z)(z^2 - \beta^4)}{\beta^4[h(0) - h(z)] + z^2 h(z)}$$

due to equality of the diagonal elements in  $I^{-1}(\theta, \xi_w^*)$ . The optimal  $z$  minimizes

$$\frac{h(0)}{z^2 h(z)} - \frac{1}{z^2} \quad (3.12)$$

and the “weights”  $\alpha_1, \alpha_2$  in the equivalence theorem are given by

$$\alpha_2 = 1 - \alpha_1 = \left[ 1 - \frac{z^2}{\beta^4} - \frac{2z h(z)}{\beta^4 h'(z)} \right]^{-1}. \tag{3.13}$$

Note that the expression in (3.12) is independent of the parameters  $\mu$  and  $\beta$  and the minimizing  $z$  is usually also unique. If (2.1) holds for these weights, then  $\xi_w^*$  is minimax optimal, otherwise the procedure has to be continued. Thus in the last step we have to consider designs of the form

$$\xi^* = \left( \begin{array}{cccc} \mu - \frac{z_1}{\beta} & \mu - \frac{z_2}{\beta} & \mu + \frac{z_2}{\beta} & \mu + \frac{z_1}{\beta} \\ \frac{w}{2} & \frac{1-w}{2} & \frac{1-w}{2} & \frac{w}{2} \end{array} \right).$$

The equality of the diagonal entries of the Fisher information matrix yields

$$w = \frac{h(z_2)(\beta^4 - z_2^2)}{h(z_2)(\beta^4 - z_2^2) + h(z_1)(z_1^2 - \beta^4)}.$$

The optimal values for  $z_1$  and  $z_2$  maximize the function

$$\frac{h(z_1)h(z_2)(z_1^2 - z_2^2)}{h(z_2)(\beta^4 - z_2^2) + h(z_1)(z_1^2 - \beta^4)}$$

under the constraint that  $z_1 > \beta^2 > z_2 > 0$  (i.e.  $0 < w < 1$ ). Furthermore, the differentiability of  $h$  implies that either  $h(z_1) = h(z_2)$  and  $h'(z_1) = h'(z_2) = 0$  or

$$\frac{z_1(h(z_1))^2}{h'(z_1)} = \frac{z_2(h(z_2))^2}{h'(z_2)} \quad \text{and} \quad \frac{h(z_1)h'(z_2)}{h(z_2)(h(z_2) - h(z_1))} = \frac{2z_2}{z_2^2 - z_1^2}.$$

The quantities  $\alpha_1$  and  $\alpha_2$  in the equivalence theorem are given by

$$\alpha_2 = 1 - \alpha_1 = \frac{\beta^4(h(z_2) - h(z_1))}{h(z_2)(\beta^4 - z_2^2) + h(z_1)(z_1^2 - \beta^4)}.$$

Recent results in the literature suggest that for most binary response models the symmetric optimal designs will have either two or three support points and are given by (3.9), (3.10) or (3.11), respectively (see Sitter and Wu (1993) for some results regarding A-, D- and Fieller-optimality and also for the construction of a model with a symmetric D-optimal design supported at four points).

For the parametrization (3.2) the approach is very similar and illustrated in the second part of Section 4.

### 3.2. Standardized minimax optimal designs

We demonstrate in Section 4 that minimax optimal designs for binary response models may become inefficient in the sense that the number of different

dose levels of the optimal design is less than the number of parameters in the model. This disadvantage is usually caused by the different size of the elements in the Fisher information matrix. The standardization used in the criterion (1.2) makes elements of different scale better comparable and does not yield degenerate optimal designs (see Lemma 2.3). It turns out that in binary response models the standardized minimax optimal designs can be easily obtained from the minimax optimal designs for special parameter combinations. More precisely, for the parametrization (3.1) a minimax optimal design  $\xi^*$  can be obtained from  $\xi^*(x) = \eta(\beta(x - \mu))$  where  $\eta$  minimizes

$$\frac{1}{c_2 c_0 - c_1^2} \max\left\{\frac{c_2}{\beta^2}, c_0 \beta^2\right\}, \quad (3.14)$$

$c_j = \int y^j h(y) d\eta(y)$  ( $j = 0, 1, 2$ ). Similarly, the standardized minimax optimal design  $\tilde{\xi}$  is given by  $\tilde{\xi}(x) = \tilde{\eta}(\beta(x - \mu))$  where  $\tilde{\eta}$  minimizes

$$\frac{1}{c_2 c_0 - c_1^2} \max\{h(0)c_2, c_0 k^2\} = \frac{\sqrt{h(0)}k}{c_0 c_2 - c_1^2} \max\left\{\frac{\sqrt{h(0)}}{k} c_2, \frac{k}{\sqrt{h(0)}} c_0\right\}. \quad (3.15)$$

This follows easily from definitions (1.2), (3.7) and (3.8). Note that the optimization problem in (3.15) does not depend on the parameters  $\mu$  and  $\beta$ . Moreover, the maximum of (3.15) is obtained by maximizing (3.14) for  $\beta^2 = k/\sqrt{h(0)}$  which proves the following result.

**Lemma 3.2.** *Let  $\xi_{\mu,\beta}^*$  denote the minimax optimal design for a binary response model with parametrization (3.1) such that (3.5) and (3.6) are satisfied. Then the standardized minimax optimal design  $\tilde{\xi}$  for this model is given by*

$$\tilde{\xi}(x) = \xi_{0,s}^*\left(\frac{\beta}{s}(x - \mu)\right),$$

where  $s = k^{1/2}(h(0))^{-1/4}$  and  $k$  is defined in (3.6).

For the parametrization (3.2) a similar result can be established. The proof is omitted for the sake of brevity.

**Lemma 3.3.** *Let  $\xi_{a,b}^*$  denote the minimax optimal design for a binary response model with parametrization (3.2) such that*

$$S = \left\{ \sqrt{h(\omega)} \binom{1}{\omega} \mid \omega \in \mathbb{R} \right\} \quad (3.16)$$

is convex and

$$\lim_{|\omega| \rightarrow \infty} \omega^2 h(\omega) = 0. \quad (3.17)$$



The standardized minimax optimal design for this model is given by

$$\tilde{\xi}(x) = \xi_{a,s}^* \left( \frac{bx}{s} \right),$$

where  $s = mc$ ,  $c$  is defined in (3.6) and

$$m^2 = m^2(a) = h(c) \inf_{\xi} e_1^T I^{-}(\theta, \xi) e_1 = \begin{cases} \frac{a^2}{c^2} & \text{if } |a| > c \\ \frac{h(c)}{h(a)} & \text{if } |a| \leq c \end{cases}. \quad (3.18)$$

The conditions (3.16) and (3.17) are satisfied for the commonly used regression models (see Sitter and Wu (1993) or Ford, Torsney and Wu (1992)). It is also worthwhile to mention that for both parametrizations (3.1) and (3.2) the standardized minimax optimal designs are scale invariant in the sense that a scaling of the explanatory variable  $x$  transfers to the standardized minimax optimal design. The experimenter only has to scale the support points of the standardized minimax optimal design according to the transformation of the  $x$ -space. Moreover, by Lemma 3.2, the same argument applies for translations in binary response models with parametrization (3.1). In this case the standardized minimax optimal designs are invariant with respect to linear transformations of the explanatory variable. This is a strong advantage of parametrization (3.1) over the alternative parametrization (3.2) in binary response models.

#### 4. Examples

In this section we illustrate the result with two examples; a double exponential model with parametrization (3.1) and a logistic regression with parametrization (3.2). We also demonstrate that the maximum variance criterion can produce inefficient designs for the binary response model in the sense that the number of different dose levels of the minimax optimal design is less than the number of parameters in the model.

##### 4.1. Double exponential model

Let

$$F(t) = \frac{1 + \text{sign}(t)}{2} - \frac{\text{sign}(t)}{2} e^{-|t|}$$

and consider the parametrization (3.1). We calculate  $h(t) = (2e^{|t|} - 1)^{-1}$  and the values for  $k$  and  $c$  in (3.6) are easily obtained as

$$c \approx 1.84141 \quad k \approx 0.5404. \quad (4.1)$$

We now apply the general procedure of Section 3.1. If  $\beta^2 \geq c$ , the optimal design  $\xi_{\beta}$  for estimating the slope is also minimax optimal and given by (3.9). In a

second step we consider the case  $\beta^2 \leq c$  and check whether the design in (3.10) is minimax optimal using  $\alpha_2 = 1 - \alpha_1 = -\beta^2 h'(\beta^2)/2h(\beta^2)$  in the inequality (2.1). A straightforward calculation shows that (2.1) is satisfied if and only if  $v_0 \leq \beta^2 \leq c$  where  $v_0$  is the unique positive solution of  $v + 2e^{-v} = 2$ ,  $v_0 \approx 1.59362$ . Finally, if  $\beta^2 < v_0$  the minimax optimal design is given by (3.11) where  $z$  and  $w$  are defined by  $v_0$  and

$$w = \frac{(v_0^2 - \beta^4)h(v_0)}{h(v_0)(v_0^2 - \beta^4) + \beta^4}, \tag{4.2}$$

respectively (note that  $h(0) = 1$ ). This can be shown by the equivalence theorem (2.1) observing the representation of  $\alpha_1$  and  $\alpha_2$  in (3.13). The minimax optimal designs are listed in Table 4.1.

The standardized minimax optimal design are easily obtained by Lemma 3.2. More precisely we obtain, for  $s = k^{1/2}$ ,  $(h(0))^{-1/4} = c\sqrt{h(c)} \approx 0.5404 < \sqrt{v_0}$ , which shows that the standardized minimax optimal design is always supported at three points, i.e.

$$\tilde{\xi} = \begin{pmatrix} \mu - \frac{v_0}{\beta} & \mu & \mu + \frac{v_0}{\beta} \\ \frac{1-\tilde{w}}{2} & \tilde{w} & \frac{1-\tilde{w}}{2} \end{pmatrix}, \tag{4.3}$$

where  $v_0 \approx 1.59326$  and  $\tilde{w}$  is obtained from (4.2) as

$$\tilde{w} = \frac{(v_0^2 - k^2)h(v_0)}{k^2 + h(v_0)(v_0^2 - k^2)} \approx 0.4653.$$

This design has equal efficiencies  $\text{eff}_\beta(\xi) = \text{eff}_\mu(\xi) = (1-w)(v_0^2 h(v_0))(c^2 h(c))^{-1} \approx 0.5258$  for the individual parameters  $\beta$  and  $\mu$ .

Table 4.1. Minimax optimal designs  $\xi^*$  for the double exponential model with parametrization (3.1) (first row). The second row gives the value  $\tilde{\alpha}_2$  in the equivalence theorem (2.2), while the third and fourth rows contain the efficiencies for estimating the individual parameters  $\mu$  and  $\beta$ . The table shows the influence of the size of the slope parameter on the minimax optimal design. The weight  $w$  is given by (4.2), the constant  $c$  by (4.1) and  $v_0$  minimizes (3.12) [ $c \approx 1.84141, v_0 \approx 1.59362$ ].

	$\beta^2 < v_0$	$v_0 \leq \beta^2 \leq c$	$\beta^2 > c$
$\xi^*$	$\begin{pmatrix} \mu - \frac{v_0}{\beta} & \mu & \mu + \frac{v_0}{\beta} \\ \frac{1-w}{2} & w & \frac{1-w}{2} \end{pmatrix}$	$\begin{pmatrix} \mu - \beta & \mu + \beta \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$	$\begin{pmatrix} \mu - \frac{c}{\beta} & \mu + \frac{c}{\beta} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$
$\alpha_2$	$\frac{\beta^4}{\beta^4 + 1.008}$	$\frac{\beta^2}{2}(1 + h(\beta^2))$	1
$\text{eff}_\mu$	$\frac{0.287}{0.287 + 0.887\beta^4}$	$h(\beta^2)$	$h(c)$
$\text{eff}_\beta$	$\frac{0.983\beta^4}{0.287 + 0.887\beta^4}$	$\frac{\beta^4 h(\beta^2)}{c^2 h(c)}$	1

A brief comparison of the minimax and standardized minimax optimal designs in the double exponential model might be appropriate at this point. The standardized minimax optimal designs are based on more points than there are parameters in the model. As a consequence tests of goodness-of-fit of the double exponential model can be performed. Note also that the minimax optimal design for a small slope  $\beta$  quickly degenerates to a one point design (see Table 4.1). For example, if  $\beta = 0.1$ , the minimax optimal design advises the experimenter to take 99.9% of the observations at the LD 50  $\mu$ . However, this design does not allow any statistical inference regarding the slope parameter. These observations can be explained by observing the “different scales” of the elements in the Fisher information matrix (3.4). For small values of the slope the dominating element in  $I^{-1}(\theta, \xi)$  is the element at position (1, 1). Consequently the minimax optimal design is “close” to the optimal design for estimating the LD 50, which is usually the one point design concentrated at  $\mu$ . Note that these phenomena contradict the intuition. If the probability of success is slowly increasing with the dose level, a good design should cover a large range of dose levels. The standardization of the maximum variance criterion in (1.2) produces efficiencies, which are of the same scale and therefore better comparable than the “pure” diagonal elements of  $I^{-1}(\theta, \xi)$ . As a consequence we do not observe degenerate standardized minimax optimal designs at all. For example, if  $\beta = 0.1$ , the standardized minimax optimal design advises the experimenter to take 26.7% of the observations at  $\mu \pm 15.93$  and the remaining 46.6% of the observations at the LD50  $\mu$  (see equation (4.3)).

**4.2. Logistic regression**

As a second example consider the logistic regression model with parametrization (3.2), i.e.

$$F(t) = [1 + e^{-t}]^{-1} \quad \text{and} \quad h(t) = \frac{e^t}{(1 + e^t)^2}, \quad t \in \mathbb{R}$$

and the value  $c > 0$  maximizing  $t^2h(t)$  is given by  $c \approx 2.39936$ . We first determine the optimal designs for  $\xi_a$  and  $\xi_b$  for estimating the individual coefficients  $a$  and  $b$  using Elfving’s theorem (Elfving (1952)). This gives

$$\begin{aligned} \inf_{\xi} e_1^T I^{-1}(\theta, \xi) e_1 &= e_1^T I^{-1}(\theta, \xi_a) e_1 = \frac{m^2}{h(a)} \\ \inf_{\xi} e_2^T I^{-1}(\theta, \xi) e_2 &= e_2^T I^{-1}(\theta, \xi_b) e_2 = \frac{b^2}{c^2 h(c)}, \end{aligned}$$

(where  $m^2$  is defined by (3.18)) and the optimal designs are given by

$$\xi_a = \begin{cases} \begin{pmatrix} 0 \\ 1 \end{pmatrix} & \text{if } |a| \leq c \\ \begin{pmatrix} -\frac{a}{b} - \frac{c}{b} & -\frac{a}{b} + \frac{c}{b} \\ \frac{1}{2} - \frac{c}{2a} & \frac{1}{2} + \frac{c}{2a} \end{pmatrix} & \text{if } |a| > c \end{cases} \quad \xi_b = \begin{pmatrix} -\frac{a}{b} - \frac{c}{b} & -\frac{a}{b} + \frac{c}{b} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

From the procedure in Section 2 it follows that  $\xi_b$  is minimax optimal if and only if

$$e_1^T I^-(\theta, \xi_b) e_1 \leq e_2^T I^-(\theta, \xi_b) e_2$$

which is equivalent to  $b^2 \geq a^2 + c^2$ . Similarly we obtain

$$\xi_a \text{ is minimax optimal} \iff b^2 \leq a^2 - c^2.$$

In all other cases ( $|b^2 - a^2| < c^2$ ) we must have equality of the diagonal elements of the inverse Fisher information matrix of the minimax optimal design. A reasonable guess for a design at two points is

$$\xi = \begin{pmatrix} -\frac{a}{b} - \frac{v}{b} & -\frac{a}{b} + \frac{v}{b} \\ 1-w & w \end{pmatrix} \quad v > 0, w \in (0, 1) \tag{4.4}$$

which implies for the weight

$$w = \begin{cases} \frac{1}{2} & \text{if } a = 0 \\ \frac{1}{2} + \frac{v^2 + (a^2 - b^2)}{4av} & \text{if } a \neq 0. \end{cases} \tag{4.5}$$

The point  $v > 0$  can be found by maximizing the function  $h(x)g(x)$  for  $x > 0$ , where

$$g(x) = 2x^2(a^2 + b^2) - x^4 - (a^2 - b^2)^2, \tag{4.6}$$

and the optimality of the design in (3.10) is established using the equivalence theorem (2.1).

The resulting cases are listed in Table 4.2 which also contains the corresponding quantities  $\alpha_j$ , used for checking the minimax optimality by (2.1). Note that the table does not include the case  $a = 0, b^2 < c^2$ . In this case the minimax optimal design has equal masses at the points  $-1$  and  $1$  which also follows from (2.1) using  $\alpha_1 = 1$ .

Table 4.2. Locally minimax optimal designs for the logistic regression model with parametrization (3.2). The table shows the dependence of the locally optimal design on the size of the quantity  $a^2 - b^2$ . Here  $g(x)$  is defined by (4.6), the constant  $c = \arg \max\{x^2 h(x)\} \approx 2.39936$  by (3.6),  $v = \arg \max\{h(x)g(x)\}$  and  $w$  by (4.5).

	$a^2 - b^2 \leq -c^2$	$ a^2 - b^2  < c^2$	$c^2 \leq a^2 - b^2$
minimax opt. design	$\begin{pmatrix} \frac{c-a}{b} & \frac{-c-a}{b} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$	$\begin{pmatrix} \frac{-v-a}{b} & \frac{v-a}{b} \\ 1-w & w \end{pmatrix}$	$\begin{pmatrix} \frac{-c-a}{b} & \frac{c-a}{b} \\ \frac{a-c}{2a} & \frac{a+c}{2a} \end{pmatrix}$
$\alpha_1$	0	$\frac{b^4 - (a^2 - v^2)^2}{g(v)}$	1
$\Phi(\xi)$	$\frac{b^2}{c^2 h(c)}$	$\frac{4a^2 b^2}{h(v)g(v)}$	$\frac{a^2}{c^2 h(c)}$

Finally, the standardized minimax optimal designs for the logistic regression with parametrization (3.2) are obtained from Lemma 3.3 and given by

$$\tilde{\xi} = \begin{pmatrix} -\frac{a}{b} - \frac{\tilde{v}}{b} & -\frac{a}{b} + \frac{\tilde{v}}{b} \\ 1 - \tilde{w} & \tilde{w} \end{pmatrix} \tag{4.7}$$

$$\tilde{v} = \arg \max\{h(x)[2x^2(a^2 + s^2) - x^4 - (a^2 - s^2)^2]; x > 0\} \tag{4.8}$$

$$\tilde{w} = \begin{cases} \frac{1}{2} & \text{if } a = 0 \\ \frac{1}{2} + \frac{\tilde{v}^2 + (a^2 - s^2)}{4a\tilde{v}} & \text{if } a \neq 0, \end{cases} \tag{4.9}$$

where  $s = s(a)$  is defined in Lemma 3.3. Note that the standardized minimax optimal design is scale invariant, because  $\tilde{v}$  and  $\tilde{w}$  defined in (4.8) and (4.9) do not depend on the slope parameter  $b$ . In Tables 4.3 and 4.4 the minimax and standardized minimax optimal designs for the logistic regression model with parametrization (3.2) are calculated for some specific values of  $a$  and  $b$  as well as their efficiencies with respect to the optimal designs for the individual parameters  $a$  and  $b$ . We observe again that for many parameter combinations the minimax optimal designs are inefficient for estimating the slope, especially for small values of the slope parameter. The standardized minimax optimal designs are equally efficient for estimating the individual parameters. This efficiency does not depend on the parameter  $b$  and is increasing with the ‘‘intercept’’  $a$ .

Table 4.3. Locally minimax optimal designs  $\xi^*$  for logistic regression with parametrization (3.2) for various values of  $a$  and  $b$ . The designs are obtained from (4.4)-(4.6) and put masses  $1 - w$  and  $w$  at the points  $-a/b - v/b$  and  $-a/b + v/b$ . The last two columns give the efficiencies of the minimax optimal design for estimating the individual parameters. The table illustrates how sensitively the locally minimax optimal design depends on the slope parameter of the binary response model. In particular it shows that for small values of the slope, the minimax optimal design becomes inefficient for estimating the slope parameter.

$a$	$b$	$v$	$w$	$\Phi(\xi^*)$	$\text{eff}_a(\xi^*)$	$\text{eff}_b(\xi^*)$
0	0.1	0.1	0.5	4.01	0.9975	0.0057
0	1	1	0.5	5.09	0.786	0.448
0	c	c	0.5	13.11	0.305	1
1	0.1	1.003	0.9975	5.10	0.9982	0.0045
1	1	1.256	0.814	6.07	0.838	0.375
1	c	2.228	0.523	13.24	0.384	0.9902
c	0.1	2.397	0.9996	13.11	1.	0.0017
c	1	2.228	0.955	13.24	0.9902	0.172
c	c	2.033	0.712	16.58	0.790	0.790

Table 4.4. Locally standardized minimax optimal designs for the logistic regression model with parametrization (3.2) for various values of  $a$  and arbitrary  $b$ . The last column gives the efficiencies of the minimax optimal design for estimating the individual parameters. The designs are obtained from (4.7)-(4.9) and put masses  $1 - \tilde{w}$  and  $\tilde{w}$  at the points  $-a/b - \tilde{v}/b$  and  $-a/b + \tilde{v}/b$ . The locally standardized minimax optimal designs and its efficiencies are invariant with respect to the slope parameter. Thus the table illustrates the impact of the intercept on the locally standardized minimax optimal designs. Note that these designs never become inefficient for estimating the individual parameters.

$a$	$\tilde{v}$	$\tilde{w}$	$\Psi(\tilde{\xi})$	$\text{eff}_a(\tilde{\xi}) = \text{eff}_b(\tilde{\xi})$
0	1.325	0.5	1.507	0.663
0.1	1.328	0.523	1.507	0.664
1	1.541	0.685	1.475	0.678
c	2.033	0.712	1.265	0.790
10	2.376	0.559	1.014	0.986

## 5. Conclusions

In this paper we considered the problem of constructing minimax optimal designs for binary response models which minimize the asymptotic maximum variance of the estimates for the individual parameters. Some general guidelines for the determination of these designs are given and illustrated in the double exponential and logistic regression model. It turns out that minimax optimal designs may become inefficient in the sense, that the number of their dose levels is less than the number of parameters in the model. For this reason a standardized version (standardized minimax optimality) of Elfving's partial minimax criterion is also considered, which is based on the variances of the estimators for the individual coefficients relative to the best variances obtainable by the choice of an experimental design, i.e. efficiencies of the estimates are considered. The standardized minimax optimal designs for binary response models can be obtained from the minimax optimal designs and produce equal efficiencies for estimating the individual parameters in the model. In some cases the standardized minimax optimal design has more support points than the minimax optimal design and a goodness-of-fit test of the binary response model can be performed. Because the standardized minimax optimal designs are scale invariant and because they cannot degenerate to one point designs, we would recommend their application instead of minimax optimal designs.

## Acknowledgement

The authors are grateful to an unknown Associate Editor and referee for their valuable comments and suggestions which led to a substantial improvement of the original manuscript.

## Appendix

**Proof of Lemma 3.1.** Let  $z_1, \dots, z_k$  be arbitrary nonnegative numbers. We show that there is a symmetric minimax optimal design with at most four support points in the class of all symmetric designs with support in the set  $\{\mu \pm z_i/\beta | i = 1, \dots, k\}$ . Therefore, let  $\xi$  be a symmetric measure on  $\{\mu \pm z_i/\beta | i = 1, \dots, k\}$  with weights  $w_i/2$  at the points  $\mu \pm z_i/\beta$ . Due to the symmetry of  $\xi$  the Fisher information matrix is diagonal, i.e.

$$I(\theta, \xi) = \text{diag}\left(\beta^2 \sum_{i=1}^k w_i h(z_i), \beta^{-2} \sum_{i=1}^k w_i z_i^2 h(z_i)\right).$$

Optimizing with respect to the criterion (1.1) is therefore equivalent to maximizing

$$\min \left\{ \beta^2 \sum_{i=1}^k w_i h(z_i), \beta^{-2} \sum_{i=1}^k w_i z_i^2 h(z_i) \right\}$$

over all possible weights, that is maximizing on the  $k$ -simplex

$$\left\{ w_1, \dots, w_k \mid w_i \geq 0, \sum_{i=1}^k w_i = 1 \right\}.$$

Since both functions are linear in the weights the Lemma follows from a general result from linear programming. More precisely, the maximum value of the minimum of two linear functions on the  $k$ -simplex is always attained at a convex combination of two extreme points. The extreme points of the  $k$ -simplex represent the symmetric two point designs with weight  $1/2$  at the points  $\mu \pm z_{i_0}/\beta$ ,  $1 \leq i_0 \leq k$ . That means there is always a minimax optimal design in the considered class with at most 4 support points.

## References

- Abdelbasit, K. M. and Plackett, R. L. (1983). Experimental design for binary data. *J. Amer. Statist. Assoc.* **78**, 90-98.
- Atkinson, A. C. and Donev, A. N. (1992). *Optimum Experimental Designs*. Oxford: Oxford University Press.
- Chaloner, K. and Larntz, K. (1989). Optimal Bayesian design applied to logistic regression experiments. *J. Statist. Plann. Inference* **21**, 191-208.
- Dette, H. (1997). Designing experiments with respect to standardized optimality criteria. *J. Roy. Statist. Soc. Ser. B* **59**, 97-110.
- Dette, H. and Studden, W. J. (1994). Optimal designs with respect to Elfving's partial minimax criterion in polynomial regression. *Ann. Inst. Statist. Math.* **46**, 389-403
- Elfving, G. (1952). Optimum allocation in linear regression theory. *Ann. Math. Statist.* **23**, 255-262.
- Elfving, G. (1959). Design of linear experiments. In *Probability and Statistics. The Harald Cramér Volume* (Edited by U. Grenander), 58-74. Stockholm: Almagrist and Wiksell.

- Ford, I., Torsney, B. and Wu, C. F. J. (1992). The use of a canonical form in the construction of locally optimal designs for non-linear problems. *J. Roy. Statist. Soc. Ser. B* **54**, 569-583.
- Gaudard, M. A., Karson, M. J., Linder, E. and Tse, S. K. (1993). Efficient designs for estimation in the power logistic quantal response model. *Statist. Sinica* **3**, 233-243.
- Krafft, O. and Schaefer, M. (1995). Exact Elfving-minimax designs for quadratic regression. *Statist. Sinica* **5**, 475-484.
- Müller, C. (1995). Maximin efficient designs for estimating nonlinear aspects in linear models. *J. Statist. Plann. Inference* **44**, 117-132.
- Murty, V. N. (1971). Minimax designs. *J. Amer. Statist. Assoc.* **66**, 319-320.
- Pukelsheim, F. (1993). *Optimal Experimental Design*. New York: Wiley.
- Silvey, S. D. (1980). *Optimal Design: An Introduction to the Theory for Parameter Estimation*. Chapman and Hall, London.
- Sitter, R. R. (1992). Robust design for binary data. *Biometrics* **48**, 1145-1155.
- 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.
- Torsney, B. and López-Fidalgo, J. (1995). MV-optimization in simple linear regression. In *MODA4-Advances in Model-Oriented Data Analysis* (Edited by C. P. Kitsos, W.G. Müller), 71-76.
- Wu, C. F. J. (1985). Efficient sequential designs with binary data. *J. Amer. Statist. Assoc.* **80**, 974-984.
- Wu, C. F. J. (1988). Optimal design for percentile estimation of a quantal response curve. In *Optimal Design and Analysis of Experiments* (Edited by J. Dodge, V. V. Fedorov, H. P. Wynn), 213-233. North Holland, Amsterdam.

Ruhr-Universität Bochum, Fakultät für Mathematik, Universitätsstr. 150, D-44780 Bochum, Germany.

E-mail: holger.dette@rz.ruhr-uni-bochum.de

E-mail: michael.sahm@ruhr-uni-bochum.de

(Received May 1997; accepted October 1997)