APPLIED STOCHASTIC MODELS IN BUSINESS AND INDUSTRY
Appl. Stochastic Models Bus. Ind. 2009; **25**:179–193
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/asmb.739

Robust designs for misspecified exponential regression models

Xiaojian Xu*,†

Department of Mathematics, Brock University, 500 Glenridge Avenue, St. Catharines, Ont., Canada L2S 3A1

SUMMARY

We consider the construction of designs for exponential regression. The response function is an only approximately known function of a specified exponential function. As well, we allow for variance heterogeneity. We find minimax designs and corresponding optimal regression weights in the context of the following problems: (1) for nonlinear least-squares (LS) estimation with homoscedasticity, determine a design to minimize the maximum value of the integrated mean-squared error (IMSE), with the maximum being evaluated for the possible departures from the response function; (2) for nonlinear LS estimation with heteroscedasticity, determine a design to minimize the maximum value of IMSE, with the maximum being evaluated over both types of departures; (3) for nonlinear weighted LS estimation, determine both weights and a design to minimize the maximum IMSE; and (4) choose weights and design points to minimize the maximum IMSE, subject to a side condition of unbiasedness. Solutions to (1)–(4) are given in complete generality. Copyright © 2009 John Wiley & Sons, Ltd.

Received 19 June 2008; Revised 30 July 2008; Accepted 12 September 2008

KEY WORDS: nonlinear least squares; regression design; Q-optimal; unbiased design

1. INTRODUCTION

Exponential regression models are widely used in many areas, such as pharmacokinetics [1], agricultural sciences [2], and life testing [3]. Melas [4] investigates the optimal design problem for exponential regression under D-optimality. Recently, the D-efficient (with respect to the local D-optimal design) designs for such model with a single covariate have been constructed by Dette *et al.* [5]. Both studies use the minimax approach to address the problem of the resulting designs' dependency on the parameters being estimated. While minimax designs prevents the worst case scenario within the parameter space, the Bayesian approach considers an average over the

*Correspondence to: Xiaojian Xu, Department of Mathematics, Brock University, 500 Glenridge Avenue, St. Catharines, Ont., Canada L2S 3A1.
†E-mail: xxu@brocku.ca

Contract/grant sponsor: Natural Sciences and Engineering Research Council of Canada; contract/grant number: 341783-07

Copyright © 2009 John Wiley & Sons, Ltd.

parameter space. Fang and Wiens [6] discuss the Bayesian minimally supported D-optimal designs for exponential regression models.

All these studies above assume that the response is described exactly by a particular exponential regression model. In practice, the assumed model is likely to be only a reasonable approximation to the true model for the response. This true model is generally not known. As indicated in [7], it is risk to design a regression experiment, which assumes that the regression model is exactly correct. By analyzing the relative importance of errors due to bias, and to variance, they find that very small deviations from an assumed model can eliminate any supposed gain arising from the use of a design, which minimizes variance alone. There is abundant literature on robust designs for an approximately known *linear* regression model, to name a few, [8–12]. For nonlinear regression, robust designs for a generalized version of linear regression model, which is an approximately known monotonic function of a linear function, in both parameters and regressors, are discussed in [13, 14].

In this paper, we study the construction of designs for regression responses to be an approximately known exponential function in a complete general setting. The regression function considered in the present paper does not necessarily approximate a known function of linear function in both parameters and regressors, as required in [13, 14]. When one is interested in the best estimation of the response function, the designs' problems are to find the optimal prediction designs. Therefore, Q-optimality is considered in this paper. Our designs are robust, in that we allow both for imprecision in the specification of the response, and for possible heteroscedasticity. We also adopt the minimax approach to address the worst possible case among all possible model departures from our assumptions.

(1) The experimenter takes n uncorrelated observations $Y_i = Y(\mathbf{x}_i)$, with $\mathbf{x}_i$ freely chosen from a design space S. Our goal is to choose these design points from S in an optimal manner in order to estimate $E(Y|\mathbf{x})$ in S.
(2) We consider the exponential regression model:

$$E(Y|\mathbf{x}) = \alpha + \beta \exp(\boldsymbol{\lambda}^{\mathrm{T}} \mathbf{z}(\mathbf{x})) + n^{-1/2} f(\mathbf{x}) \tag{1}$$

for p regressors $\mathbf{z}(\mathbf{x}) = (z_1(\mathbf{x}), z_2(\mathbf{x}), \ldots, z_p(\mathbf{x}))^{\mathrm{T}}$, depending on a q-dimensional vector $\mathbf{x}$ of independent variables. We assume that $\|\mathbf{z}(\mathbf{x})\|$ is bounded on S. The response contaminant f represents uncertainty about the exact nature of the regression response and is unknown and arbitrary, subject to certain restrictions. We estimate $\boldsymbol{\theta} = (\alpha, \beta, \boldsymbol{\lambda}^{\mathrm{T}})^{\mathrm{T}}$ but not f; this leads to possibly biased estimation $\hat{Y}(\mathbf{x}) = \hat{\alpha} + \hat{\beta} \exp(\hat{\boldsymbol{\lambda}}^{\mathrm{T}} \mathbf{z}(\mathbf{x}))$ of $E(Y|\mathbf{x})$. The factor $n^{-1/2}$ is necessary for a sensible asymptotic treatment (see [13]). We note that due to the existence of α and β, the model (1) is no longer an approximation of a known function of linear function in parameters.
(3) The observations Y_i are possibly heteroscedastic, with $\mathrm{Var}\{Y(\mathbf{x}_i)\} = \sigma^2 g(\mathbf{x}_i)$ for a function g satisfying conditions given below.

We estimate $\boldsymbol{\theta}$ by nonlinear least squares (LS), possibly weighted with weights $w(\mathbf{x})$. Our loss function is n times the integrated mean-squared error (IMSE) of $\hat{Y}(\mathbf{x})$ in estimating $E(Y|\mathbf{x})$. This depends on the design measure $\xi = n^{-1} \Sigma_{i=1}^{n} \delta_{\mathbf{x}_i}$ as well as on w, f and g:

$$\mathrm{IMSE}(f, g, w, \xi) = n \int_S E\{[\hat{Y}(\mathbf{x}) - E(Y|\mathbf{x})]^2\} \,\mathrm{d}\mathbf{x}$$

Copyright © 2009 John Wiley & Sons, Ltd.

DOI: 10.1002/asmb

We denote unweighted LS by $w=\mathbf{1}$ and homogeneous variances by $g=\mathbf{1}$. The following problems are addressed:

(P1) For ordinary LS (OLS) estimation under homoscedasticity, determine designs to minimize the maximum value, over f, of $\text{IMSE}(f, \mathbf{1}, \mathbf{1}, \xi)$.
(P2) For OLS estimation under possible heteroscedasticity, determine designs to minimize the maximum value, over f and g, of $\text{IMSE}(f, g, \mathbf{1}, \xi)$.
(P3) For weighted LS (WLS) estimation, determine designs and weights to minimize the maximum value, over f and g, of $\text{IMSE}(f, g, w, \xi)$.
(P4) Choose weights and design points to minimize $\max_{f,g} \text{IMSE}(f, g, w, \xi)$, subject to a side condition of unbiasedness.

The remainder of this paper is organized as follows. Some mathematical preliminaries are detailed in Section 2. The maximization part of the minimax designs construction is provided in Section 3. The designs for P1 are provided in Section 4. The designs and weights that constitute solutions to P2 and P3 are given in Section 5 and those for P4 are given in Section 6. Computation and discretization for these designs are presented in Section 7.

2. PRELIMINARIES AND NOTATION

We define the 'target' parameter $\boldsymbol{\theta}_0=(\alpha_0, \beta_0, \boldsymbol{\lambda}_0^{\mathrm{T}})^{\mathrm{T}}$ to be the value, which gives the best agreement, in the L_2-sense, between $\alpha+\beta\exp(\boldsymbol{\lambda}^{\mathrm{T}}\mathbf{z}(\mathbf{x}))$ and $E(Y|\mathbf{x})$:

$$\boldsymbol{\theta}_0=\arg\min_{\boldsymbol{\theta}}\left\{\int_S [\alpha+\beta\exp(\boldsymbol{\lambda}^{\mathrm{T}}\mathbf{z}(\mathbf{x}))-E(Y|\mathbf{x})]^2\,\mathrm{d}\mathbf{x}\right\}$$

According to (1),

$$f_n(\mathbf{x})=\sqrt{n}[E(Y|\mathbf{x})-\alpha-\beta\exp(\boldsymbol{\lambda}^{\mathrm{T}}\mathbf{z}(\mathbf{x}))]$$

and we denote

$$\tilde{\mathbf{z}}^{\mathrm{T}}(\mathbf{x})=-\beta\mathbf{z}^{\mathrm{T}}(\mathbf{x})\exp(\boldsymbol{\lambda}^{\mathrm{T}}\mathbf{z}(\mathbf{x})),\quad \mathbf{t}^{\mathrm{T}}(\mathbf{x})=(1, \exp(\boldsymbol{\lambda}^{\mathrm{T}}\mathbf{z}(\mathbf{x})), \tilde{\mathbf{z}}^{\mathrm{T}}(\mathbf{x}))$$

Then, we have $\int_S \mathbf{t}(\mathbf{x}) f_n(\mathbf{x})\,\mathrm{d}\mathbf{x}=\mathbf{0}$. We drop the subscript on f whenever it is possible.

We shall assume that $f_n=f$ is an unknown member of the class

$$\mathscr{F}=\left\{f\,\middle|\,\int_S f^2(\mathbf{x})\,\mathrm{d}\mathbf{x}\leqslant\eta_S^2<\infty, \int_S \mathbf{t}(\mathbf{x}) f(\mathbf{x})\,\mathrm{d}\mathbf{x}=\mathbf{0}\right\} \tag{2}$$

where η_S is a positive constant. The second condition, which comprises $p+2$ integrals, is required in order that $\boldsymbol{\theta}_0$ can be uniquely defined, and in fact can be derived through the definition.

The departure from homogeneity of variances is measured by $g(\mathbf{x})$, which is assumed to be an unknown member of the class

$$\mathscr{G}=\left\{g\,\middle|\,\int_S g^2(\mathbf{x})\,\mathrm{d}\mathbf{x}\leqslant\int_S \mathrm{d}\mathbf{x}<\infty\right\} \tag{3}$$

Letting $\Omega := (\int_S \mathrm{d}\mathbf{x})^{-1}$, the condition in (3) is equivalent to defining $\sigma^2=\sup_g\{\int_S \mathrm{Var}^2[\varepsilon(\mathbf{x})]\Omega\,\mathrm{d}\mathbf{x}\}^{1/2}$.

Copyright © 2009 John Wiley & Sons, Ltd.

DOI: 10.1002/asmb

To ensure the nonsingularity of a number of relevant matrices, we assume that the regressors and design space satisfy

(Δ) For each $\mathbf{a}\neq\mathbf{0}$, the set $\{\mathbf{x}\in S:\mathbf{a}^{\mathrm{T}}\mathbf{t}(\mathbf{x})=0\}$ has the Lebesgue measure zero.

We propose to estimate $\boldsymbol{\theta}_0$ using LS to fit $E(\hat{Y}|\mathbf{x})=\alpha_0-\beta_0\exp(\boldsymbol{\lambda}_0^{\mathrm{T}}\mathbf{z}(\mathbf{x}))$ with nonnegative weights $w(\mathbf{x})$.

We make use of the following matrices and vectors:

$$\mathbf{A}_S=\int_S \mathbf{t}(\mathbf{x})\mathbf{t}^{\mathrm{T}}(\mathbf{x})\,\mathrm{d}\mathbf{x},\quad \mathbf{B}=\int_S \mathbf{t}(\mathbf{x})\mathbf{t}^{\mathrm{T}}(\mathbf{x})w(\mathbf{x})\xi(\mathrm{d}\mathbf{x})$$

$$\mathbf{b}_{f,S}=\int_S \mathbf{t}(\mathbf{x})f(\mathbf{x})w(\mathbf{x})\xi(\mathrm{d}\mathbf{x}),\quad \mathbf{D}=\int_S \mathbf{t}(\mathbf{x})\mathbf{t}^{\mathrm{T}}(\mathbf{x})w^2(\mathbf{x})g(\mathbf{x})\xi(\mathrm{d}\mathbf{x})$$

It follows from (Δ) that $\mathbf{A}_S$ is nonsingular, and that $\mathbf{B}$ is nonsingular as well when ξ is absolutely continuous. The LS estimator of $\boldsymbol{\theta}_0$ is

$$\hat{\boldsymbol{\theta}}=\arg\min\sum_{i=1}^{n}[Y_i-\alpha-\beta\exp(\boldsymbol{\lambda}^{\mathrm{T}}\mathbf{z}(\mathbf{x}_i))]^2 w(\mathbf{x}_i)$$

and satisfies $\sum_{i=1}^{n}\dot{\boldsymbol{\phi}}_i(\hat{\boldsymbol{\theta}})=0$ for

$$\dot{\boldsymbol{\phi}}_i(\boldsymbol{\theta})=[Y_i-\alpha-\beta\exp(\boldsymbol{\lambda}^{\mathrm{T}}\mathbf{z}(\mathbf{x}_i))]w(\mathbf{x}_i)\mathbf{t}(\mathbf{x}_i)$$

In addition, the Hessian $\ddot{\Phi}(\boldsymbol{\theta})$ is given by

$$\sum_{i=1}^{n}\ddot{\boldsymbol{\phi}}_i(\boldsymbol{\theta})=\sum_{i=1}^{n}[Y_i-\alpha-\beta\exp(\boldsymbol{\lambda}^{\mathrm{T}}\mathbf{z}(\mathbf{x}_i))]w(\mathbf{x}_i)\dot{\mathbf{t}}(\mathbf{x}_i)-\sum_{i=1}^{n}w(\mathbf{x}_i)\mathbf{t}(\mathbf{x}_i)\mathbf{t}^{\mathrm{T}}(\mathbf{x}_i)$$

The information matrix is

$$\mathscr{I}(\boldsymbol{\theta}_0)=\lim_{n\to\infty}E\left(-\frac{1}{n}\ddot{\Phi}(\boldsymbol{\theta}_0)\right)=\mathbf{B}$$

since

$$E\left\{\frac{1}{n}\sum_{i=1}^{n}[Y_i-\alpha-\beta\exp(\boldsymbol{\lambda}^{\mathrm{T}}\mathbf{z}(\mathbf{x}_i))]w(\mathbf{x}_i)\dot{\mathbf{t}}(\mathbf{x}_i)\right\}=n^{-1/2}\cdot\frac{1}{n}\sum_{i=1}^{n}f(\mathbf{x}_i)w(\mathbf{x}_i)\dot{\mathbf{t}}(\mathbf{x}_i)$$

is $O(n^{-1/2})$ by virtue of our assumptions on f, and $\mathbf{z}$.

By Taylor's Theorem,

$$\mathbf{0}=\sum_{i=1}^{n}\dot{\boldsymbol{\phi}}_i(\hat{\boldsymbol{\theta}})=\sum_{i=1}^{n}\{\dot{\boldsymbol{\phi}}_i(\boldsymbol{\theta}_0)+\ddot{\boldsymbol{\phi}}_i(\tilde{\boldsymbol{\theta}})(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0)\}$$

where $\tilde{\boldsymbol{\theta}}$ lies between $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\theta}_0$. Then,

$$\sqrt{n}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0)=\left(-\frac{1}{n}\sum_{i=1}^{n}\ddot{\boldsymbol{\phi}}_i(\tilde{\boldsymbol{\theta}})\right)^{-1}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\dot{\boldsymbol{\phi}}_i(\boldsymbol{\theta}_0)\right)$$

Copyright © 2009 John Wiley & Sons, Ltd.

DOI: 10.1002/asmb

Note that $n^{-1/2}\sum_{i=1}^{n}\dot{\boldsymbol{\phi}}_i(\boldsymbol{\theta}_0)$ is asymptotically normal, with asymptotic mean $\mathbf{b}_{f,S}$ and covariance

$$\mathrm{Cov}\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\dot{\boldsymbol{\phi}}_i(\boldsymbol{\theta}_0)\right]=\frac{1}{n}\sum_{i=1}^{n}\sigma^2 g(\mathbf{x}_i)\mathbf{t}(\mathbf{x}_i,\boldsymbol{\theta}_0)\mathbf{t}^{\mathrm{T}}(\mathbf{x}_i,\boldsymbol{\theta}_0)w^2(\mathbf{x}_i)=\sigma^2\mathbf{D}$$

As at [15] (Section 12.2), it follows that the asymptotic distribution of $\sqrt{n}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0)$ is

$$\sqrt{n}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0)\sim AN(\mathbf{B}^{-1}\mathbf{b}_{f,S},\sigma^2\mathbf{B}^{-1}\mathbf{D}\mathbf{B}^{-1})$$

Let $h(\boldsymbol{\theta})=\alpha-\beta\exp(\boldsymbol{\lambda}^{\mathrm{T}}\mathbf{z}(\mathbf{x}))$, and then by the delta method,

$$\begin{aligned}&\sqrt{n}([\hat{\alpha}-\hat{\beta}\exp(\hat{\boldsymbol{\lambda}}^{\mathrm{T}}\mathbf{z}(\mathbf{x}))]-[\alpha_0-\beta_0\exp(\boldsymbol{\lambda}_0^{\mathrm{T}}\mathbf{z}(\mathbf{x}))])\\&\quad=\sqrt{n}(h(\hat{\boldsymbol{\theta}})-h(\boldsymbol{\theta}_0))\sim AN(\mathbf{t}^{\mathrm{T}}(\mathbf{x})\mathbf{B}^{-1}\mathbf{b}_{f,S},\sigma^2\mathbf{t}^{\mathrm{T}}(\mathbf{x})\mathbf{B}^{-1}\mathbf{D}\mathbf{B}^{-1}\mathbf{t}(\mathbf{x}))\end{aligned}$$

Under Q-optimality, the loss function IMSE splits into terms due to (squared) bias and variance:

$$\begin{aligned}\mathrm{IMSE}(f,g,w,\xi)&=n\int_S\{[\hat{Y}(\mathbf{x})-E(Y|\mathbf{x})]^2\}\,\mathrm{d}\mathbf{x}\\&=n\int_S E\left\{\left((h(\hat{\boldsymbol{\theta}})-h(\boldsymbol{\theta}_0))-\frac{1}{\sqrt{n}}f(\mathbf{x})\right)^2\right\}\\&=\mathrm{IB}(f,w,\xi)+\mathrm{IV}(g,w,\xi)+\int_S f^2(\mathbf{x})\,\mathrm{d}\mathbf{x}\end{aligned}$$

where the integrated-squared bias (IB) and the integrated variance (IV) are

$$\mathrm{IB}(f,w,\xi)=\int_S\{\sqrt{n}E[h(\hat{\boldsymbol{\theta}}^{\mathrm{T}})-h(\boldsymbol{\theta}_0^{\mathrm{T}})]\}^2\,\mathrm{d}\mathbf{x}$$

and

$$\mathrm{IV}(g,w,\xi)=n\int_S\mathrm{Var}(\hat{Y}(\mathbf{x}))\,\mathrm{d}\mathbf{x}=n\int_S\mathrm{Var}(h(\hat{\boldsymbol{\theta}}^{\mathrm{T}}))\,\mathrm{d}\mathbf{x}$$

Asymptotically,

$$\begin{aligned}\mathrm{IB}(f,w,\xi)&=\mathbf{b}_{f,S}^{\mathrm{T}}\mathbf{B}^{-1}\mathbf{A}_S\mathbf{B}^{-1}\mathbf{b}_{f,S}\\\mathrm{IV}(g,w,\xi)&=\sigma^2\,\mathrm{tr}(\mathbf{A}_S\mathbf{B}^{-1}\mathbf{D}\mathbf{B}^{-1})\end{aligned}$$

We have defined ξ to be a discrete measure, $\xi(\mathbf{x}_i)=n^{-1}$ at the design points $\mathbf{x}_i$ (possibly repeated). We now adopt the viewpoint of approximate design theory and allow ξ to be any probability measure on S. As the class $\mathscr{F}$ is so broad, only absolutely continuous measures ξ can have finite maximum loss, see Lemma 1 in [10] and the italic statement in [10] (p. 335) for the rationale. The requirement of absolute continuity excludes exact, implementable designs, and thus, approximations are necessary, see [16, 17] for various methods for implementing designs

Copyright © 2009 John Wiley & Sons, Ltd.

DOI: 10.1002/asmb

with continuous measures. For example, a practical implementation for univariate x is to place the n design points at the quantiles $x_i = \xi^{-1}((i-1)/(n-1))$, and it is also used in Section 7 of this paper.

We note that one might then, in order to avoid the exclusion for discrete designs and without loss of generality, require f to be bounded or continuous. According to the exploration in [14], the discretized version of such optimal designs, among those being absolutely continuous, performs well in terms of a small variance with a small bias when f is uniformly bounded. Therefore, in this paper we still continue to focus on the design with continuous measure ξ, which can be implemented by its discrete version, see Section 7 for the example.

Let $k(\mathbf{x})$ be the density of ξ and define $m(\mathbf{x}) = k(\mathbf{x})w(\mathbf{x})$. Without loss of generality, we assume that the mean weight is $\int_S w(\mathbf{x})\xi(\mathrm{d}\mathbf{x}) = 1$. Then, $m(\mathbf{x})$ is also a density on S that satisfies

$$\int_S \frac{m(\mathbf{x})}{w(\mathbf{x})}\,\mathrm{d}\mathbf{x} = 1 \tag{4}$$

and we have

$$\mathbf{B} = \int_S \mathbf{t}(\mathbf{x})\mathbf{t}^{\mathrm{T}}(\mathbf{x})m(\mathbf{x})\,\mathrm{d}\mathbf{x}$$

$$\mathbf{b}_{f,S} = \int_S \mathbf{t}(\mathbf{x})f(\mathbf{x})m(\mathbf{x})\,\mathrm{d}\mathbf{x}$$

$$\mathbf{D} = \int_S \mathbf{t}(\mathbf{x})\mathbf{t}^{\mathrm{T}}(\mathbf{x})m(\mathbf{x})w(\mathbf{x})\,\mathrm{d}\mathbf{x}$$

From the definitions of $\mathbf{B}, \mathbf{b}_{f,S}$, we note that $\mathrm{IB}(f, w, \xi)$ depends on (w, ξ) only through m and $\mathrm{IV}(g, w, \xi)$ through m and w. Hence, we can optimize over m and w subject to (4) rather than over k and w. In the following four sections, we exhibit solutions to P1–P4.

3. MAXIMIZATION OVER $f \in \mathscr{F}$ AND $g \in \mathscr{G}$

In this section we derive the maxima of IMSE for fixed functions $m(\mathbf{x})$ and $w(\mathbf{x})$. The minimizing m and w then constitute the solutions to P1–P4. The maximum in Theorem 1 is obtained in a manner similar to that used in Theorem 1 of [10], and hence its derivation is omitted.

Define positive semidefinite matrices

$$\mathbf{K} = \int_S \mathbf{t}(\mathbf{x})\mathbf{t}^{\mathrm{T}}(\mathbf{x})m^2(\mathbf{x})\,\mathrm{d}\mathbf{x}$$

$$\mathbf{H}_S = \mathbf{B}^{-1}\mathbf{A}_S\mathbf{B}^{-1}, \quad \mathbf{G} = \mathbf{K} - \mathbf{B}\mathbf{A}_S^{-1}\mathbf{B}$$

and $v = \sigma^2/\eta_S^2$, representing the relative importance of variance versus bias. We note that

$$\mathbf{G} = \int_S [(m(\mathbf{x})\mathbf{I} - \mathbf{B}\mathbf{A}_S^{-1})\mathbf{t}(\mathbf{x})][(m(\mathbf{x})\mathbf{I} - \mathbf{B}\mathbf{A}_S^{-1})\mathbf{t}(\mathbf{x})]^{\mathrm{T}}\,\mathrm{d}\mathbf{x} \tag{5}$$

hence, $\mathbf{G}$ is positive semidefinite. Let $\lambda_{\max}(\cdot)$ be the largest eigenvalue of a matrix. In this notation, we have the following theorem.

Copyright © 2009 John Wiley & Sons, Ltd.

Theorem 1
The maximum squared bias is

$$\sup_{f \in \mathscr{F}} \mathrm{IB}(f, m) = \eta_S^2 \lambda_{\max}(\mathbf{G}\mathbf{H}_S)$$

The maximum is attained at

$$f_m(\mathbf{x}) = \eta_S \mathbf{t}^{\mathrm{T}}(\mathbf{x})\{m(\mathbf{x})\mathbf{I} - \mathbf{A}_S^{-1}\mathbf{B}\}\mathbf{G}^{-1/2}\mathbf{a}_0$$

where $\mathbf{a}_0$ the eigenvector corresponding to $\lambda_{\max}(\mathbf{G}\mathbf{H}_S)$ and satisfies $\mathbf{a}_0^{\mathrm{T}}\mathbf{a}_0 = 1$.

From this result, we obtain Theorem 2 immediately and obtain Theorem 3 by applying the Cauchy–Schwarz inequality to the IV part of $\mathrm{IMSE}(f, g, w, m)$. Theorem 2 gives the maximum IMSE under homoscedasticity while Theorem 3 gives this quantity under heteroscedasticity.

Theorem 2
The maximum mean-squared error in problem P1 is

$$\sup_{f \in \mathscr{F}} \mathrm{IMSE}(f, \mathbf{1}, \mathbf{1}, m) = \eta_S^2\{1 + \lambda_{\max}(\mathbf{G}\mathbf{H}_S) + v\,\mathrm{tr}(\mathbf{A}_S\mathbf{B}^{-1})\} \tag{6}$$

attained at f_m.

Theorem 3
Define $l_m(\mathbf{x}) = [\mathbf{t}^{\mathrm{T}}(\mathbf{x})\mathbf{H}_S\mathbf{t}(\mathbf{x})]$. Then:

(i) the maximum mean-squared error in problem P2 is

$$\sup_{f \in \mathscr{F}, g \in \mathscr{G}} \mathrm{IMSE}(f, g, \mathbf{1}, m) = \eta_S^2 \left\{ 1 + \lambda_{\max}(\mathbf{G}\mathbf{H}_S) + v\Omega^{-1/2} \left[\int_S \{l_m(\mathbf{x})m(\mathbf{x})\}^2 \,\mathrm{d}\mathbf{x} \right]^{1/2} \right\}$$

attained at f_m and

$$g_m(\mathbf{x}) \propto l_m(\mathbf{x})m(\mathbf{x})$$

(ii) the maximum mean-squared error in problems P3–P4 is

$$\sup_{f \in \mathscr{F}, g \in \mathscr{G}} \mathrm{IMSE}(f, g, w, m) = \eta_S^2 \left\{ 1 + \lambda_{\max}(\mathbf{G}\mathbf{H}_S) + v\Omega^{-1/2} \left[\int_S \{w(\mathbf{x})l_m(\mathbf{x})m(\mathbf{x})\}^2 \,\mathrm{d}\mathbf{x} \right]^{1/2} \right\} \tag{7}$$

attained at f_m and

$$g_{m,w}(\mathbf{x}) \propto w(\mathbf{x})l_m(\mathbf{x})m(\mathbf{x})$$

The following theorem can be derived from Theorem 3(ii) by applying the Cauchy–Schwarz inequality to the last term in (7), it gives the minimax weights for a fixed $m(\mathbf{x})$.

Copyright © 2009 John Wiley & Sons, Ltd.

DOI: 10.1002/asmb

Theorem 4
Define $\alpha_m = \int_S [l_m(\mathbf{x}) m^2(\mathbf{x})]^{2/3}\,d\mathbf{x}$. For fixed $m(\mathbf{x})$ the weights minimizing $\sup_{f\in\mathscr{F}, g\in\mathscr{G}} \mathrm{IMSE}(f, g, w, m)$ subject to (4) are given by

$$w_m(\mathbf{x}) = \alpha_m [l_m^2(\mathbf{x}) m(\mathbf{x})]^{-1/3} I[m(\mathbf{x})>0]$$

Then $\min_w \{\sup_{f\in\mathscr{F}, g\in\mathscr{G}} \mathrm{IMSE}(f, g, w, m)\} = \eta_S^2\{1+\lambda_{\max}(\mathbf{GH}_S) + v\Omega^{-1/2}\alpha_m^{3/2}\}$.

4. OPTIMAL DESIGNS WITH HOMOSCEDASTICITY: SOLUTION TO P1

Problem P1 has become that of finding a density $m_*(\mathbf{x})$, which minimizes (6). The solution is given by Theorem 5, which reduces the problem to a $2(p+1)$-dimensional numerical problem.

Theorem 5
The design density $m_*(\mathbf{x})$ minimizing (6) for the OLS estimation under homoscedasticity is of the form

$$m_*(\mathbf{x}) = \left[\frac{\mathbf{t}^{\mathrm{T}}(\mathbf{x})\mathbf{P}\mathbf{t}(\mathbf{x}) + \delta}{\mathbf{t}^{\mathrm{T}}(\mathbf{x})\mathbf{Q}\mathbf{t}(\mathbf{x})}\right]^+$$

where $(z)^+ = \max(z, 0)$. The $(p+2)\times(p+2)$ symmetric matrices $\mathbf{P}$, $\mathbf{Q}(\geqslant \mathbf{0})$ and a constant δ satisfy $\int_S m_*(\mathbf{x})\,d\mathbf{x} = 1$ and minimize (6).

The proofs of this and the other two theorems in this section are similar to those of Theorems 1, 2 and 3 in [13], respectively, and thus are omitted.

Some special cases of the general exponential regression model considered in this paper have their particular applications. Certainly, the results of this and other sections apply to all special cases. Note that, in general $\boldsymbol{\theta}$ is a $p+2$-dimensional vector, so is $\mathbf{t}(\mathbf{x})$. For special cases, $\boldsymbol{\theta}$ may be either a $(p+1)\times 1$ or a $p\times 1$ vector. For instance, when $\alpha=0, \beta=1$, or $\alpha=-\beta=1$, the vector $\boldsymbol{\theta}$ becomes $(\beta, \boldsymbol{\lambda}^{\mathrm{T}})^{\mathrm{T}}$, $(\alpha, \boldsymbol{\lambda}^{\mathrm{T}})^{\mathrm{T}}$, or $\boldsymbol{\lambda}$, respectively. The following presents four models commonly used in crop growth analysis (see [18]), pharmacokinetics and other areas (see [5]). These models serve as typical examples throughout this section and the sections hereafter. We consider that the regression function assumed model is an approximation to the truth. However, all the results derived from this paper are in complete generality and apply for all multiple covariate case as well, see Example 2 for an example with two covariates.

Model 1:

$$E(Y|x) \approx \alpha + \mathrm{e}^{-\lambda x}$$

with $\mathbf{t}(x) = (1, -x\mathrm{e}^{-\lambda x})^{\mathrm{T}}$, where $x\in[0, 1]$.

Model 2:

$$E(Y|x) \approx \alpha(1-\mathrm{e}^{-\lambda x})$$

with $\mathbf{t}(x) = (1-\mathrm{e}^{-\lambda x}, -\alpha x\mathrm{e}^{-\lambda x})^{\mathrm{T}}$, where $x\in[0, 1]$.

Model 3:

$$E(Y|x) \approx 1-\mathrm{e}^{-\lambda x} \quad \text{with } \lambda>0$$

with $t(x) = x\mathrm{e}^{-\lambda x}$, where $x\in[0, 1]$.

Copyright © 2009 John Wiley & Sons, Ltd.

DOI: 10.1002/asmb

Model 4:

$$E(Y|x) \approx \beta \mathrm{e}^{-\lambda x}$$

with $\mathbf{t}(x) = (\mathrm{e}^{-\lambda x}, -\beta x \mathrm{e}^{-\lambda x})^{\mathrm{T}}$, where $x \in [0, 1]$.

Model 5: below is a general version of the above models,

$$E(Y|x) \approx \alpha + \beta \mathrm{e}^{-\lambda x}$$

with $\mathbf{t}(x) = (1, \mathrm{e}^{-\lambda x}, -\beta x \mathrm{e}^{-\lambda x})^{\mathrm{T}}$, where $x \in [0, 1]$.

Example 1

For Model 1, by Theorem 5 the optimal design density has the form

$$m_*(x) = \left[\frac{a_1 + a_2 x \mathrm{e}^{-\lambda x} + a_3 x^2 \mathrm{e}^{-2\lambda x}}{a_4 + a_5 x \mathrm{e}^{-\lambda x} + a_6 x^2 \mathrm{e}^{-2\lambda x}} \right]^+ \tag{8}$$

Example 2

Consider an approximate exponential regression model with two covariates: $E(Y|x) \approx \mathrm{e}^{-\lambda_1 x_1 - \lambda_2 x_2}$, and the design space $S = [0, 1] \times [0, 1]$. Applying Theorem 5, the resulting minimax design density is in the form

$$m_*(x) = \left[\begin{array}{c} (a_1 x_1^2 + a_2 x_1 x_2 + a_3 x_2^2) \mathrm{e}^{-2(\lambda_1 x_1 + \lambda_2 x_2)} + a_4 \\ (a_5 x_1^2 + a_6 x_1 x_2 + a_7 x_2^2) \mathrm{e}^{-2(\lambda_1 x_1 + \lambda_2 x_2)} \end{array} \right]^+$$

Example 3

We consider Model 5 and apply Theorem 6. The resulting minimax design density is in the form

$$m_*(x) = \left[\frac{a_1 + a_2 \mathrm{e}^{-\lambda x} + a_3 x \mathrm{e}^{-\lambda x} + a_4 \mathrm{e}^{-2\lambda x} + a_5 x \mathrm{e}^{-2\lambda x} + a_6 x^2 \mathrm{e}^{-2\lambda x}}{b_1 + b_2 \mathrm{e}^{-\lambda x} + b_3 x \mathrm{e}^{-\lambda x} + b_4 \mathrm{e}^{-2\lambda x} + b_5 x \mathrm{e}^{-2\lambda x} + b_6 x^2 \mathrm{e}^{-2\lambda x}} \right]^+$$

5. OPTIMAL DESIGNS WITH HETEROSCEDASTICITY

The problems P2 and P3 have become the following:

(P2) Find a density $m_*(\mathbf{x})$ that minimizes

$$\eta_S^{-2} \sup_{f \in \mathscr{F}, g \in \mathscr{G}} \mathrm{IMSE}(f, g, \mathbf{1}, m) = 1 + \lambda_{\max}(\mathbf{K}\mathbf{H}_S) + v\Omega^{-1/2} \left[\int_S \{l_m(\mathbf{x}) m(\mathbf{x})\}^2 \, \mathrm{d}\mathbf{x} \right]^{1/2} \tag{9}$$

with $\lambda_{\max}$ and $l_m(\mathbf{x})$ as defined in Theorems 1 and 3, respectively. Then, $k_*(\mathbf{x}) = m_*(\mathbf{x})$ is the optimal design density for the OLS estimation.

(P3) Find a density $m_*(\mathbf{x})$ that minimizes

$$\eta_S^{-2} \sup_{f \in \mathscr{F}, g \in \mathscr{G}} \mathrm{IMSE}(f, g, w_m, m) = 1 + \lambda_{\max}(\mathbf{K}\mathbf{H}_S) + v\Omega^{-1/2} \left[\int_S \{l_m(\mathbf{x}) m^2(\mathbf{x})\}^{2/3} \, \mathrm{d}\mathbf{x} \right]^{3/2} \tag{10}$$

Copyright © 2009 John Wiley & Sons, Ltd.

DOI: 10.1002/asmb

Then the weights

$$w_*(\mathbf{x}) = \alpha_{m_*}\{l^2_{m_*}(\mathbf{x})m_*(\mathbf{x})\}^{-1/3} I[m_*(\mathbf{x})>0] \tag{11}$$

and the density

$$k_*(\mathbf{x}) = \alpha^{-1}_{m_*}[l_{m_*}(\mathbf{x})m^2_*(\mathbf{x})]^{2/3} \tag{12}$$

with α_{m_*} defined in Theorem 4 are optimal for the WLS estimation.

5.1. *Minimax designs for OLS: solution to P2*

The solution to P2 is provided by Theorem 6 below.

Theorem 6
The density $m_*(\mathbf{x})$ minimizing (9) for the OLS estimation under heteroscedasticity is of the form

$$m_*(\mathbf{x}) = \left[\frac{\mathbf{t}^{\mathrm{T}}(\mathbf{x})\mathbf{P}\mathbf{t}(\mathbf{x})+\delta}{\mathbf{t}^{\mathrm{T}}(\mathbf{x})\mathbf{Q}\mathbf{t}(\mathbf{x})+\left\{\mathbf{t}^{\mathrm{T}}(\mathbf{x})\mathbf{U}\mathbf{t}(\mathbf{x})\right\}^2}\right]^+ \tag{13}$$

The $(p+2)\times(p+2)$ symmetric matrices $\mathbf{P}, \mathbf{Q}(\geqslant\mathbf{0}), \mathbf{U}(>\mathbf{0})$, and a constant δ satisfy $\int_S m_*(\mathbf{x})\,\mathrm{d}\mathbf{x}=1$, and minimize (9).

Example 4
For Model 4, we apply Theorem 6. The resulting minimax design density is in the form

$$m_*(x) = \left(\frac{(a_1+b_1x+c_1x^2)\mathrm{e}^{-2\lambda x}+d}{(a_2+b_2x+c_2x^2)\mathrm{e}^{-2\lambda x}+(a_3+b_3x+c_3x^2)^2\mathrm{e}^{-4\lambda x}}\right)^+$$

5.2. *Minimax designs for WLS: solution to P3*

The solution to P3 is provided by Theorem 7.

Theorem 7
The minimizing $m_*(\mathbf{x})$ in (10) for the WLS estimation is of the form

$$m_*(x) = \frac{[a(\mathbf{x})-b(\mathbf{x})]^+}{c(\mathbf{x})} \tag{14}$$

where, for constant symmetric matrices $\mathbf{P}$, $\mathbf{Q}(\geqslant\mathbf{0})$, $\mathbf{U}(>\mathbf{0})$ and a constant d we have

$$a(\mathbf{x}) = \mathbf{t}^{\mathrm{T}}(\mathbf{x})\mathbf{P}\mathbf{t}(\mathbf{x})+d, \quad c(\mathbf{x}) = \mathbf{t}^{\mathrm{T}}(\mathbf{x})\mathbf{Q}\mathbf{t}(\mathbf{x})$$

and where b satisfies the cubic equation

$$b^3+\frac{r^3}{c}b-\frac{r^3a}{c}=0$$

Copyright © 2009 John Wiley & Sons, Ltd.

with $r(\mathbf{x})=\{\mathbf{t}^{\mathrm{T}}(\mathbf{x})\mathbf{U}\mathbf{t}(\mathbf{x})\}^{2/3}$. Explicitly,

$$b(\mathbf{x})=r\left[\left\{\frac{a}{2c}+\sqrt{\left(\frac{a}{2c}\right)^2+\left(\frac{r}{3c}\right)^2}\right\}^{1/3}+\left\{\frac{a}{2c}-\sqrt{\left(\frac{a}{2c}\right)^2+\left(\frac{r}{3c}\right)^2}\right\}^{1/3}\right]$$

The constants in a, b and c satisfy $\int_S m_*(\mathbf{x})\,\mathrm{d}\mathbf{x}=1$ and minimize (10). Then (11) and (12) provide the optimal regression weights and design density for the WLS estimation, respectively.

Example 5
For Model 2, by Theorem 7, the optimal minimax $m_*(x)=k_*(x)w_*(x)$ for WLS has the form (14) with

$$a(x)=a_1(1-\mathrm{e}^{-\lambda x})^2+a_2x\mathrm{e}^{-\lambda x}(1-\mathrm{e}^{-\lambda x})+a_3x^2\mathrm{e}^{-2\lambda x}+d$$

$$c(x)=c_1(1-\mathrm{e}^{-\lambda x})^2+c_2x\mathrm{e}^{-\lambda x}(1-\mathrm{e}^{-\lambda x})+c_3x^2\mathrm{e}^{-2\lambda x}$$

$$r(x)=[r_1(1-\mathrm{e}^{-\lambda x})^2+r_2x\mathrm{e}^{-\lambda x}(1-\mathrm{e}^{-\lambda x})+r_3x^2\mathrm{e}^{-2\lambda x}]^{2/3}$$

The minimax design ξ_* has density $k_*(x)$ computed from (12). The minimax weights $w_*(x)$ are obtained from (11).

6. OPTIMAL UNBIASED DESIGNS: SOLUTION TO P4

We say that a design/weights pair (ξ, w) is *unbiased* if it satisfies $\mathrm{IB}(f, w, \xi)=0$ for all $f\in\mathscr{F}$, so that $\sup_{f\in\mathscr{F}}\mathrm{IB}(f, w, \xi)=0$. The following theorem gives a necessary and sufficient condition for unbiasedness.

Theorem 8
The pair (w, ξ) is unbiased if and only if $m(\mathbf{x})\equiv\Omega=1/\int_S \mathrm{d}\mathbf{x}$.

Proof of Theorem 8
For the sufficiency, $m(\mathbf{x})\equiv\Omega$ implies $\mathbf{G}=0$. According to Theorem 1, $\sup_{f\in\mathscr{F}}\mathrm{IB}(f, w, \xi)=0$. For the necessity, $\sup_{f\in\mathscr{F}}\mathrm{IB}(f, w, \xi)=0$ implies $\mathbf{G}=0$. By (5), $(m(\mathbf{x})\mathbf{I}-\mathbf{B}\mathbf{A}_S^{-1})\mathbf{t}(\mathbf{x})=0$ almost everywhere $\mathbf{x}\in S$. The rest part of the proof can be done using the same technique as that in the proof of Theorem 2(b) in [11].

We can construct the optimal unbiased design $m_0(\mathbf{x})$ by forcing $\sup_{f\in\mathscr{F}}\mathrm{IB}(f, w, \xi)=0$ through the choice $k=\Omega/w$, and then minimizing $\sup_{g\in\mathscr{G}}\mathrm{IV}(g, w, \xi)$ over w. From Theorem 4, the optimal weight function is

$$w_0(\mathbf{x})=\Omega\alpha_0[\mathbf{t}^{\mathrm{T}}(\mathbf{x})\mathbf{A}_S^{-1}\mathbf{t}(\mathbf{x})]^{-4/3}$$

and the optimal unbiased design density is

$$k_0(\mathbf{x})=\alpha_0^{-1}[\mathbf{t}^{\mathrm{T}}(\mathbf{x})\mathbf{A}_S^{-1}\mathbf{t}(\mathbf{x})]^{4/3}$$

with

$$\alpha_0=\int_S[\mathbf{t}^{\mathrm{T}}(\mathbf{x})\mathbf{A}_S^{-1}\mathbf{t}(\mathbf{x})]^{4/3}\,\mathrm{d}\mathbf{x}$$

Copyright © 2009 John Wiley & Sons, Ltd.

DOI: 10.1002/asmb

The minimax IMSE is

$$\min_{(w,\xi)} \sup_{f\in\mathscr{F},g\in\mathscr{G}} \mathrm{IMSE}(f,g,w,m)=\eta_S^2+\min_{(w,\xi)}\sup_{g\in\mathscr{G}}\mathrm{IV}(g,w_m,\xi)=\eta_S^2\{1+v\Omega^{-1/2}\alpha_0^{3/2}\}$$

We summarize these results below.

Theorem 9
The density $k_0(\mathbf{x})$ of the optimal unbiased design measure ξ_0 and optimal weights w_0, which minimize $\sup_{f\in\mathscr{F},g\in\mathscr{G}}\mathrm{IMSE}(f,g,w,\xi)$ subject to $\sup_{f\in\mathscr{F}}\mathrm{IB}(f,w,\xi)=0$ are given by

$$k_0(\mathbf{x})=\frac{[\mathbf{t}^{\mathrm{T}}(\mathbf{x})\mathbf{A}_S^{-1}\mathbf{t}(\mathbf{x})]^{2/3}}{\int_S[\mathbf{t}^{\mathrm{T}}(\mathbf{x})\mathbf{A}_S^{-1}\mathbf{t}(\mathbf{x})]^{2/3}\,\mathrm{d}\mathbf{x}}$$

and $w_0(\mathbf{x})=\Omega/k_0(\mathbf{x})$. Minimax IMSE is

$$\sup_{f\in\mathscr{F},g\in\mathscr{G}}\mathrm{IMSE}(f,g,w_0,\xi_0)=\eta_S^2\left\{1+v\Omega^{-1/2}\left[\int_S[\mathbf{t}^{\mathrm{T}}(\mathbf{x})\mathbf{A}_S^{-1}\mathbf{t}(\mathbf{x})]^{2/3}\,\mathrm{d}\mathbf{x}\right]^{3/2}\right\} \tag{15}$$

attained at $g_0(\mathbf{x})=w_0^{-1/2}(\mathbf{x})$.

This is an immediate result from Theorems 4 and 8.

Example 6
Consider Model 3. Note that the designs provided by Theorem 9 for this example depend only on λ but not on α. To deal with this issue, we search for 'locally most robust' designs as discussed in [14]. We consider a neighbourhood Θ. By setting a start value $\lambda^{(0)}\in\Theta$, we first construct the design $k_0(x,\lambda^{(0)})$ and weights $\Omega/k_0(x,\lambda^{(0)})$ provided by Theorem 9. We then find the least favourable λ^{LF} in Θ. From Theorem 3, we find that this is equivalent to maximizing

$$\begin{aligned}\int_S\frac{\{\mathbf{t}^{\mathrm{T}}(x)\mathbf{A}_S^{-1}\mathbf{t}(x)\}^2}{k_0^2(x,\lambda^{(0)})}\,\mathrm{d}x&=\frac{3\lambda^2}{\lambda^{(0)}(1-\mathrm{e}^{-2\lambda})^2}\int_S\mathrm{e}^{-4\lambda x}\left(\frac{1-\mathrm{e}^{-4\lambda^{(0)}x}/3}{\mathrm{e}^{-4\lambda^{(0)}x/3}}\right)\mathrm{d}x\\&=\frac{3\lambda^2}{\lambda^{(0)}(1-\mathrm{e}^{-2\lambda})^2}\left(\frac{\mathrm{e}^{4(\lambda^{(0)}/3-\lambda)}-1}{4(\lambda^{(0)}/3-\lambda)}+\frac{\mathrm{e}^{-4\lambda}-1}{4\lambda}\right)\end{aligned}$$

over the occurrences of λ in Θ. We then construct the unbiased optimal design for λ^{LF}, and iterate to convergence. With $S=[0,1]$, for a simple demonstration of the procedure described above, we start at $\lambda^{(0)}=0.75$ and take $\Theta=[0.5,1]$. The iteration converges to $\lambda^{\mathrm{LF}}=0.5$. The unbiased minimax design density is

$$k_0(x)=1.37\mathrm{e}^{-(2/3)x}$$

and the corresponding optimal weights are $w_0(x)=0.78\mathrm{e}^{(2/3)x}$.

Copyright © 2009 John Wiley & Sons, Ltd.

DOI: 10.1002/asmb

7. COMPUTATION AND DISCRETIZATION

In this section, we first demonstrate the computation of numerical values of the constants in our constructed design density using one typical example: optimal design for P1 with Model 1; then we present the discretized version of these designs for implementation.

As indicated in Example 1, the optimal design density has the form of (8). Note that (8) is over-parameterized—if one of a_1–a_6 is nonzero—then we can assume that it is unity. We let $a_4=1$, and choose a_1, a_2, a_3, a_5 and a_6 such that (6) can be minimized subject to $\int_S m_*(\mathbf{x})\,\mathrm{d}\mathbf{x}=1$.

The optimal design m_* depends on λ, but not on α. To address this issue one may adopt a mixture of minimax and local approaches (see [13]): (1) start at some $\lambda_1=\lambda_1^{(0)}$ and find the corresponding optimal design density: $m_*^{(0)}(x)$; (2) maximize (6) with $m=m_*^{(0)}$ over an interval containing $\lambda^{(0)}$ to find the least favourable value of λ: $\lambda^{(1)}$; (3) iterate between minimizing over designs and maximizing over λ until attaining convergence, say to λ^{LF}; and (4) finally, employ Theorem 5 to construct the 'locally most robust design' density $m_*(x;\lambda^{\mathrm{LF}})$.

We consider a region $\lambda\in[0.5,1]$ as that in [5] and carry out the process described above for several values of v, each time starting at $\lambda^{(0)}=0.75$. In each case we obtain $\lambda^{\mathrm{LF}}=2$, see Table I for the numerical values of the constants for both locally optimal design density at $\lambda=0.75$ and locally most robust design density when $\lambda\in[0.5,1]$ and see Figure 1 for their plots.

Table I. Numerical values for Example 1.

	$\lambda^{(0)}=0.75$					$\lambda\in[0.5,1]$					
v	a_1	a_2	a_3	a_5	a_6	a_1	a_2	a_3	a_5	a_6	λ^{LF}
0.5	1.266	0.0113	−0.771	1.348	−2.030	1.206	1.646	22.065	2.014	23.541	0.5
1	1.628	0.0000142	−4.328	7.644	−0.189	1.988	0.0375	−4.188	7.628	−14.634	0.5
2	2.870	0.0627	−11.522	17.310	−40.568	2.716	0.0117	−5.237	14.675	−26.076	0.5

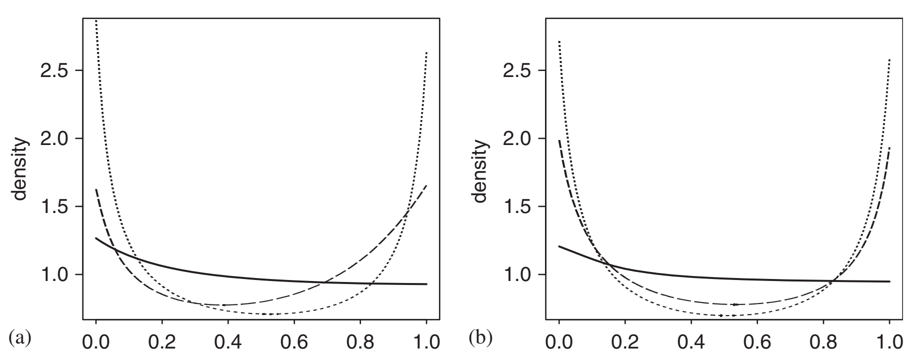

Figure 1. Optimal minimax design densities $m_*(x;\lambda)$ in Example 1: (a) locally most robust design densities for $\lambda=\lambda^{\mathrm{LF}}$ in $[0.5,1]$ and (b) locally robust design densities for $\lambda=\lambda^{(0)}=0.75$. Each plot uses three values of $v: v=0.5$ (solid line), $v=1$ (broken line), $v=5$ (dotted line).

Copyright © 2009 John Wiley & Sons, Ltd.

DOI: 10.1002/asmb

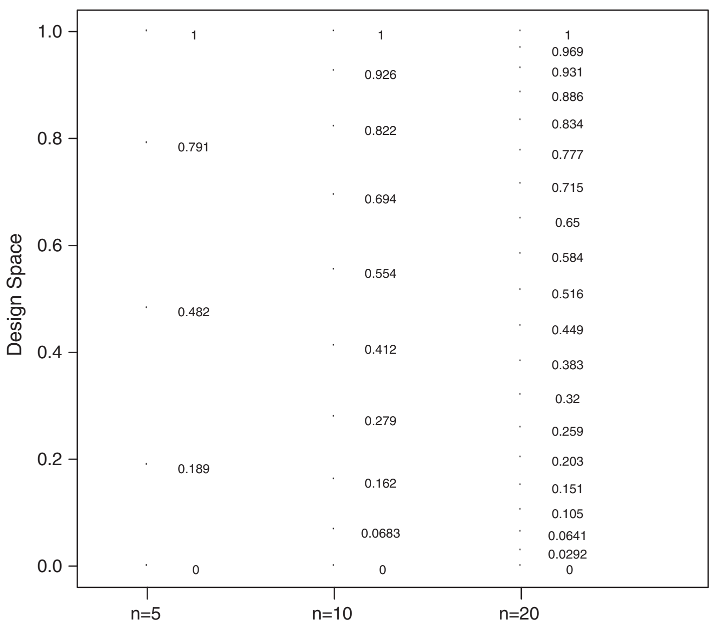

Figure 2. Support points of the discretized locally most robust designs for Model 1 when λ^{LF} and $v=5$.

We discretize the optimal design ξ constructed above by placing mass n^{-1} at each of the design points: $\xi^{-1}((i-1)/(n-1))$. For $n=5, 10, 20$ and $v=1$, the support points of the discretized locally most robust designs are shown in Figure 2.

We have provided the methods of constructing optimally robust designs for general exponential regression under consideration of model uncertainties in both response function and variance component. As v increases, the designs should have more emphasis on variance minimization and less on protection from bias. As we would expect, the experimenter should then place relatively more design points closer to the boundary of the design space. As v decreases, the designs should have more emphasis on protection from bias and less on variance reduction; hence, is more uniform.

ACKNOWLEDGEMENTS

This research is supported by the Natural Sciences and Engineering Research Council of Canada (341783-07). The author would like to thank two anonymous referees for their helpful comments on the previous version of this paper.

REFERENCES

1. Landaw EW, DiStefano III JJ. Multiexponential, multicompartmental, and noncompartmental modeling. II. Data analysis and statistical consideration. *American Journal of Physiology* 1984; **246**:665–677.
2. Liebig HP. Temperature integration by Kohlrabi growth. *Acta Horticulturae* 1988; **230**:371–380.
3. Nelson W. *Accelerated Testing—Statistical Models, Test Plans, and Data Analyses*. Wiley: New York, 1990.

Copyright © 2009 John Wiley & Sons, Ltd.

4. Melas VB. Optimal designs for exponential regression. *Math. Operationsforsch*, *Statist.*, *Ser. Statistics* 1978; **9**:45–59.
5. Dette H, Lopez IM, Rodriguez IMO, Pepelyshev A. Maximin efficient design of experiment for exponential regression models. *Journal of Statistical Planning and Inference* 2006; **136**:4397–4418.
6. Fang Z, Wiens DP. Bayesian minimally supported D-optimal designs for an exponential regression model. *Communications in Statistics*, *Theory and Methods* 2004; **33**:1187–1204.
7. Box GEP, Draper NR. A basis for the selection of a response surface design. *Journal of the American Statistical Association* 1959; **54**:622–654.
8. Kiefer J. Optimal designs for fitting multiresponse surface. *Symposium on Multivariate Analysis III*. Academic Press: New York, 1973; 287–297.
9. Huber PJ. Robustness and designs. In *A Survey of Statistical Design and Linear Models*, Srivastava JN (ed.). Elsevier: Amsterdam, North Holland, 1975; 287–303.
10. Wiens DP. Minimax designs for approximately linear regression. *Journal of Statistical Planning and Inference* 1992; **31**:353–371.
11. Wiens DP. Minimax robust designs and weights for approximately specified regression models with heteroscedastic errors. *Journal of the American Statistical Association* 1998; **93**:1440–1450.
12. Chang YJ, Notz WI. Model robust designs. In *Handbook of Statistics*, Ghosh S, Rao CR (eds), vol. 13. Elsevier: Amsterdam, North Holland, 1996; 1055–1098.
13. Wiens DP, Xu X. Robust prediction and extrapolation designs for misspecified generalized linear regression models. *Journal of Statistical Planning and Inference* 2008; **138**:30–46.
14. Wiens DP, Xu X. Robust designs for one-point extrapolation. *Journal of Statistical Planning and Inference* 2008; **138**:1339–1357.
15. Seber GAF, Wild CJ. *Nonlinear Regression*. Wiley: New York, 2003.
16. Heo G, Schmuland B, Wiens DP. Restricted minimax robust designs for misspecified regression models. *Canadian Journal of Statistics* 2001; **29**:117–128.
17. Xu X. Robust prediction and extrapolation designs, with applications to accelerated life testing. *Ph.D. Thesis*, University of Alberta, Canada, 2006.
18. Krug H, Liebig HP. Static regression models for planning greenhouse production. *Acta Horticulturae* 1998; **230**:427–433.

Copyright © 2009 John Wiley & Sons, Ltd.

DOI: 10.1002/asmb