

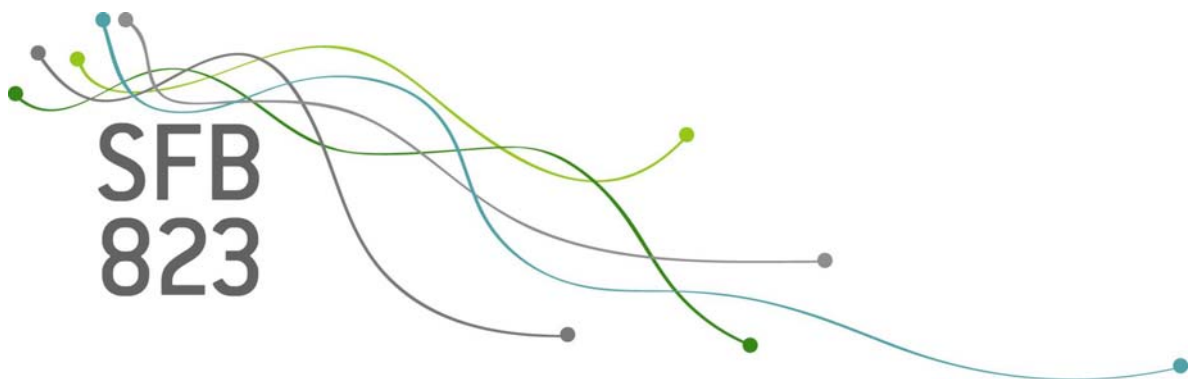
SFB
823

Optimal designs for quantile regression models

Holger Dette, Matthias Trampisch

Nr. 26/2011

Discussion Paper



OPTIMAL DESIGNS FOR QUANTILE REGRESSION MODELS

HOLGER DETTE / MATTHIAS TRAMPISCH

ABSTRACT. Despite of their importance optimal designs for quantile regression models have not been developed so far. In this paper we investigate the D -optimal design problem for the location scale nonlinear quantile regression model. We provide a necessary condition to check for the optimality of a given design and use it to determine bounds for the number of support points of locally D -optimal designs. The results are illustrated determining locally, Bayesian and standardized maximin D -optimal designs for quantile regression analysis in the Michaelis-Menten and EMAX model, where the location and the scale function are related by a known link function.

1. INTRODUCTION

Quantile regression was introduced by Koenker and Bassett (1978) as an alternative to least squares estimation and yields a far-reaching extension of regression analysis by estimating families of conditional quantile curves. Since its introduction, quantile regression has found great attraction in statistics because of its ease of interpretation, its robustness and its numerous applications which include such important areas as medicine, economics, environment modeling, toxicology or engineering [see Buchinsky (1994); Cade et al. (1999) or Wei et al. (2006) among many others]. For a detailed description of quantile regression analysis we refer to the monograph Koenker (2005), which also provides a variety of additional examples. We focus on the general univariate nonlinear quantile regression model

$$(1.1) \quad y(x) = g(x, \theta) + \sigma(x, \theta)\epsilon,$$

where the functions g and σ represent the location and scale function, respectively, $\theta^T = (\theta_1, \dots, \theta_p)$ is a vector of unknown parameters which varies in a compact parameter space, say Θ , and ϵ denotes a random variable with distribution function F which has τ -quantile 0, that is $F^{-1}(\tau) = 0$. The quantity x denotes the explanatory variable which can vary in a compact design space, say \mathcal{X} . We assume that in principle for each experimental condition x an observation $y(x)$ is available according to the model (1.1), where different observations are independent. This means that the conditional τ -quantile of $y(x)$ at experimental condition $x \in \mathcal{X}$ is given by $g(x, \theta)$. The model (1.1) has found considerable attention in the literature. Koenker and Zhao (1994) and Zhou and Portnoy (1998) proposed a weighted quantile regression estimate for the parameter θ where the (random) weights are determined from a preliminary estimate for the parameter of the scale function. Alternatively, one can use unweighted quantile regression to estimate the parameters of the conditional quantile function for heteroscedastic data, and in the present paper we investigate design problems arising in the context of this estimation method.

For least squares and maximum likelihood estimation it is well known that the efficiency of statistical analysis can be improved substantially by the application of optimal designs. Such designs

Key words and phrases. quantile regression, Michaelis-Menten model, EMAX model, heteroscedasticity, locally optimal design, robust designs

AMS Subject classification: 62K05, 62J02.

result in a substantial reduction of costs and are therefore important from an ethical and economical point of view [see Atkinson (1996)]. For example, it was shown by Dette et al. (2008) in the context of estimating the minimum effective dose in a phase II dose finding study that an optimal design can result in a 50% reduction of sample size without losing any accuracy in the statistical inference. Despite of these advantages for least squares and maximum likelihood estimation there exist - to the knowledge of the authors - no results on optimal designs for statistical analysis with quantile regression models.

It is the purpose of the present paper to determine optimal designs in quantile regression models. In Section 2 we introduce the optimal design problem in the context of nonlinear quantile regression. For the sake of transparency, we focus on the D -optimality criterion, which determines the design such that the (asymptotic) confidence ellipsoid for the parameters in the conditional quantile function has minimal volume. It turns out that - in contrast to classical design theory for maximum likelihood and least squares estimation - the optimal design problems resulting from the quantile regression model (1.1) are not convex. Therefore standard theory is not directly applicable which makes the solution substantially harder. In the first part of Section 3 we present some general results for the D -optimality criterion. The resulting designs are locally optimal designs in the sense of Chernoff (1953) because they require a specification of the unknown parameters. There are many situations where such preliminary knowledge about the parameters is available, such that the application of locally optimal designs is well justified [such as phase II dose finding trials, see Dette et al. (2008)]. Moreover, locally optimal designs usually serve as a benchmark for many commonly used designs and are the basis for more sophisticated design strategies, which require less precise knowledge about the model parameters, such as Bayesian or standardized maximin optimality criteria [see Chaloner and Verdinelli (1995) and Dette (1997) among others]. Optimal designs with respect to these robust criteria are discussed in the second part of Section 3.

We illustrate the new concepts in the construction of optimal designs for quantile regression analysis in Section 4. In particular, we determine locally optimal, Bayesian and standardized maximin optimal designs for the Michaelis-Menten and EMAX quantile regression model, where the scale and the location function are related by a known link function. The benefits of an efficient design of experiment are demonstrated in Section 5, where we compare the efficiency of a commonly used uniform design with the designs derived in this paper in an explicit application. Finally, all technical details are deferred to an appendix in Section 6.

2. OPTIMAL DESIGNS FOR QUANTILE REGRESSION

Let $x_1, \dots, x_N \in \mathcal{X}$ denote fixed predictors and suppose that N observations $y(x_1), \dots, y(x_N)$ at these experimental conditions are available. We assume that the observations are generated by the model (1.1), that is

$$(2.1) \quad y(x_i) = g(x_i, \theta) + \sigma(x_i, \theta)\varepsilon_i; \quad i = 1, \dots, N,$$

where $\varepsilon_1, \dots, \varepsilon_N$ are independent identically distributed random variables with distribution function F and τ -quantile 0, i.e. $F^{-1}(\tau) = 0$.

For a given $\tau \in (0, 1)$, the (unweighted) quantile regression estimate of the parameter θ in model (2.1) is defined by

$$(2.2) \quad \hat{\theta}_N(\tau) = \arg \min_{\theta \in \Theta} \sum_{i=1}^N \rho_\tau(y(x_i) - g(x_i, \theta)),$$

where ρ_τ denotes the check function $\rho_\tau(u) = u(\tau - I(u < 0))$ [see Koenker (2005)]. Throughout this paper we assume that the location function g is differentiable with respect to the parameter θ and denote by

$$(2.3) \quad \dot{g}(x, \theta) = \left(\frac{\partial g(x, \theta)}{\partial \theta_1}, \dots, \frac{\partial g(x, \theta)}{\partial \theta_p} \right)^T$$

the corresponding vector of partial derivatives. Additionally, we assume that the following standard assumptions of nonlinear quantile regression are satisfied [see Koenker (2005)].

- **A1.** The distribution function F of the random variables $\varepsilon_1, \dots, \varepsilon_N$ in model (2.1) is absolutely continuous, with continuous density f such that $0 < f(0) < \infty$.
- **A2.** There exist positive definite matrices D_0 and D_1 such that the following limits exist
 - (1) $\lim_{N \rightarrow \infty} N^{-1} \sum_{i=1}^N \dot{g}(x_i, \theta) \dot{g}^T(x_i, \theta) = D_0$
 - (2) $\lim_{N \rightarrow \infty} N^{-1} f(0) \sum_{i=1}^N \frac{1}{\sigma(x_i, \theta)} \dot{g}(x_i, \theta) \dot{g}^T(x_i, \theta) = D_1$
 - (3) $\max_{i=1, \dots, N} \|\dot{g}(x_i, \theta)\| / \sqrt{N} \rightarrow 0$
- **A3.** There exist constants $k_0, k_1 > 0$ and $N_0 \in \mathbb{N}$ such that, for $\theta_1, \theta_2 \in \Theta$, and $N > N_0$,

$$k_0 \|\theta_1 - \theta_2\| \leq \left(N^{-1} \sum_{i=1}^N ((g(x_i, \theta_1) - g(x_i, \theta_2))^2) \right)^{1/2} \leq k_1 \|\theta_1 - \theta_2\|$$

Under these conditions it can be shown [see Koenker (2005)] that the estimate defined by (2.2) is asymptotically normal distributed, that is

$$(2.4) \quad \sqrt{n} \left(\hat{\theta}_N(\tau) - \theta(\tau) \right) \xrightarrow{\mathcal{D}} N \left(0, \tau(1 - \tau) D_1^{-1} D_0 D_1^{-1} \right),$$

where the symbol $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution. We observe that the asymptotic variance of the quantile regression estimate (2.2) depends on the experimental conditions x_i and a “good” experimental design should provide “small” variances of this estimate. Optimal design problems of this type have mainly been considered in the context of maximum likelihood and least squares estimation and – to the best knowledge of the authors – no results are available for quantile regression. The following discussion provides a first insight in optimal design problems for these models.

In order to find optimal designs for quantile regression analysis, we consider approximate designs in the sense of Kiefer (1974), which are defined as probability measures on the design space \mathcal{X} with finite support. The support points of an (approximate) design ξ give the locations where observations are taken, while the weights give the corresponding relative proportions of total observations to be taken at these points. If the design ξ has masses $w_i > 0$ at the different points x_i ($i = 1, \dots, k$) and N observations can be made by the experimenter, the quantities $w_i N$ are rounded to integers, say N_i , satisfying $\sum_{i=1}^k N_i = N$, and the experimenter takes N_i observations at each location x_i ($i = 1, \dots, k$). In this case the assumptions A1, A2 and A3 are satisfied if $\dot{g}(x_i, \theta)$ and $\sigma(x_i, \theta)$ are well-defined and $\sigma(x_i, \theta) > 0$ for $i = 1, \dots, k$. The asymptotic covariance

matrix in (2.4) is given by

$$\frac{\tau(1-\tau)}{f^2(0)}H(\xi, \theta),$$

where f denotes the density of the errors, the matrix H is defined by

$$(2.5) \quad H(\xi, \theta) = D_1^{-1}(\xi, \theta)D_0(\xi, \theta)D_1^{-1}(\xi, \theta)$$

and the matrices $D_0(\xi, \theta)$ and $D_1(\xi, \theta)$ are given by

$$(2.6) \quad D_0(\xi, \theta) = \int_{\mathcal{X}} \dot{g}(x, \theta)\dot{g}^T(x, \theta)d\xi(x)$$

and

$$(2.7) \quad D_1(\xi, \theta) = \int_{\mathcal{X}} \frac{1}{\sigma(x, \theta)}\dot{g}(x, \theta)\dot{g}^T(x, \theta)d\xi(x).$$

Consequently, an optimal (approximate) design for quantile regression analysis in model (1.1) maximizes an appropriate concave functional of the inverse $H^{-1}(\xi, \theta)$ of the asymptotic covariance matrix, and there are numerous criteria which can be used for discriminating between competing designs [see Silvey (1980) or Pukelsheim (2006) among others]. In contrast to classical optimal design theory the mapping

$$\xi \longrightarrow H^{-1}(\xi, \theta)$$

is in general not concave. This fact makes the optimal design problem for quantile regression analysis considered in this paper substantially harder, because standard convex optimization theory is in general not applicable. For the sake of transparency, we concentrate here on the well known D -optimality criterion, which determines the optimal design such that the determinant

$$(2.8) \quad \det(H^{-1}(\xi, \theta)) = \left(\det D_1(\xi, \theta)\right)^2 \det D_0^{-1}(\xi, \theta)$$

is maximal. A design ξ_θ^* which maximizes (2.8) is called locally D -optimal for quantile regression analysis in model (2.1) and minimizes the volume of the asymptotic confidence ellipsoid for the parameter θ with respect to the choice of the design ξ . Note that in nonlinear regression models the information matrix $H(\xi, \theta)$ - and as a consequence the corresponding optimal designs - depend on the unknown parameter vector θ , which justifies the name locally optimal designs [see Chernoff (1953)]. These designs require an initial guess of the unknown parameters in the model, and there are numerous situations where such knowledge is available [see for example Dette et al. (2008)]. Moreover, locally optimal designs serve as benchmarks for many commonly used designs, and they form the basis for many robust optimality criteria based on the Bayesian and maximin principle. To be precise, let ξ_θ^* denote the locally D -optimal design for quantile regression analysis and define

$$(2.9) \quad \text{eff}(\xi, \theta) = \left(\frac{\det H(\xi_\theta^*, \theta)}{\det H(\xi, \theta)}\right)^{1/p} \in [0, 1]$$

as the D -efficiency of a given design ξ . Let π denote a prior distribution on the parameter space Θ ; following Chaloner and Verdinelli (1995) we call a design ξ_π^* Bayesian D -optimal design for

quantile regression analysis in model (2.1) if it maximizes the functional

$$(2.10) \quad \begin{aligned} \Phi_\pi(\xi) &= \int_{\Theta} \log(\text{eff}(\xi, \theta)) \pi(d\theta) \\ &= c_\Theta + \frac{1}{p} \int_{\Theta} \left(2 \log(\det D_1(\xi, \theta)) - \log(\det D_0(\xi, \theta)) \right) \pi(d\theta), \end{aligned}$$

where the constant c_Θ does not depend on the design ξ . Similarly, a (standardized) maximin D -optimal design ξ_{MM}^* for quantile regression analysis in model (2.1) maximizes the functional

$$(2.11) \quad \Psi_\Theta(\xi) = \min\{\text{eff}(\xi, \theta) \mid \theta \in \Theta\}$$

[see Dette (1997)]. Bayesian and maximin optimal designs have been investigated in the context of least squares and maximin likelihood estimation [see Chaloner and Verdinelli (1995); Haines (1995) or Imhof (2001) among others]. Corresponding results for quantile regression analysis are discussed in Section 3 and 4. Furthermore, we illustrate the benefits of robust design strategies in Section 5 reanalyzing a data example.

3. D-OPTIMAL DESIGNS

In this section, we investigate D -optimal designs for the estimating the parameters in the quantile regression model (1.1), which maximize the determinant of the matrix $H^{-1}(\xi, \theta)$ as defined in (2.5). Note that in the case of homoscedasticity ($\sigma(x, \theta) = \sigma$) we have $\sigma D_1(\xi, \theta) = D_0(\xi, \theta)$, and the optimization problem reduces to a “classical” D -optimal design problem for least squares or maximum likelihood estimation in a nonlinear regression model [see for example Ford et al. (1992); He et al. (1996) or Fang and Hedayat (2008) among many others]. However, for non-constant scale functions the optimal design problem for quantile and least squares regression are different. In particular, optimal design problems for quantile regression with non-constant scale function are substantially more difficult to solve because they are usually not convex. In order to investigate D -optimal designs in this context recall the definition of the matrices D_0 and D_1 in (2.6) and (2.7), and define

$$(3.1) \quad d(x, \xi, \theta) = \frac{2}{\sigma(x, \theta)} d_1(x, \xi, \theta) - d_0(x, \xi, \theta)$$

as generalized dispersion function, where the functions d_0 and d_1 are given by

$$d_i(x, \xi) = \dot{g}^T(x, \theta) D_i^{-1}(\xi, \theta) \dot{g}(x, \theta), \quad i = 0, 1.$$

This function plays a fundamental role in the following theorem, which gives a necessary condition for a design to be locally D -optimal for quantile regression analysis in model (2.1). The proof can be found in the Appendix.

Theorem 3.1. *Any locally D -optimal design design ξ_θ^* for quantile regression analysis in model (2.1) satisfies the inequality*

$$(3.2) \quad d(x, \xi_\theta^*, \theta) \leq p$$

for all $x \in \mathcal{X}$. Moreover, the maximum in (3.2) is attained for any support point of the locally D -optimal design.

Throughout this paper we call a design with k support points a k -point design and an optimal k -point design ξ^* maximizes an appropriate concave functional of the matrix $H^{-1}(\xi, \theta)$ defined in

(2.5) in the class of all k -point designs. Designs which have the same number of support points as the dimension p of the parameter vector θ are called saturated designs. The following result shows that locally D -optimal saturated designs are always uniform designs and is proved in the Appendix.

Lemma 3.2. *The locally D -optimal saturated design for quantile regression analysis in model (2.1) has equal masses at its support points.*

We conclude this section with a statement of a corresponding necessary condition for Bayesian and standardized maximin D -optimality. The results are proved by similar arguments as given in the proof of Theorem 3.1 and the proofs are therefore omitted.

Theorem 3.3.

- (a) *If the design ξ_π^* is Bayesian D -optimal with respect to the prior π for quantile regression analysis in model (2.1), then the inequality*

$$(3.3) \quad \int_{\Theta} d(x, \xi_\pi^*, \theta) \pi(d\theta) = \int_{\Theta} \dot{g}^T(x, \theta) \left\{ 2 \frac{D_1^{-1}(\xi_\pi^*, \theta)}{\sigma(x, \theta)} - D_0^{-1}(\xi_\pi^*, \theta) \right\} \dot{g}(x, \theta) \pi(d\theta) \leq p$$

holds for all $x \in \mathcal{X}$. Moreover, in this case there is equality in (3.3) for all support points of the Bayesian D -optimal design ξ_π^ .*

- (b) *If the design ξ_{MM}^* is standardized maximin D -optimal for quantile regression analysis in model (2.1), then there exists a distribution π^* on the set*

$$(3.4) \quad \Pi(\xi_{\text{MM}}^*) = \{\theta \in \Theta \mid \text{eff}(\xi_{\text{MM}}^*, \theta) = \Psi_{\Theta}(\xi_{\text{MM}}^*)\}$$

such that the inequality

$$(3.5) \quad \int_{\Pi(\xi_{\text{MM}}^*)} d(x, \xi_{\text{MM}}^*, \theta) \pi^*(d\theta) \leq p$$

holds for all $x \in \mathcal{X}$. Moreover, there is equality in (3.5) for all support points of the standardized maximin D -optimal design ξ_{MM}^ .*

4. APPLICATION TO NONLINEAR MODELS

In this section, we investigate locally D -optimal designs for quantile regression analysis for two nonlinear models which are widely used in applications, namely the Michaelis-Menten model

$$(4.1) \quad g_1(x, \theta) = \frac{\theta_1 x}{\theta_2 + x}$$

and the EMAX model

$$(4.2) \quad g_2(x, \theta) = \theta_0 + \frac{\theta_1 x}{\theta_2 + x},$$

where $x \in \mathcal{X} = [x_l, x_u] \subset \mathbb{R}_0^+$. In the Michaelis-Menten model (4.1) the function g_1 represents the velocity of a chemical reaction, the parameter $\theta_1 \geq 0$ denotes the maximum velocity, the predictor $x \geq 0$ reflects the concentration of a substrate and the parameter $\theta_2 \geq 0$ is the half-saturated constant, the concentration x where the velocity is half-maximal. Specific applications of the Michaelis-Menten model in such important areas as pharmacology, medical research and biological

sciences, particularly in biochemistry and in enzyme reaction studies, can be found in Johansen (1984) or Cornish-Browden (1995) among many others.

The EMAX model (4.2) is frequently used in pharmaceutical studies to model a dose-response relationship of a drug where the maximum effect is achieved asymptotically at large dose levels. The model is used to describe the relationship of drug-receptor interactions and therefore deduced from the chemical equilibrium equation [see Boroujerdi (2002)]. Various applications of the EMAX model have been discussed by Demana et al. (1997) among others. In model (4.2) the parameter $\theta_0 \geq 0$ usually represents the response at placebo, $\theta_1 \geq 0$ is the maximum achievable increase above the placebo response and $\theta_2 \geq 0$ is the dose which produces 50% of the effect.

In the following two subsections we investigate properties of locally D -optimal designs for quantile regression analysis in model (4.1) and (4.2), where the location function $g(x, \theta)$ and the scale function $\sigma(x, \theta)$ are related by a known link function, say h , i.e.

$$(4.3) \quad \sigma(x, \theta) = h(g(x, \theta))$$

for some differentiable function h . Robust design strategies are discussed in Section 4.3

4.1. Locally D -optimal designs for the Michaelis-Menten. Optimal designs for maximum likelihood or least squares estimation in the Michaelis-Menten model have been studied by numerous authors, see e.g. Dunn (1988); Rasch (1990); Dette and Wong (1999); Lopez-Fidalgo and Wong (2002) or Dette and Biedermann (2003) among many others. The following results give some first insights into the structure of locally D -optimal saturated designs in the Michaelis-Menten model (4.1) with scale functions of the form (4.3) if quantile regression is used to estimate the parameters. All proofs are deferred to the Appendix, see Section 6.2.2.

Theorem 4.1. *Any locally D -optimal saturated design ξ_θ^* on the design space $\mathcal{X} = [x_l, x_u] \subset \mathbb{R}_0^+$ for quantile regression analysis in the Michaelis-Menten model (4.1) with scale function (4.3) has equal masses at its support points. The design ξ_θ^* is supported at the point x_u if the inequality*

$$(4.4) \quad \frac{2}{g_1(x, \theta)} \geq \frac{h'(g_1(x, \theta))}{h(g_1(x, \theta))},$$

holds for all $x \in \mathcal{X}$. Moreover, if additionally the inequality

$$(4.5) \quad \frac{\partial}{\partial x} \left(\frac{h'(g_1(x, \theta))}{h(g_1(x, \theta))} \right) \geq 0$$

holds for all $x \in \mathcal{X}$, then the locally D -optimal saturated design ξ_θ^* is unique and supported at the points $\max\{t^{-1}(z^*), x_l\}$ and x_u , where z^* is the unique solution of the equation

$$(4.6) \quad l(z) = \frac{1}{z} - \frac{1}{t(x_u) - z} - \frac{h'(z)}{h(z)} = 0$$

and the function t is defined by $t(x) = \theta_1 x / (\theta_2 + x)$.

It is easy to see that the assumption of Theorem 4.1 is satisfied for the link functions

$$(4.7) \quad h_1(z) = 1/z^n, n \geq 0$$

$$(4.8) \quad h_2(z) = \exp(-nz), n \geq 0.$$

Note that the case of a homoscedastic quantile regression model appears for the choice $n = 0$ in both cases. Moreover, if $n < 0$, the response at $x = 0$ is identically 0 and we assume for notational

convenience throughout this paper that $x_l > 0$ if $n < 0$. Note also that an optimal design for the Michaelis-Menten model never advises the experimenter to take observations at $x = 0$ because the corresponding conditional τ -quantile is 0 and contains no information about the unknown parameters.

In the following discussion we derive some explicit optimal designs for quantile regression analysis in the Michaelis-Menten model for this type of link functions. We begin with a result on the number of support points of the locally D -optimal design for the link function (4.7).

Theorem 4.2. *The locally D -optimal design on the design space $\mathcal{X} = [x_l, x_u] \subset \mathbb{R}_0^+$ for quantile regression analysis in the Michaelis-Menten model (4.1) with scale function satisfying (4.3) and (4.7) is supported at at most 3 support points, where one of the support points is located at the upper boundary point of the design space \mathcal{X} .*

Moreover, if $n \in \{-2, -1, 0, 1\}$, the locally D -optimal design for the Michaelis-Menten model (4.1) with scale function satisfying (4.3) and (4.7) is supported at exactly two points including the upper boundary point x_u of the design space.

By the previous result any locally D -optimal design for quantile regression analysis in the Michaelis-Menten model with link function (4.7) is supported at at most three support points. If $n \in \{-2, -1, 0, 1\}$ a saturated design is also locally D -optimal in the class of all designs. Moreover, in an extensive numerical study we investigated the general case where n is not necessarily an integer satisfying $n \geq -2$. We could not find any value of n where the design is supported at three or more support points. A similar observation can be made for the link function (4.8) and we state the following conjecture.

Conjecture 4.1. *The locally D -optimal design on the design space $\mathcal{X} = [x_l, x_u] \subset \mathbb{R}_0^+$ for quantile regression analysis in the Michaelis-Menten model (4.1) with scale function satisfying (4.3) and (4.7) with $n \geq -2$ or (4.8) with $n \geq 0$ is supported at exactly 2 support points, where one of the support points is located at the upper boundary point of the design space.*

We now derive analytical results for D -optimal saturated designs in the Michaelis-Menten model with link functions h_1 and h_2 defined in (4.7) and (4.8), respectively.

Theorem 4.3. *Consider the Michaelis-Menten (4.1) model on the design space $\mathcal{X} = [x_l, x_u] \subset \mathbb{R}_0^+$ with scale function satisfying (4.3).*

- (a) *If the link function is given by (4.7) then the locally D -optimal saturated design for quantile regression analysis has equal masses at the points*

$$x_1^* = \max \left\{ \frac{(n+1)x_u\theta_2}{(n+2)(\theta_2 + x_u)}, x_l \right\}, \quad \text{and} \quad x_2^* = x_u.$$

If $n \in \{-2, -1, 0, 1\}$ this design is also locally D -optimal in the class of all designs, where the cases $n = -1$ and $n = -2$ are interpreted as $x_1^ = x_l$.*

- (b) *If the link function is given by (4.8) then the locally D -optimal saturated design for quantile regression analysis has equal masses at the points*

$$x_1^* = \max \left\{ \frac{-2\theta_2 + \theta_1 n x_u + \sqrt{(2\theta_2 + 2x_u)^2 + (\theta_1 n x_u)^2}}{2(\theta_1 n + 2 + x_u/\theta_2)}, x_l \right\} \quad \text{and} \quad x_2^* = x_u.$$

Note that for the case $n = 0$ Theorem 4.3 yields in both cases the locally D -optimal design ξ_θ^* for the homoscedastic Michaelis-Menten model with equal masses at the points

$$x_1^* = \max\left\{\frac{\theta_2 x_u}{2\theta_2 + x_u}, x_l\right\} \quad \text{and} \quad x_2^* = x_u,$$

which coincides with the locally D -optimal design for maximum likelihood estimation under normality assumption [see Rasch (1990)]. On the other hand, if $n > 0$ the locally D -optimal saturated designs for quantile regression analysis in the Michaelis-Menten model with link functions h_1 and h_2 defined in (4.7) and (4.8) are different.

4.2. Locally D -optimal designs for the EMAX model. Optimal designs for least squares and maximum likelihood estimation in the EMAX model have been studied by various authors [see Merle and Mentre (1995); Wang (2006); Dette et al. (2008) or Fang and Hedayat (2008) among many others]. Our first result establishes some general properties of locally D -optimal saturated designs for quantile regression analysis in model (4.2) with variance function of the form (4.3).

Lemma 4.4. *Any locally D -optimal saturated design ξ_θ^* on the design space $\mathcal{X} = [x_l, x_u] \subset \mathbb{R}_0^+$ for quantile regression analysis in the EMAX model (4.2) with variance function defined in (4.3) has the point x_u as support point if*

$$(4.9) \quad \frac{2}{g_2(x, \theta)} \geq \frac{h'(g_2(x, \theta))}{h(g_2(x, \theta))}$$

holds for all $x \in \mathcal{X}$.

We now provide some analytical results for D -optimal saturated designs in the EMAX model (4.2), if the link function is defined by (4.7) or (4.8).

Theorem 4.5. *In the EMAX model (4.2) model on the design space $\mathcal{X} = [x_l, x_u] \subset \mathbb{R}_0^+$ with scale function satisfying (4.3) the locally D -optimal saturated design for quantile regression analysis has equal masses at its support points.*

Moreover, if the link function is given by (4.7) define

$$(4.10) \quad \kappa = \frac{\theta_2 (a (a - \sqrt{3}) \theta_1 x_u - (9 + \sqrt{3}a + 6n) \theta_0 (\theta_2 + x_u))}{\sqrt{3}a\theta_0\theta_2 + (3 + 2n)(3\theta_0 + 2(2 + n)\theta_1)\theta_2 + (9 + \sqrt{3}a + 6n) (\theta_0 + \theta_1)x_u}$$

with $a = \sqrt{3 + 8n + 4n^2}$.

(a) *If $\kappa \leq 0$, then the support points of the locally D -optimal saturated design for quantile regression analysis are given by $x_1^* = x_l$, $x_3^* = x_u$ and*

$$(4.11) \quad x_2^* = \theta_2 \frac{(1 + n)x_+ - \theta_0 + \sqrt{((1 + n)x_+ - \theta_0)^2 - (2 + n)(nx_- - 2\theta_0 x_+)}}{(2 + n)\theta_1 - (1 + n)x_+ + \theta_0 - \sqrt{((1 + n)x_+ - \theta_0)^2 - (2 + n)(nx_- - 2\theta_0 x_+)}}$$

where

$$(4.12) \quad x_+ = \frac{1}{2} \left(\frac{\theta_1 x_l}{\theta_2 + x_l} + \frac{\theta_1 x_u}{\theta_2 + x_u} \right) \quad \text{and} \quad x_- = \frac{\theta_1 x_l}{\theta_2 + x_l} \cdot \frac{\theta_1 x_u}{\theta_2 + x_u}.$$

- (b) If $0 \leq x_l \leq \kappa$, then the support points of the locally D -optimal saturated design for quantile regression analysis in the EMAX model are given by $x_1^* = \kappa$, $x_3^* = x_u$ and

$$x_2^* = \theta_2 \frac{(3 + 4n(2 + n))\theta_1 x_u - b}{(3 + 2n)(2(2 + n)\theta_2 + 3x_u)\theta_1 + b}$$

with $b = 3(3 + 2n)\theta_0(\theta_2 + x_u) - (\theta_0\theta_2 + (\theta_0 + \theta_1)x_u)\sqrt{9 + 12n(2 + n)}$.

- (c) If $0 < \kappa \leq x_l$, then the support points of the locally D -optimal saturated design for quantile regression analysis in the EMAX model are given by $x_1^* = x_l$, $x_3^* = x_u$ and x_2^* is defined in (4.11).

Theorem 4.6. In the EMAX model (4.2) model on the design space $\mathcal{X} = [x_l, x_u] \subset \mathbb{R}_0^+$ with scale function satisfying (4.3) the locally D -optimal saturated design for quantile regression analysis has equal masses at its support points. Moreover, if the link function is given by (4.8) define

$$\kappa = -\theta_2 \frac{(3 + \sqrt{3})(\theta_2 + x_u) - 2n\theta_1 x_u}{(3 + \sqrt{3})(\theta_2 + x_u) + 2n\theta_1\theta_2}$$

- (a) If $\kappa \leq 0$ then the support points of the locally D -optimal saturated design for quantile regression analysis in the EMAX model are given by $x_1^* = x_l$, $x_3^* = x_u$ and

$$(4.13) \quad x_2^* = \theta_2 \frac{nx_+ - 1 + \sqrt{1 + n^2 x_-^2}}{n(\theta_1 - x_+) + 1 - \sqrt{1 + n^2 x_-^2}}$$

where x_+ is defined (4.12) and

$$x_- = \frac{1}{2} \left(\frac{\theta_1 x_l}{\theta_2 + x_l} - \frac{\theta_1 x_u}{\theta_2 + x_u} \right)$$

- (b) If $0 \leq x_l \leq \kappa$ then the support points of the locally D -optimal saturated design for quantile regression analysis in the EMAX model are given by $x_1^* = \kappa$, $x_3^* = x_u$ and

$$x_2^* = -\theta_2 \frac{(-3 + \sqrt{3})(\theta_2 + x_u) + 2n\theta_1 x_u}{(-3 + \sqrt{3})(\theta_2 + x_u) - 2n\theta_1\theta_2}$$

- (c) If $0 \leq \kappa \leq x_l$ then the support points of the locally D -optimal saturated design for quantile regression analysis in the EMAX model are given by $x_1^* = x_l$, $x_3^* = x_u$ and the third support point x_2^* is given by (4.13).

It is worthwhile to mention that for $n = 0$ the value of κ in Theorem 4.5 and Theorem 4.6 is always negative, and equation (4.11) or equation (4.13) (using l'Hopital's rule) yield the locally D -optimal saturated design ξ^* with design points

$$x_1^* = x_l, \quad x_2^* = \frac{x_u(x_l + \theta_2) + x_l(x_u + \theta_2)}{(x_l + \theta_2) + (x_u + \theta_2)} \quad \text{and} \quad x_3^* = x_u.$$

It follows by similar arguments as given in Dette et al. (2010) that this design is also locally D -optimal in the class of all designs. On the other hand, if $n > 0$, the locally D -optimal saturated designs for the link functions h_1 and h_2 defined in (4.7) and (4.8) are different.

4.3. Robust designs for the Michaelis-Menten model. In this section, we illustrate how standardized maximin and Bayesian D -optimal designs for quantile regression analysis can be constructed. For the sake of brevity we restrict ourselves to the case of the Michaelis-Menten

model (4.1) with link function (4.7). The following results give the Bayesian and the standardized maximin D -optimal saturated designs for quantile regression analysis for this model on the design space $\mathcal{X} = [0, x_u]$.

Theorem 4.7. *Consider the the Michaelis-Menten model on the design space $\mathcal{X} = [0, x_u]$ with scale function defined by (4.7) such that $n > -1$, and let π denote a prior on the set $\Theta = [0, \infty) \times [0, \infty)$ with marginals π_1 and π_2 such that $\text{supp}(\pi_2) \subset [\bar{x} \cdot x_u, \infty)$, where $\bar{x} = \frac{1}{2} \left(-1 + \sqrt{-9 + 6\sqrt{3}} \right) \approx 0.0899798$. Furthermore, let the integral $\int_{\Theta} |\log \theta_2| \pi_2(d\theta)$ be finite. The Bayesian D -optimal saturated design for quantile regression analysis with respect to the prior π has equal masses at the points x_u and the point x^* , which is the unique solution of the equation*

$$(4.14) \quad -\frac{1}{1-x} + \frac{1+n}{x} - (2+n) \int_{\Theta} \frac{\pi_2(d\theta)}{\theta_2/x_u + x} = 0$$

in the interval $(0, x_u)$.

Theorem 4.8. *Consider the the Michaelis-Menten model on the design space $\mathcal{X} = [0, x_u]$ with scale function defined by (4.7). The standardized maximin D -optimal saturated design ξ_{MM} for quantile regression analysis with respect to the set $\Theta = [0, \infty) \times [\theta_{2,l}, \theta_{2,u}]$ has equal masses at the points x_u and the point*

$$(4.15) \quad x^* = \frac{\theta_{2,u} \left(\theta_{2,l} (x_u + \theta_{2,l})^{n+1} \right)^{\frac{1}{2+n}} - \theta_{2,l} \left(\theta_{2,u} (x_u + \theta_{2,u})^{n+1} \right)^{\frac{1}{2+n}}}{\left(\theta_{2,u} (x_u + \theta_{2,u})^{n+1} \right)^{\frac{1}{2+n}} - \left(\theta_{2,l} (x_u + \theta_{2,l})^{n+1} \right)^{\frac{1}{2+n}}}.$$

Moreover, for this design the set $\Pi(\xi^*)$ defined in (3.4) is given by

$$\mathcal{N}(\xi_{\text{MM}}) = \left\{ \frac{\theta_{2,l}}{x_u}, \frac{\theta_{2,u}}{x_u} \right\}.$$

5. DATA EXAMPLE

In this section, we reanalyze data from hormone receptor assays, where the relationship between the response and explanatory variable is usually modeled by the Michaelis-Menten equation [see Cressie and Keightley (1979)]. These authors investigated estrogen receptor assays, which are used in the study of human breast cancer. The response Y represents the amount of estrogen bound to receptor and x is the amount of hormone not bound to receptor and the Michaelis-Menten model was used to evaluate the data. In the specific example the range for the variable x is the design space $\mathcal{X} = [0, x_u] = [0, 2000]$.

To be precise, consider model (4.1) with link function (4.7), where the assumed design space is given by $\mathcal{X} = [0, 2000]$. Robust designs for maximum likelihood estimation have been determined by Dette and Biedermann (2003) under the assumption of homoscedastic data ($n = 0$), and in this section we investigate designs for quantile regression analysis. We consider the optimal design problem under the assumptions of homoscedasticity ($n = 0$) and two cases of heteroscedasticity ($n = 1$ and $n = 5$) where two possible regions for the parameter θ are considered, namely

$$(5.1) \quad \Theta_1 = [0, \infty) \times [100, 2000]$$

$$(5.2) \quad \Theta_2 = [0, \infty) \times [500, 5000].$$

Θ	n	ξ_{MM}		eff	ξ_{MM}^*			eff
Θ_1	0	267.4	2000	0.7208	109.6	635.8	2000	0.7925
		.500	.500		.235	.321	.444	
Θ_1	1	499.2	2000	0.6469	211.2	846.3	2000	0.7438
		.500	.500		.198	.353	.449	
Θ_1	5	1,041.0	2000	0.5733	489.0	1,256.8	2000	0.6199
		.500	.500		.107	.430	.463	
Θ_2	0	548.6	2000	0.9052	548.6	2000		0.9052
		.500	.500		.500	.500		
Θ_2	1	872.0	2000	0.8756	872.0	2000		0.8756
		.500	.500		.500	.500		
Θ_2	5	1,408.1	2000	0.8433	1,408.1	2000		0.8433
		.500	.500		.500	.500		

TABLE 1. Standardized maximin D -optimal saturated design ξ_{MM} derived by Theorem 4.8 and standardized maximin D -optimal design ξ_{MM}^* (determined in the class of all designs) for quantile regression analysis in the Michaelis-Menten model. The corresponding minimum D -efficiency (2.11) is listed in the columns labeled 'eff'.

Θ	n	$\xi_{\pi_1}^*$		$\xi_{\pi_2}^*$		$\xi_{\pi_3}^*$	
Θ_1	0	451.2	2000	552.5	2000	359.5	2000
		.500	.500	.500	.500	.500	.500
Θ_1	1	754.4	2000	871.8	2000	630.0	2000
		.500	.500	.500	.500	.500	.500
Θ_1	5	1,306.8	2000	1,402.3	2000	1,183.1	2000
		.500	.500	.500	.500	.500	.500
Θ_2	0	686.0	2000	759.4	2000	615.0	2000
		.500	.500	.500	.500	.500	.500
Θ_2	1	1,028.7	2000	1,103.0	2000	948.9	2000
		.500	.500	.500	.500	.500	.500
Θ_2	5	1,526.4	2000	1,575.0	2000	1,467.6	2000
		.500	.500	.500	.500	.500	.500

TABLE 2. Bayesian D -optimal designs with respect to the priors $\pi_i, i = 1, 2, 3$ for quantile regression analysis in the Michaelis-Menten model. All saturated designs are Bayesian D -optimal in the class of all designs.

Note that the efficiency (2.9) does not depend on the parameter θ_1 . We calculate the standardized maximin D -optimal saturated design ξ_{MM} derived by Theorem 4.8 and the standardized maximin D -optimal design in the class of all designs by numerical methods, which is denoted by ξ_{MM}^* . For the parameter ranges in (5.1) and (5.2), these designs have two or three support points. The results are listed in Table 1 which also shows the worst D -efficiency defined in (2.9). For example, if the set Θ in (2.11) is given by Θ_1 and $n = 1$, then the standardized maximin D -optimal saturated design ξ_{MM} for quantile regression analysis in the Michaelis-Menten model with scale function $\sigma(x, \theta) = (\theta_2 + x)/(\theta_1 x)$ advises the experimenter to take 50% of the observations at the points 499.2 and 2000 and its worst efficiency in the set Θ_1 is given by 64.7%. The standardized maximin

D -optimal design ξ_{MM}^* in the class of all designs has masses 19.8%, 35.3% and 44.9% at the points 211.2, 846.3 and 2000 and its worst efficiency in the set Θ_1 is given by 74.4%.

The corresponding results for the Bayesian D -optimality criterion are depicted in Table 2, where $\xi_{\pi_1}^*$ denotes the Bayesian D -optimal saturated designs derived by Theorem 4.7 with respect to a uniform prior, and $\xi_{\pi_2}^*$ and $\xi_{\pi_3}^*$ denote the Bayesian D -optimal saturated designs with respect to the absolute continuous priors

$$\begin{aligned}\pi_2(d\theta_2) &= 2 \frac{\theta_2 - \theta_{2,l}}{(\theta_{2,u} - \theta_{2,l})^2} d\theta_2, & \theta_2 \in [\theta_{2,l}, \theta_{2,u}], \\ \pi_3(d\theta_2) &= -2 \frac{\theta_2 - \theta_{2,u}}{(\theta_{2,l} - \theta_{2,u})^2} d\theta_2, & \theta_2 \in [\theta_{2,l}, \theta_{2,u}].\end{aligned}$$

In all cases the saturated designs are globally Bayesian D -optimal (this was verified numerically). The necessary conditions specified by the equations (3.3) and (3.5) in Theorem 3.3 are displayed in the left panel of Figure 5.1 for the set Θ_1 and the scale function $\sigma(x, \Theta) = (\theta_1 x / (\theta_2 + x))^{-n}$ corresponding to the case $n = 1$, while the necessary condition for the set Θ_2 and $n = 5$ is shown in the right panel of of this figure.

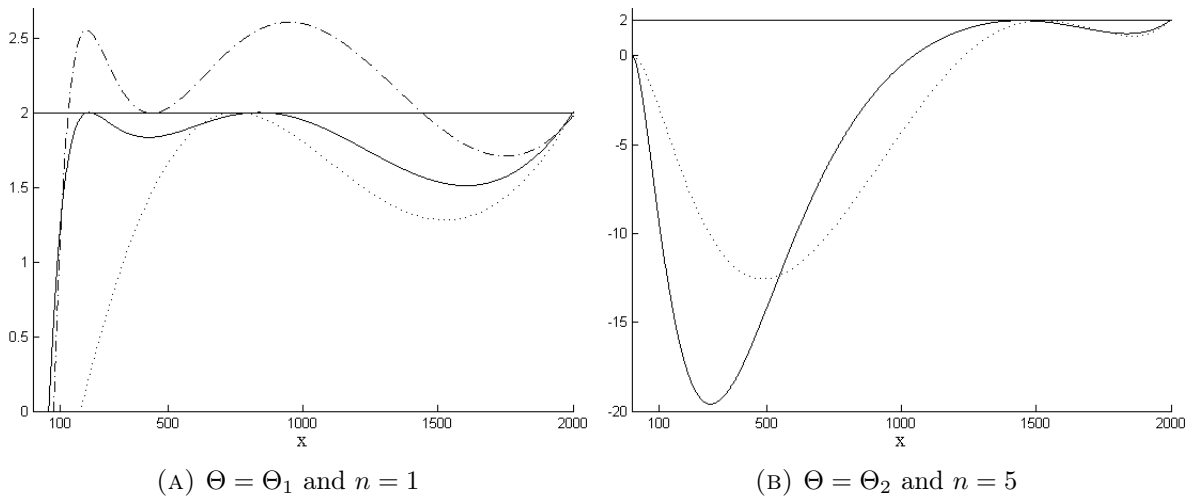


FIGURE 5.1. Plots of the function (3.3) for the design $\xi_{\pi_1}^*$ (\cdots) and of the function (3.5) for the designs ξ_{MM}^* ($—$) and ξ_{MM} ($- \cdots -$) for two different sets of parameters and link functions $h_1(z) = z^{-n}$. The designs are Bayesian or standardized maximin D -optimal if the curve stays below the line $y = 2$.

Finally, we investigate the efficiencies of the robust designs ξ_{MM}^* and $\xi_{\pi_1}^*$ as a function of the parameter θ_2 . The results are depicted in Figure 5.2, where we also show the efficiency of the uniform design ξ_U with equal masses at the 10 support points $\{200 \cdot i\}_{i=1}^{10}$ and the standardized maximin D -optimal design ξ_{MM}^0 derived under the assumption of a constant scale function (i.e. $n = 0$). We observe that the robust designs yield substantially better efficiencies than the uniform design ξ_U . Additionally, the standardized maximin D -optimal design ξ_{MM}^0 derived under the (false) assumption of homoscedasticity also yields a considerably loss of efficiency. This effect is even more visible in the right panel of Figure 5.2 for the above-mentioned two designs. These results indicate that the variation in the data has to be addressed carefully in the design of experiments for quantile regression analysis.

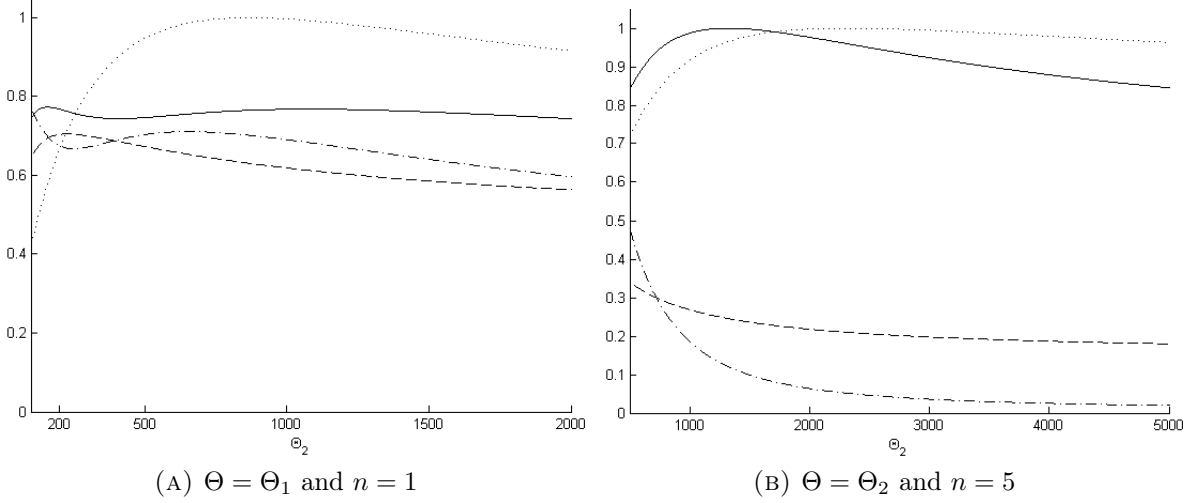


FIGURE 5.2. D -efficiencies (2.9) for the designs $\xi_{\pi_1}^*$ (\cdots), ξ_{MM}^* ($—$), ξ_{MM}^0 ($- \cdot - \cdot$) and ξ_U ($- - -$) for two different sets of parameters.

Acknowledgements. This work has been supported in part by the Collaborative Research Center “Statistical modeling of nonlinear dynamic processes” (SFB 823, Teilprojekt C2) of the German Research Foundation (DFG).

6. APPENDIX: PROOFS

6.1. Proof of the results in Section 3.

Proof of Theorem 3.1. Assume that ξ_θ^* is locally D -optimal for quantile regression analysis in model (2.1), that is ξ_θ^* maximizes the determinant defined in (2.8). For $\alpha \in [0, 1]$ and a further design ξ consider the design $\xi_\alpha = (1 - \alpha)\xi_\theta^* + \alpha\xi$. By a straightforward calculation we obtain the directional derivative of the function

$$m(\xi) = \log \det(H^{-1}(\xi, \theta)) = 2 \log \det D_1(\xi, \theta) - \log \det D_0(\xi, \theta)$$

at ξ_θ^* in the direction of $\xi - \xi_\theta^*$ as

$$\begin{aligned} \frac{d}{d\alpha} m(\xi_\alpha) \Big|_{\alpha=0} &= \frac{d}{d\alpha} (2 \log |D_1(\xi_\alpha, \theta)| - \log |D_0(\xi_\alpha, \theta)|) \Big|_{\alpha=0} \\ &= 2 \text{tr} \left(D_1^{-1}(\xi_\theta^*, \theta) [D_1(\xi, \theta) - D_1(\xi_\theta^*, \theta)] \right) - \text{tr} \left(D_0^{-1}(\xi_\theta^*, \theta) [D_0(\xi, \theta) - D_0(\xi_\theta^*, \theta)] \right) \\ (6.1) \quad &= 2 \text{tr} \left(D_1^{-1}(\xi_\theta^*, \theta) D_1(\xi, \theta) \right) - \text{tr} \left(D_0^{-1}(\xi_\theta^*, \theta) D_0(\xi, \theta) \right) - p, \end{aligned}$$

where $\text{tr}(B)$ denotes the trace of the matrix B . Because the design ξ_θ^* is D -optimal, the derivative is not positive for all designs ξ and using Dirac measures δ_x with mass 1 at the points $x \in \mathcal{X}$ for the design ξ we have for $j = 0, 1$

$$\text{tr} \left(D_j^{-1}(\xi_\theta^*, \theta) D_j(\delta_x, \theta) \right) = \text{tr} \left(D_j^{-1}(\xi_\theta^*, \theta) \frac{1}{(\sigma(x, \theta))^j} g(x, \theta) \dot{g}^T(x, \theta) \right) = \frac{1}{(\sigma(x, \theta))^j} d_j(x, \xi_\theta^*, \theta).$$

Therefore, we obtain from (6.1) the inequality $d(x, \xi_\theta^*) \leq p$ for all $x \in \mathcal{X}$. Next, assume for the locally D -optimal design ξ_θ^* that $\max_{x \in \mathcal{X}} d(x, \xi_\theta^*, \theta) < p$, which would yield

$$(6.2) \quad \int_{\mathcal{X}} d(x, \xi_\theta^*, \theta) d\xi_\theta^*(x) < p \int_{\mathcal{X}} d\xi_\theta^*(x) = p.$$

On the other hand, it follows from the definition of the functions d_0 and d_1 and a straightforward calculation that for any design ξ

$$\int_{\mathcal{X}} d_0(x, \xi, \theta) d\xi(x) = \int_{\mathcal{X}} \frac{1}{\sigma(x, \theta)} d_1(x, \xi, \theta) d\xi(x) = p,$$

which yields $\int_{\mathcal{X}} d(x, \xi_\theta^*) d\xi_\theta^*(x) = p$ and contradicts (6.2). Therefore, $\max_{x \in \mathcal{X}} d(x, \xi_\theta^*) = p$ and from $\int_{\mathcal{X}} d(x, \xi_\theta^*, \theta) d\xi_\theta^*(x) = p$ it follows that the maximum is attained at each support point of the locally D -optimal design ξ_θ^* , which concludes the proof of Theorem 3.1. \square

Proof of Lemma 3.2. For any saturated design ξ with weights w_1, \dots, w_p at support points x_1, \dots, x_p , let X denote the $p \times p$ matrix with i -th row given by $\dot{g}^T(x_i, \theta)$, $W = \text{diag}(w_1, \dots, w_p)$ and $L = \text{diag}(\sigma(x_1, \theta), \dots, \sigma(x_p, \theta))$. We obtain from (2.8) by a simple calculation that

$$(6.3) \quad \det(H^{-1}(\xi, \theta)) = (\det X)^2 (\det L)^{-2} \det W,$$

and the assertion follows by a standard calculation maximizing $\det W = w_1 \cdot \dots \cdot w_p$ with respect to the choice of the weights. \square

6.2. Proof of results on locally D -optimal designs.

An important argument in the proof of most results on locally optimal designs is a transformation from the original design space \mathcal{X} onto a new design space \mathcal{Z} . It turns out that the optimization problems corresponding to the D -optimal design problem on the new design space \mathcal{Z} have some favorable properties for both the Michaelis-Menten and the EMAX model. In particular, we obtain a very simple representation of the support points of the locally D -optimal designs on the design space \mathcal{Z} . Furthermore, the statements of the theorems can easily be obtained by transforming the designs on the design space \mathcal{Z} back onto the design space \mathcal{X} .

6.2.1. The induced design space \mathcal{Z} .

To be precise, note that the gradient \dot{g} defined in (2.3) with respect to the parameter vector θ for the Michaelis-Menten model (4.1) is given by

$$\dot{g}_1(x, \theta) = \frac{\partial}{\partial \theta} g_1(x, \theta) = \left(\frac{x}{\theta_2 + x}, -\frac{\theta_1 x}{(\theta_2 + x)^2} \right)^T.$$

We introduce the transformation

$$(6.4) \quad z = t(x) = \frac{\theta_1 x}{\theta_2 + x},$$

which gives $\sigma(t^{-1}(z), \theta) = h_j(g_1(t^{-1}(z), \theta)) = h_j(z)$ for $j = 1, 2$ and

$$\dot{g}_1(t^{-1}(z), \theta) = \left(\frac{z}{\theta_1}, \frac{-z(\theta_1 - z)}{\theta_1 \theta_2} \right)^T = \frac{z}{\theta_1} \left(1, \frac{z - \theta_1}{\theta_2} \right)^T = A(z, z^2)^T$$

on the design space $\mathcal{Z} = [z_l, z_u]$ with $z_l = \theta_1 x_l / (x_l + \theta_2)$ and $z_u = \theta_1 x_u / (x_u + \theta_2)$ and a nonsingular matrix A . For a design ξ on the design space \mathcal{X} let $\tilde{\xi}$ denote the design on $\mathcal{Z} = t(\mathcal{X}) = [z_l, z_u]$ induced by the transformation $x \rightarrow t(x) = z$, that is $d\tilde{\xi}(z) = d\xi(t^{-1}(z))$. Then it is easy to see that the matrix $H(\xi, \theta)$ defined in (2.5) can be represented as

$$(6.5) \quad H(\xi, \theta) = \tilde{H}(\tilde{\xi}) := (A^T)^{-1} \tilde{D}_1^{-1}(\tilde{\xi}) \tilde{D}_0(\tilde{\xi}) \tilde{D}_1^{-1}(\tilde{\xi}) A^{-1},$$

where the matrices \tilde{D}_0 and \tilde{D}_1 do not depend on the parameter vector θ and are defined by

$$(6.6) \quad \tilde{D}_j(\tilde{\xi}) = \int_{\mathcal{Z}} z^2 f_1(z) f_1^T(z) \frac{d\tilde{\xi}(z)}{(h(z))^j} \quad j = 0, 1,$$

the function h is given in (4.3) and $f_1^T(z) = (1, z)$. Consequently, the locally D -optimal design problem for the Michaelis-Menten model on the design space \mathcal{X} can be solved by maximizing $2 \log \det \tilde{D}_1(\tilde{\xi}) - \log \det \tilde{D}_0(\tilde{\xi})$ on the design space \mathcal{Z} and transforming the results back to the original design space \mathcal{X} . Moreover, the corresponding analogue of the function (3.1) in Theorem 3.1 is given by

$$(6.7) \quad \tilde{d}(z, \tilde{\xi}) = z^2 \left\{ \frac{2}{h(z)} f_1^T(z) \tilde{D}_1^{-1}(\tilde{\xi}) f_1(z) - f_1^T(z) \tilde{D}_0^{-1}(\tilde{\xi}) f_1(z) \right\}.$$

For the EMAX model (4.2), the gradient with respect to the parameter vector θ on the design space \mathcal{X} is given by

$$\dot{g}_2(x, \theta) = \frac{\partial}{\partial \theta} g_2(x, \theta) = \left(1, \frac{x}{\theta_2 + x}, -\frac{\theta_1 x}{(\theta_2 + x)^2} \right)^T,$$

and the situation is slightly more complicated. The same transformation defined by (6.4) yields

$$(6.8) \quad \dot{g}_2(t^{-1}(z), \theta) = \left(1, \frac{z}{\theta_1}, \frac{z(z - \theta_1)}{\theta_1 \theta_2} \right)^T = B(1, z, z^2)^T$$

for some appropriate non-singular matrix B and $g_2(t^{-1}(z), \theta) = \theta_0 + z$. By the same arguments as before D -optimal designs for the EMAX model can be found by maximizing $2 \log \det \bar{D}_1(\tilde{\xi}) - \log \det \bar{D}_0(\tilde{\xi})$, where the matrices \bar{D}_0 and \bar{D}_1 are given by

$$\bar{D}_j(\tilde{\xi}) = \int_{\mathcal{Z}} f_2(z) f_2^T(z) \frac{d\tilde{\xi}(z)}{h(\theta_0 + z)^j} \quad j = 1, 2,$$

and $f_2^T(z) = (1, z, z^2)$. The analogue of the function (3.1) is given by

$$\bar{d}(z, \tilde{\xi}) = \frac{2}{h(z + \theta_0)} f_2^T(z) \bar{D}_1^{-1}(\tilde{\xi}) f_2(z) - f_2^T(z) \bar{D}_0^{-1}(\tilde{\xi}) f_2(z).$$

Note that we do not reflect the dependence of $\bar{D}_j, j = 1, 2$ on the parameter θ_0 in the notation.

6.2.2. Proof of results on locally D -optimal designs for the Michaelis-Menten model.

Proof of Theorem 4.1. Note that the induced design space is given by $\mathcal{Z} = [z_l, z_u] = [t(x_l), t(x_r)] \subset \mathbb{R}_0^+$. By Lemma 3.2 the optimal design has equal masses at its support points. For any saturated design $\tilde{\xi}$ with equal weights at support points $z_1 < z_2$ on the design space \mathcal{Z} a straightforward calculation yields for the determinant of the inverse of the matrix \tilde{H} defined in (6.5)

$$(6.9) \quad \det \tilde{H}^{-1}(\tilde{\xi}) = c \cdot n(z_1, z_2) := c \cdot (z_1 z_2 (z_2 - z_1) h^{-1}(z_1) h^{-1}(z_2))^2,$$

with a constant c not depending on the points z_1 and z_2 . Taking the logarithmic derivative of $n(z_1, z_2)$ with respect to the upper support point z_2 yields

$$(6.10) \quad \frac{\partial}{\partial z_2} \log n(z_1, z_2) = 2 \left(\frac{1}{z_2} + \frac{1}{z_2 - z_1} - \frac{h'(z_2)}{h(z_2)} \right).$$

On the design space \mathcal{Z} the condition (4.4) shows that the inequality $\frac{z}{z} - \frac{h'(z)}{h(z)} \geq 0$ holds for all $z \in \mathcal{Z}$. Observing that for $z_2 > z_1 > 0$ we have $\frac{1}{z_2 - z_1} > \frac{1}{z_2}$ and it follows that $\frac{\partial}{\partial z_2} \log(n(z_1, z_2)) > 0$

for all $z_2 > z_1 > 0$. Therefore, the function $n(z_1, z_2)$ is increasing with respect to $z_2 > z_1$ and z_u is a support point of the best saturated design maximizing $\det \tilde{H}^{-1}(\tilde{\xi})$, which proves the first part of the assertion.

For the proof of the second part let $\tilde{\xi}$ denote a saturated design with equal weights at support points z_1, z_u on the design space \mathcal{Z} . Consider the logarithmic derivative of the function $n(z_1, z_u)$ with respect to z_1

$$(6.11) \quad l(z_1) = \frac{\partial}{\partial z_1} \log \det n(z_1, z_u) = \frac{1}{z_1} - \frac{1}{z_u - z_1} - \frac{h'(z_1)}{h(z_1)}.$$

The function $j(z) = \frac{1}{z} - \frac{1}{z_u - z}$ is continuous and strictly decreasing on $\mathcal{Z} = [z_l, z_u] \subset \mathbb{R}^+$. By assumption (4.5) the function $\frac{h'(z)}{h(z)}$ is non-decreasing and therefore the equation $l(z_1) = 0$ is solved by exactly one $z^* \in \mathbb{R}^+$. If $z^* \in \mathcal{Z} = [z_l, z_u]$, then the best saturated design is supported at z^* and z_u , otherwise the support points are given by z_l and z_u . The assertion follows transforming these results to the original design space \mathcal{X} . \square

Proof of Theorem 4.2. By the discussion in Section 6.2.1 the D -optimality of the design $\tilde{\xi}$ implies that the inequality $\tilde{d}(z, \tilde{\xi}) \leq 2$ is satisfied for all $z \in \mathcal{Z}$, where the function \tilde{d} is defined in (6.7). Moreover, there is equality for all support points of the D -optimal design $\tilde{\xi}$. It is easy to see that for the link function $h_1(z) = z^{-n}$ this inequality is of the form

$$(6.12) \quad H(z) = z^{n+2}P_2(z) + z^2Q_2(z) - 2 \leq 0,$$

where $P_2(z) = a_2z^2 + a_1z + a_0$ and $Q_2(z) = b_2z^2 + b_1z + b_0$ are nonnegative quadratic polynomials with $a_0, a_2, b_1 \geq 0$; $a_1, b_0, b_2 \leq 0$.

We first consider the case $n \geq 2$, where the derivative of the polynomial $H(z)$ is a polynomial of degree $n + 3$ whose sequence of coefficients has at most 5 sign changes. By Descartes' rule of signs [see Polya and Szegö (1971)] this means that $H'(z)$ has either exactly 5 positive roots or the number of roots is less or equal than 3 (counting multiplicities). A straightforward counting argument now shows that the function H in (6.12) has at most 3 positive roots including the boundary point z_u . Consequently, the locally D -optimal design has at most 3 support points including the point z_u . The remaining cases $n \in \{-2, -1, 0, 1\}$ have to be considered separately and we restrict ourselves to the case $n = 1$ (all other cases are treated similarly). If $n = 1$ we obtain for the derivative of the function H in (6.12) that $H'(z) = zQ_3(z)$ where Q_3 is a polynomial of degree 3. Because 0 cannot be a support point of the locally D -optimal design (note that $H(0) = -2$), the assumption of a locally D -optimal 3-point design would yield that the derivative $H'(z)$ had at least 4 positive roots contradicting to the fact that the degree of the polynomial Q_3 is 3. Therefore, the locally D -optimal design is supported at two points. The same argument shows (observing that $H(z) \rightarrow \infty$ as $z \rightarrow \infty$) that the right boundary point z_u must be a support point. \square

Proof of Theorem 4.3. Note that for the weight function h_1 in (4.7) we have $h'(z)/h(z) = -n/z$ and therefore in the case $n \geq 0$ Theorem 4.1 is directly applicable and shows that the locally D -optimal saturated design has equal masses at the points $\max\{x^*, x_l\}$ and x_u , where $x^* = t^{-1}(z^*)$ and z^* is a solution of the equation $l(z^*) = 0$ where the function l is defined in (6.11); that is ($z_u = t(x_u)$)

$$z^* = \frac{(1+n)z_u}{2+n} = \frac{(1+n)\frac{\theta_1 x_u}{x_u + \theta_2}}{2+n}.$$

Now an application of the inverse of the transformation (6.4) yields $x_1^* = t^{-1}(z_1^*) = \frac{(n+1)x_u\theta_2}{(n+2)\theta_2+x_u}$ and completes the proof of the first part of Theorem 4.3. The arguments for the remaining case $n \in [-2, 0)$ are similar and are left to the reader.

For the weight function h_2 in (4.8) we have $h'(z)/h(z) = -n$ and Theorem 4.1 is also applicable. The solution of the equation $l(z^*) = 0$ is given by

$$z_1^* = \frac{-2 + nz_2 + \sqrt{4 + n^2z_2^2}}{2n},$$

and the proof is now completed by the same arguments as in the previous paragraph. \square

6.2.3. Proof of results on locally D -optimal designs for the EMAX model.

Proof of Lemma 4.4. The proof follows by similar arguments as the proof of Theorem 4.1 and we only mention the main differences. Let $\tilde{\xi}$ denote a saturated design on the induced design space \mathcal{Z} with equal weights at the points $0 \leq z_l \leq z_1 < z_2 < z_3 \leq z_u$. It is easy to see that the determinant $\det H^{-1}(\tilde{\xi}, \theta)$ is proportional to the function

$$(6.13) \quad m(z_1, z_2, z_3) = ((z_1 - z_2)(z_1 - z_3)(z_2 - z_3)h^{-1}(z_1 + \theta_0)h^{-1}(z_2 + \theta_0)h^{-1}(z_3 + \theta_0))^2.$$

Taking the logarithmic derivative of $m(z_1, z_2, z_3)$ with respect to the largest support point z_3 yields

$$\frac{\partial}{\partial z_3} \log m(z_1, z_2, z_3) = \frac{1}{z_3 - z_1} + \frac{1}{z_3 - z_2} - \frac{h'(\theta_0 + z_3)}{h(\theta_0 + z_3)} > 0,$$

where the inequality follows from condition (4.9) (transferred to the induced design space) and a straightforward calculation. Transforming these results back to the design space \mathcal{X} concludes the proof. \square

Proof of Theorem 4.5 and Theorem 4.6. We only prove Theorem 4.5 since the proof of Theorem 4.6 follows by similar arguments. Note that for the link function h_1 in (4.7) Lemma 4.4 is applicable and shows that the locally D -optimal design has upper support point $z_3^* = z_u$. Recalling the definition of the function $m(z_1, z_2, z_3)$ in (6.13) we determine the solution of the equation $(\partial/\partial z_1, \partial/\partial z_2)^T m(z_1, z_2, z_u) = 0$ with respect to z_1 and z_2 . There exist two (symmetric) solutions and the vector (\bar{z}_-, \bar{z}_+) with

$$\bar{z}_\pm = \frac{-3(3 + 2n)\theta_0 + (3 + 4n(2 + n))z_u \pm \sqrt{3}\sqrt{(3 + 4n(2 + n))(\theta_0 + z_u)^2}}{2(2 + n)(3 + 2n)}$$

corresponds to a maximum of the function m . Thus, if $\bar{z}_- \geq z_l$, then we obtain $z_1^* = \bar{z}_-$ and $z_2^* = \bar{z}_+$ for the support points of the D -optimal design, and transforming back to the design space \mathcal{X} yields part a) and c) of the theorem.

In the opposite case note that $\kappa_{\mathcal{Z}} = \bar{z}_-$ is the analogue of the quantity κ defined in (4.10) on the induced design space \mathcal{Z} . If $\kappa_{\mathcal{Z}} = \bar{z}_- < z_l$, then $z_1^* = z_l$, and solving $\frac{\partial}{\partial z_2} m(z_l, z_2, z_u) = 0$ and transforming these results back to the design space \mathcal{X} yields part b) of the theorem. \square

6.3. Proof of results for robust designs.

For the sake of transparency we assume in this section that the design space is given by $\mathcal{X} = [0, 1]$, i.e. $x_u = 1$. The general statement can easily be obtained by a linear transformation. We also note that the efficiency defined in (2.9) does not depend on the parameter θ_1 .

Proof of Theorem 4.7. The Bayesian D -optimal design depends only on the marginal distribution π_2 of the prior π . It is also easy to check that the optimal weights of a saturated design must be

equal and that the Bayesian optimality criterion $\Phi_\pi(\xi)$ for a saturated design with equal weights is an increasing function of the larger support point. Consequently, we can restrict the optimization to saturated designs with equal masses at the points x_1 and 1, for which the optimality criterion (2.10) reduces to

$$(6.14) \quad p \cdot \Phi_\pi(\xi) = c_\Theta + \log(1 - x_1) + (1 + n) \log(x_1) - (2 + n) \int_\Theta \log(\theta_2 + x_1) \pi_2(d\theta_2),$$

where the constant c_Θ does not depend on x_1 . Since $\int_\Theta |\log \theta_2| \pi_2(d\theta_2) < \infty$, this function is continuously differentiable with respect to x_1 and the first derivative is given by

$$b(x_1) := \frac{\partial}{\partial x_1} \Phi_\pi(\xi) = -\frac{1}{1 - x_1} + \frac{1 + n}{x_1} - (2 + n) \int_\Theta \frac{\pi_2(d\theta_2)}{\theta_2 + x_1}.$$

Because $\lim_{x_1 \rightarrow 0} b(x_1) = \infty$ and $\lim_{x_1 \rightarrow 1} b(x_1) = -\infty$, there exists at least one zero in the interval $(0, 1)$. In the following we prove that the solution is unique. This implies that it corresponds to a global maximum of (6.14) in the interval $(0, 1)$, and proves the theorem. For this purpose recall the definition of \bar{x} in Theorem 4.7 and note that

$$\begin{aligned} b'(x_1) &= -\frac{1}{(1 - x_1)^2} - \frac{1 + n}{x_1^2} + (2 + n) \int_\Theta \frac{\pi_2(d\theta_2)}{(\theta_2 + x_1)^2} \\ &< \frac{-x_1^2(x_1 + \bar{x})^2 - (n + 1)(1 - x_1)^2(x_1 + \bar{x})^2 + (n + 2)(1 - x_1)^2 x_1^2}{(1 - x_1)^2 x_1^2 (x_1 + \bar{x})^2}. \end{aligned}$$

It now follows by a straightforward calculation that the definition of \bar{x} implies $b'(x_1) < 0$ for all $x_1 \in (0, 1)$, which completes the proof of Theorem 4.7 in the case $x_u = 1$. \square

Proof of Theorem 4.8. Because the efficiency in (2.9) does not depend on the parameter θ_1 the standardized maximin optimal design maximizes

$$(6.15) \quad (\Psi_\Theta(\xi))^p = \min \left\{ \frac{|M(\xi, \theta_2)|}{|M(\xi_{\theta_2}^*, \theta_2)|} \mid \theta_2 \in [\theta_{2,l}, \theta_{2,u}] \right\},$$

where

$$(6.16) \quad |M(\xi_{\theta_2}^*, \theta_2)| = \frac{1}{4\theta_2^2} \left(\frac{(1 + n)}{(2 + n)^2} \right)^{2+n} \left(\frac{1}{1 + \theta_2} \right)^{6+4n}$$

by Theorem 4.3. It follows by a standard argument that the weights of a standardized D -optimal saturated design must be equal, and for a design ξ_{x_1, x_2} with equal masses at the points $x_1, x_2 \in [0, 1]$ the determinant of the information matrix defined in (2.5) is given by

$$|M(\xi_{x_1, x_2}, \theta_2)| = \frac{\left(\frac{x_1}{\theta_2 + x_1} \right)^{2(1+n)} (x_1 - x_2)^2 \left(\frac{x_2}{\theta_2 + x_2} \right)^{2(1+n)}}{4(\theta_2 + x_1)^2 (\theta_2 + x_2)^2}.$$

Therefore, the standardized maximin D -optimality criterion reduces to

$$(\Psi_\Theta(\xi_{x_1, x_2}))^p = d_n \cdot \min \left\{ \theta_2^2 (1 + \theta_2)^{6+4n} \frac{x_1^{2(1+n)} (x_1 - x_2)^2 x_2^{2(1+n)}}{(\theta_2 + x_1)^{2(n+2)} (\theta_2 + x_2)^{2(n+2)}} \mid \theta_2 \in [\theta_{2,l}, \theta_{2,u}] \right\},$$

where the constant $d_n = (2 + n)^{4+2n} / (1 + n)^{2+2n}$ does not depend on x_1, x_2 and θ_2 and we have used the representation for the determinant in (6.16). Taking derivatives it can be shown that for fixed x_1 the function $\Psi_\Theta(\xi_{x_1, x_2})$ is strictly increasing in $x_2 \in [x_1, 1]$ and, consequently, the boundary point 1 is a support point of the best saturated design. We introduce the notation $\xi_x = \xi_{x,1}$ and

the criterion reduces further to

$$(6.17) \quad (\Psi_{\Theta}(\xi_x))^p = d_n \cdot \min \left\{ u(x, \theta_2) \mid \theta_2 \in [\theta_{2,l}, \theta_{2,u}] \right\},$$

where the function u is defined by

$$u(x, \theta_2) = \frac{\theta_2^2(1 + \theta_2)^{2(1+n)}x^{2(1+n)}(x - 1)^2}{(\theta_2 + x)^{2(2+n)}}.$$

Recalling the definition of the point x^* in (4.15) for $x_u = 1$, it can be shown by similar arguments as in Dette and Biedermann (2003) that

- (1) for any fixed $x \in (0, 1)$ and $n > -1$, the function $\theta_2 \rightarrow \log u(x, \theta_2)$ is unimodal, and
- (2) if ξ_{x^*} maximizes the function Ψ_{Θ} in the class of all saturated designs, then $u(x^*, \theta_{2,l}) = u(x^*, \theta_{2,u})$.

Solving the last equation with respect to x^* yields by a straightforward calculation that

$$x^* = x^*(\theta_{2,l}, \theta_{2,u}) = \frac{\theta_{2,u}(\theta_{2,l}(1 + \theta_{2,l})^{1+n})^{\frac{1}{2+n}} - \theta_{2,l}(\theta_{2,u}(1 + \theta_{2,u})^{1+n})^{\frac{1}{2+n}}}{(\theta_{2,u}(1 + \theta_{2,u})^{1+n})^{\frac{1}{2+n}} - (\theta_{2,l}(1 + \theta_{2,l})^{1+n})^{\frac{1}{2+n}}},$$

which proves the assertion of Theorem 4.8 in the case $x_u = 1$. □

REFERENCES

- Atkinson, A. C. (1996). The usefulness of optimum experimental designs. with discussion and a reply by the author. *Journal of the Royal Statistical Society, Ser. B*, 58(1):59–76, 95–111.
- Boroujerdi, M. (2002). *Pharmacokinetics: Principles and Applications*. New York: Mc-Graw Hill.
- Buchinsky, M. (1994). Changes in the united-states wage structure 1963-1987 - application of quantile regression. *Econometrica*, 62(2):405–458.
- Cade, B., Terrell, J., and Schroeder, R. (1999). Estimating effects of limiting factors with regression quantiles. *Ecology*, 80(1):311–323.
- Chaloner, K. and Verdinelli, I. (1995). Bayesian experimental design: a review. *Statistical Science*, 10(3):273–304.
- Chernoff, H. (1953). Locally optimal designs for estimating parameters. *Annals of Mathematical Statistics*, 24:586–602.
- Cornish-Browden, A. (1995). *Fundamentals of Enzyme Kinetics (Rev. Ed.)*. Portland Press, London.
- Cressie, N. and Keightley, D. (1979). The underlying structure of a direct linear plot with applications to the analysis of hormone-receptor interactions. *Journal of Steroid Biochemistry*, 11:1173–1180.
- Demana, P., Smith, E., Walker, R., Haigh, J., and Kanfer, I. (1997). Evaluation of the proposed fda pilot dose-response methodology for topical corticosteroid bioequivalence testing. *Pharmaceutical Research*, 14:303–308.
- Dette, H. (1997). Designing experiments with respect to “standardized” optimality criteria. *Journal of the Royal Statistical Society, Ser. B*, 59:97–110.
- Dette, H. and Biedermann, S. (2003). Robust and efficient designs for the Michaelis-Menten model. *Journal of the American Statistical Association*, 98:679–686.
- Dette, H., Bretz, F., Pepelyshev, A., and Pinheiro, J. (2008). Optimal designs for dose-finding studies. *Journal of the American Statistical Association*, 103(483):1225–1237.

- Dette, H., Kiss, C., Bevanda, M., and Bretz, F. (2010). Optimal designs for the EMAX, log-linear and exponential models. *Biometrika*, 97(2):513–518.
- Dette, H. and Wong, W. (1999). E-optimal designs for the Michaelis-Menten model. *Statistics and Probability Letters*, 44:405–408.
- Dunn, G. (1988). Optimal designs for drug, neurotransmitter and hormone receptor assays. *Statistics in Medicine*, 7:805–815.
- Fang, X. and Hedayat, A. (2008). Locally d-optimal designs based on a class of composed models resulted from blending emax and one-compartment models. *Annals Of Statistics*, 36:428–444.
- Ford, I., Torsney, B., and Wu, C. F. J. (1992). The use of canonical form in the construction of locally optimum designs for nonlinear problems. *Journal of the Royal Statistical Society, Ser. B*, 54:569–583.
- Haines, L. M. (1995). A geometric approach to optimal design for one-parameter non-linear models. *Journal of the Royal Statistical Society, Series B*, 57(3):575–598.
- He, Z., Studden, W. J., and Sun, D. (1996). Optimal designs for rational models. *Annals of Statistics*, 24:2128–2147.
- Imhof, L. A. (2001). Maximin designs for exponential growth models and heteroscedastic polynomial models. *Annals of Statistics*, 29:561–576.
- Johansen, S. (1984). Functional relations, random coefficients and nonlinear regression, with application to kinetic data. *Lecture Notes in Statistics*, 22.
- Kiefer, J. (1974). General equivalence theory for optimum design (approximate theory). *The Annals of Statistics*, 2:849–879.
- Koenker, R. (2005). *Quantile Regression*. Cambridge University Press, New York.
- Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica*, 46(1):33–50.
- Koenker, R. and Zhao, Q. S. (1994). l -estimation for linear heteroscedastic models. *Journal of Nonparametric Statistics*, 3:223–235.
- Lopez-Fidalgo, J. and Wong, W. (2002). Optimal designs for the Michaelis-Menten model. *Journal of Theoretical Biology*, 215:1–11.
- Merle, Y. and Mentre, F. (1995). Bayesian design criteria - computation, comparison, and application to a pharmacokinetic and a pharmacodynamic model. *Journal of Pharmacokinetics And Biopharmaceutics*, 23:101–125.
- Polya, G. and Szegő, G. (1971). *Aufgaben und Lehrsätze aus der Analysis, Bd. I, II*. Springer, Berlin.
- Pukelsheim, F. (2006). *Optimal Design of Experiments*. SIAM, Philadelphia.
- Rasch, D. (1990). Optimum experimental design in nonlinear regression. *Communications in Statistics*, 19:4789–4806.
- Silvey, S. D. (1980). *Optimal Design*. Chapman & Hall, London.
- Wang, J. (2006). Optimal parametric design with applications to pharmacokinetic and pharmacodynamic trials. *Journal of Applied Statistics*, 33, 8:837–852.
- Wei, Y., Pere, A., Koenker, R., and He, X. (2006). Quantile regression methods for reference growth charts. *Statistics In Medicine*, 25(8):1369–1382.
- Zhou, K. and Portnoy, S. (1998). Statistical inference on heteroscedastic models based on regression quantiles. *Journal of Nonparametric Statistics*, 9:239–260.

