Published in final edited form as:

Can J Stat. 2018 June ; 46(2): 336–354. doi:10.1002/cjs.11355.

Locally optimal designs for some binary dose-response models

Yi Zhai^{1,2} and Zhidong Fang^{2,*}

¹School of Information, Qilu University of Technology, Jinan, Shandong, China

²Biostatistics Program, School of Public Health, Louisiana State University Health Sciences Center, New Orleans, USA

Abstract

In this paper, we consider the problem of seeking locally optimal designs for nonlinear dose-response models with binary outcomes. Applying the theory of Tchebycheff Systems and other algebraic tools, we show that the locally D -, A -, and c -optimal designs for three binary dose-response models are minimally supported in finite, closed design intervals. The methods to obtain such designs are presented along with examples. The efficiencies of these designs are also discussed.

Key words and phrases

Binary; cancer research; dose-response; locally optimal designs; Tchebycheff systems

1. Introduction

In a study involving a binary dose-response model, one observes with random error a response variable $y(x)$, which takes the value 1 or 0, and has expected value

$$P(x, \theta) = \eta(x, \theta), \quad (1)$$

corresponding to the input variable (dose level) x lying in an interval $\mathcal{X} = [x_L, x_U]$, $0 \leq x_L < x_U < \infty$. The function η is usually nonlinear and twice differentiable. The vector θ consists of unknown parameters, to be estimated. Model (1) has been widely employed in cancer/toxicity studies, where $P(x, \theta)$ represents the lifetime probability of cancer at dose level x , or the probability of a subject responding at dose level x when exposed to (environmental) chemicals. Among the numerous choices for η in the literature, the logistic model is the most popular one. Another model is the so-called multi-stage model, assuming that the mechanism of carcinogenesis can be interpreted as a series of multiple somatic mutations at the cellular level. After going through this series of mutational stages, a cell becomes malignant and proceeds to develop into a tumor. The biological background of the model can be seen, for example, in Armitage & Doll (1954), Portier & Hoel (1983), Portier &

* Author to whom correspondence may be addressed. zfang@lsuhsc.edu.

Elder (1990) and references therein, and the design problem for the model is discussed in Fang, Wiens & Wu (2006).

Generally, the design problem in such studies can be defined as follows: if n subjects are tested at n dose levels, $x_1, \dots, x_n$, then a design, denoted as ξ , is defined as the empirical distribution function of x_j . In other words, assume without loss of generality that $x_1, \dots, x_k$ are distinct values of these dose levels, then $\xi(x_j) = \frac{n_j}{n}$ if n_j subjects are tested at dose level x_j and $\sum_{j=1}^k n_j = n$. Denote $f(x, \theta) = \frac{\partial P(x, \theta)}{\partial \theta}$. To obtain the maximum likelihood estimate of θ where $y(x_j) \sim \text{binomial}(1, P(x_j, \theta))$, we have the likelihood function

$$L(\theta) = \prod_{j=1}^n P(x_j, \theta)^{y(x_j)} (1 - P(x_j, \theta))^{1 - y(x_j)},$$

the log-likelihood function

$$l(\theta) = \sum_{j=1}^n \left(y(x_j) \log P(x_j, \theta) + (1 - y(x_j)) \log (1 - P(x_j, \theta)) \right),$$

and Hessian matrix

$$\ddot{l}(\theta) = \frac{\partial^2 l(\theta)}{\partial \theta \partial \theta^T} = \sum_{j=1}^n \left(y(x_j) \frac{\frac{\partial^2 P(x_j, \theta)}{\partial \theta \partial \theta^T} P(x_j, \theta) - \mathbf{f}(x_j, \theta) \mathbf{f}^T(x_j, \theta)}{P^2(x_j, \theta)} - (1 - y(x_j)) \frac{\frac{\partial^2 P(x_j, \theta)}{\partial \theta \partial \theta^T} (1 - P(x_j, \theta)) + \mathbf{f}(x_j, \theta) \mathbf{f}^T(x_j, \theta)}{(1 - P(x_j, \theta))^2} \right).$$

Thus the Fisher information matrix of a design ξ is given by

$$\begin{aligned} I(\theta, \xi) &= E[-\ddot{l}(\theta)] \\ &= \sum_{i=1}^n \frac{\mathbf{f}(x_i, \theta) \mathbf{f}^T(x_i, \theta)}{P(x_i, \theta)(1 - P(x_i, \theta))} \\ &= \int_{x_L}^{x_U} \frac{\mathbf{f}(x, \theta) \mathbf{f}^T(x, \theta)}{P(x, \theta)(1 - P(x, \theta))} d\xi(x). \end{aligned}$$

By the theory of optimal designs, the criterion for screening for the optimal design is generally defined as some scalar function of the information matrix. The optimal design is

the one which maximizes (or minimizes) the pre-chosen criterion among a class of candidate designs. For the binary dose-response model (1), the information matrix usually depends on an unknown vector of parameters θ . This makes the design problem more difficult than those for linear regression models. One way to handle this difficulty is to assume a prior distribution $\delta(\theta)$ on θ , and thus construct Bayesian optimal designs. For example, the Bayesian D -optimal design for model (1) maximizes, over ξ , the function

$$\phi_B(\xi) = \int_{\Theta} |I(\theta, \xi)| \delta(d\theta),$$

where Θ is the parameter space of θ . Other Bayesian optimal criteria can be defined similarly. When $\delta(\theta^*) = 1$ for some given θ^* or when Θ is a singleton, the Bayesian optimal design reduces to the locally optimal design.

In this paper, for some binary dose-response models we construct locally D -, A -, and c -optimal designs, which minimize the determinant, the trace of $I^{-1}(\theta, \xi)$, $c^T I^{-1}(\theta, \xi) c$ separately, with c being a non-zero vector. Each model has its own motivation for being considered in practice. The first model we consider has the probability of response at dose level x given by

$$\eta(c, \theta) = 1 - e^{-\theta_0 - \theta_1 x - \theta_2 x^2}, \quad (2)$$

where $\theta_i > 0$, $i = 0, 1, 2$. We call model (2) an Extended Weibull model. Note that this model can be viewed as a special multi-stage model if θ_2 is a fixed integer (Fang, Wiens & Wu, 2006). However, the information matrix in model (2) has higher dimension than in the corresponding multi-stage model, and this makes the construction of an optimal design more complicated. When $\theta_0 = 0$ in model (2), we obtain the Weibull model (Krewski & Van Ryzin, 1981):

$$\eta(x, \theta) = 1 - e^{-\theta_1 x - \theta_2 x^2}, \quad (3)$$

where $\theta_1, \theta_2 > 0$. This model can also be applied in microbial risk assessment to model the probability of infection at dose x (Buchanan, Smith & Long, 2000), where θ_1 represents the infectivity and θ_2 is the shape parameter. The third model we consider is the Beta-Poisson model (Haas, 1983; Buchanan, Smith & Long, 2000), which takes the form

$$\eta(x, \theta) = 1 - \left(1 + \frac{x}{\theta_2}\right)^{-\theta_1}, \quad (4)$$

with $\theta_1 > 0$ being the infectivity and $\theta_2 > 0$ the shape parameter.

In addition to their formulations, these models also differ in terms of their baseline effects. Models (3) and (4) assume no baseline effect in that the probability of response is zero at dose level 0. This is not the case for model (2), which assumes a non-zero baseline effect.

Locally optimal designs for binary dose-response models have been introduced and widely discussed in Chernoff (1953), Fedorov (1972), Silvey (1980), Ford, Titterington & Kitsos (1989), Ford, Torsney & Wu (1992), Sitter & Wu (1993), Dette & Haines (1994), Sitter & Fainaru (1997), Fedorov & Leonov (2001, 2014), Biedermann, Dette & Zhu (2006), Fang, Wiens & Wu (2006), Yang & Stufken (2009), Stufken & Yang (2012), and many others. They differ with respect to model selection, optimality criteria and techniques of design construction. To our knowledge, the design problem for models (2) - (4) has not been considered in the literature.

Note that the design defined above with mass $\xi(x_j) = \frac{n_j}{n}$ is called an exact design and is the one implemented in practical situations. Since the exact design problem often leads to an intractable integer programming problem, we restrict ourselves to approximate designs, that is, discrete probability measures on the design space. Approaches to implement such approximate designs are discussed by Pukelsheim (1993, chapter 12) and others.

2. Optimal Designs for Extended Weibull Model

In this section, we construct locally D -, A -, and c -optimal designs for model (2). Using the theory of Tchebycheff Systems (T -systems, Karlin and Studden, 1966), we show that locally D -, A -, and c -optimal designs have minimal number of support points. Hereafter, for any monotone function $z = \phi(x)$ in $\mathcal{X} = [x_L, x_U]$, denote the range as $\mathcal{Z}$ with endpoints $z_L = \phi(x_L)$ and $z_U = \phi(x_U)$.

2.1. Locally D -optimal designs

We consider only designs with at least three support points so that the Fisher information matrix is invertible. The following theorem shows that the locally D -optimal design is minimally supported. Consequently, it has equal weight on each support point (Silvey, 1980).

Theorem 1. Under model (2), the locally D -optimal design is supported on three points with equal weights, including the left boundary point x_L of the design interval.

Proof. Let $z = \phi(x) = \theta_1 x^{\theta_2}$. The Fisher information matrix of a single observation at z can be written in the form

$$\mu(z, \theta) = \frac{\exp(-\theta_0 - z)}{1 - \exp(-\theta_0 - z)} P(\theta) \mathbf{q}(z) \mathbf{q}^T(z) P^T(\theta),$$

where $\mathbf{q}(z) = (1, z, z \log z)^T$, and

$$P(\boldsymbol{\theta}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\theta_1} & 0 \\ 0 & -\frac{\log \theta_1}{\theta_2} & \frac{1}{\theta_2} \end{pmatrix}.$$

Thus, for any design ξ , the variance function $d(z, \xi, \boldsymbol{\theta})$ is

$$d(z, \xi, \boldsymbol{\theta}) = \text{tr}(\mu(z, \boldsymbol{\theta})I^{-1}(\boldsymbol{\theta}, \xi)) = \frac{\exp(-\theta_0 - z)}{1 - \exp(-\theta_0 - z)} \mathbf{q}^T(z)P^T(\boldsymbol{\theta})I^{-1}(\boldsymbol{\theta}, \xi)P(\boldsymbol{\theta})\mathbf{q}(z).$$

By Kiefer-Wolfowitz's Equivalence Theorem, a design ξ_D is the locally D -optimal design if and only if $d(z, \xi_D, \boldsymbol{\theta}) \leq 3$. That is,

$$\mathbf{q}^T(z)P^T(\boldsymbol{\theta})I^{-1}(\boldsymbol{\theta}, \xi_D)P(\boldsymbol{\theta})\mathbf{q}(z) - 3e^{\theta_0}e^z + 3 \leq 0 \quad (5)$$

with equality being held at the support points of ξ_D . The left-hand side of (5) can be written as

$$H_D(z) = k_1 + k_2z + k_3z^2 + k_4z^2 \log z + k_5z \log z + k_6z^2 \log^2 z + k_7e^z, \quad (6)$$

where $k_1, k_3, k_6 > 0$ (because the matrix $I^{-1}(\boldsymbol{\theta}, \xi_D)$ is positive definite), and $k_7 < 0$. Now, it is easy to show that the sequence of functions (see Appendix A for details)

$$\{1, z, z^2, z^2 \log z, z \log z, z^2 \log^2 z, e^z\} \quad (7)$$

is a T -system in $\mathcal{R}^+$. Thus, by Theorem 4.2 in Chapter I of Karlin and Studden (1966), the Tchebycheff polynomial $H_D(z)$ in (6) has at most six zeros in $\mathcal{R}^+$, counting multiplicity (the non-nodal zeros are counted twice). Obviously, ξ_D should have at least 3 support points in $[z_L, z_U]$ for $I(\boldsymbol{\theta}, \xi_D)$ to be invertible. Assume that there are more than 3 support points in $[z_L, z_U]$. Then, since $H_D(z) = 0$ in $[z_L, z_U]$, we have two non-nodal zeros inside $[z_L, z_U]$. This, along with the fact that $H_D(z)$ has at most six zeros in $[z_L, z_U]$, implies that the two boundaries are nodal zeros. Thus, $H_D(z)$ cannot have other zeros outside $[z_L, z_U]$, which contradicts the fact that $\lim_{z \rightarrow +\infty} H_D(z) = -\infty$. Hence the equation in (5) has exactly three distinct zeros in $[z_L, z_U]$. Therefore, ξ_D has three support points in terms of z .

Assume that ξ_D has support points, $z_1 < z_2 < z_3$, in $[z_L, z_U]$. The logarithm of the determinant of the information matrix,

$$\begin{aligned}
\delta(z_1, z_2, z_3) &= \log |I(\boldsymbol{\theta}, \xi_D)| \\
&= \log \left(\frac{|P(\boldsymbol{\theta})|^2}{27} \frac{e^{-\theta_0 - z_1}}{1 - e^{-\theta_0 - z_1}} \frac{e^{-\theta_0 - z_2}}{1 - e^{-\theta_0 - z_2}} \frac{e^{-\theta_0 - z_3}}{1 - e^{-\theta_0 - z_3}} |(\mathbf{q}(z_1)\mathbf{q}(z_2)\mathbf{q}(z_3))|^2 \right) \\
&= -z_1 - z_2 - z_3 - \log(1 - e^{-\theta_0 - z_1}) - \log(1 - e^{-\theta_0 - z_2}) - \log(1 - e^{-\theta_0 - z_3}) + 2 \log \left(z_2 z_3 \log \frac{z_3}{z_2} - z_1 z_3 \log \frac{z_3}{z_1} + z_1 z_2 \log \frac{z_2}{z_1} \right) + \text{constant},
\end{aligned}$$

is a decreasing function of z_1 , given z_2, z_3 (proof in Appendix B). Therefore, (since $z = \phi(x)$ is strictly increasing) ξ_D has three support points, including x_L . This proves the theorem.

The other two design points of the locally D -optimal design can be determined by maximizing $\delta(z_L, z_2, z_3)$. The design points in x -interval can be obtained through $x = \phi^{-1}(z)$ (Ford, Torsney & Wu, 1992). As examples, Table 1 presents the support points of locally D -optimal designs for some selected values of θ_0, θ_1 and θ_2 when the design interval is $\chi = [1, 4]$. The values of $\log |I(\boldsymbol{\theta}, \xi_D)|$ are also included in the table. We observe that 1) the right boundary point may or may not be a design point; 2) the two flexible design points tend to become larger as θ_0 increases and smaller as θ_1 or θ_2 increases. The design points are most sensitive to changes in θ_2 , and least sensitive to θ_0 .

2.2. Locally A-optimal designs

For A -optimality, we also assume designs with at least three support points. Similar to locally D -optimal design, we show that the locally A -optimal design is minimally supported by using the Kiefer-Wolfowitz's Equivalence Theorem.

Theorem 2. Under model (2), the locally A -optimal design is minimally supported on three points.

Proof. Using the same notation as in Theorem 1, a design ξ_A is locally A -optimal if and only if

$$\text{tr}(\mu(z, \boldsymbol{\theta}) I^{-2}(\boldsymbol{\theta}, \xi_A)) - \text{tr}(I^{-1}(\boldsymbol{\theta}, \xi_A)) \leq 0, \quad (8)$$

that is,

$$\mathbf{q}^T(z) P^T(\boldsymbol{\theta}) I^{-2}(\boldsymbol{\theta}, \xi_A) P(\boldsymbol{\theta}) \mathbf{q}(z) - \text{tr}(I^{-1}(\boldsymbol{\theta}, \xi_A)) e^{\theta_0} e^z + \text{tr}(I^{-1}(\boldsymbol{\theta}, \xi_A)) \leq 0$$

with equality being held at the support points of ξ_A . Since $I^{-1}(\boldsymbol{\theta}, \xi_A) > 0$, the left hand side has the form

$$H_A(z) = \tilde{k}_1 + \tilde{k}_2 z + \tilde{k}_3 z^2 + \tilde{k}_4 z^2 \log z + \tilde{k}_5 z \log z + \tilde{k}_6 z^2 \log^2 z + \tilde{k}_7 e^z,$$

with $\tilde{k}_1, \tilde{k}_3, \tilde{k}_6 > 0$, and $\tilde{k}_7 < 0$, which is similar to $H_D(z)$ in (6). Following the lines after (6) in Theorem 1, we prove this theorem.

To determine the design points, x_1, x_2, x_3 , and the corresponding weights, w_1, w_2, w_3 , of ξ_A , we have, by Corollary 8.8 of Pukelsheim (1993),

$$w_i = \frac{\sqrt{u_{ii}}}{\sum_{j=1}^3 \sqrt{u_{jj}}}, \quad (9)$$

where u_{ii} is the i^{th} diagonal element of matrix $U = (XX^T)^{-1}$, and X is a (3×3) matrix with its i^{th} row as $f^T(x_i, \theta) / \sqrt{P(x_i, \theta)(1 - P(x_i, \theta))}$, $i = 1, 2, 3$. Thus, we need only to find the minimizer of a function of three variables, x_1, x_2, x_3 . Some numerical examples are presented in Table 2. We can see that the left boundary point of the design interval is one support point (we conjecture that this is generally true), but the right boundary point may or may not be a support point. When the initial values of the parameters change, the change of the support points of locally A -optimal designs is similar to that of locally D -optimal designs in both direction and sensitivity. However, the weights are no longer equal and are subject to the initial values of the parameters. Note that here the objective function of the minimization problem is $\text{tr}(I^{-1}(\theta, \xi_A)) = (\sum_{j=1}^3 \sqrt{u_{jj}})^2$, as stated in Corollary 8.8 of Pukelsheim (1993).

2.3. Locally c -optimal designs

We do not have a general result for the number of support points of a locally c -optimal design for any vector $c \neq 0$. However, we find that for special c such as standard basis vectors, locally c -optimal designs (for the individual regression coefficients) are minimally supported.

For model (2), using $z = \phi(x) = x^{\theta_2}$, the term $\frac{f^T(x, \theta)}{\sqrt{P(x, \theta)(1 - P(x, \theta))}}$ can be expressed as

$$g(z) = \sqrt{\frac{-\theta_0 - \theta_1 z}{1 - e^{-\theta_0 - \theta_1 z}}} \left(1, z, \frac{\theta_1}{\theta_2} z \log(z) \right)^T,$$

which is a T -system in $\mathcal{Z} = [z_L, z_U]$ because $\sqrt{\frac{-\theta_0 - \theta_1 z}{1 - e^{-\theta_0 - \theta_1 z}}} > 0$ and $\{1, z, \frac{\theta_1}{\theta_2} z \log(z)\}$ is a T -

system in $\mathcal{Z}$. We consider the c -optimal design in $\mathcal{Z}$. By Theorem 10.2 in Chapter II of Karlin & Studden (1966), there exists a Tchebycheff polynomial

$$W(z) = \sqrt{\frac{-\theta_0 - \theta_1 z}{1 - e^{-\theta_0 - \theta_1 z}}} (a_1 + a_2 z + a_3 \frac{\theta_1}{\theta_2} z \log(z)) \text{ with } |W(z)| \leq 1, \forall z \in \mathcal{Z}, \text{ and there exist three}$$

points, $z_1 < z_2 < z_3$, called Tchebycheff points, such that $W(z_i) = (-1)^i$, $i = 1, 2, 3$. Furthermore, we can show that $z_1 = z_L$.

In fact, due to the fact that $W(z)$ has at most two zeros in $\mathcal{Z}$ (Theorem 4.1 in Chapter I of Karlin & Studden (1966)) and that $W(z)$ has at least two zeros because $W(z_i) = (-1)^i$, $i = 1, 2, 3$, we conclude that $W(z)$ has exactly two zeros in $\mathcal{Z}$. Now, without loss of generality, we

assume $a_3 < 0$. Thus, the function $h_1(z) = \sqrt{\frac{-\theta_0 - \theta_1 z}{1 - e^{-\theta_0 - \theta_1 z}}}$ is decreasing, and the function

$h_2(z) = a_1 + a_2 z + a_3 \frac{\theta_1}{\theta_2} z \log(z)$ is concave in $\mathcal{Z}$. These indicate that $W(z) (= h_1(z)h_2(z))$ is strictly increasing in the interval between z_L and its first zero. Thus, $z_1 = z_L$.

The coefficients of the Tchebycheff polynomial, a_1, a_2, a_3 , and the other two Tchebycheff points, z_2, z_3 , can be determined as follows. Solve the system in $[z_L, \infty)$:

$$W(z_L) = -1, W(z_i) = (-1)^i, W'(z_i) = 0, i = 2, 3.$$

If $z_2 < z_3 < z_U$, then the Tchebycheff points are $\{z_L, z_2, z_3\}$. Otherwise, they are $\{z_L, z_2, z_U\}$, where z_2 and a_i are the solutions of the system,

$$W(z_L) = -1, W(z_2) = 1, W(z_U) = -1, W'(z_2) = 0.$$

As stated in Studden (1968), Pukelsheim & Studden (1993), Imhof & Studden (2001) and Dette, Melas & Pepelyshev (2004), there are many vectors c for which the c -optimal designs are supported on these Tchebycheff points. The corresponding weights of such a c -optimal design can be calculated by (Dette, Melas & Pepelyshev, 2004)

$$w_i = \frac{|e_i^T J G^{-1} c|}{\sum_{j=1}^3 |e_j^T J G^{-1} c|}, \quad i = 1, 2, 3, \quad (10)$$

where $G(g_i(z_j))_{i,j=1}^3$, $J = \text{diag}(-1, 1, -1)$, and $e_j \in \mathbb{R}^3$ is the j^{th} standard basis vector. By

Lemma 2.3 in Dette, Melas & Pepelyshev (2004), for any $\mathcal{Z} = [z_L, z_U] \subset \mathbb{R}^+$, the design supported by Tchebycheff points with weights (10) is e_j -optimal if $\{g_i | i \in \{1, 2, 3\} \setminus \{j\}\}$ is a T -system on $\mathcal{Z}$. Now, since $\{z, z \log z\}$ and $\{1, z\}$ are T -systems on $\mathbb{R}^+$, such designs are e_1 - and e_3 -optimal separately. But since $\{1, z \log z\}$ is a T -system on $[e^{-1}, \infty)$ only, the design is e_2 -optimal when $z_L > e^{-1}$.

Optimal design is first calculated in $\mathcal{Z}$, and then transformed back to the original χ by $x = \phi^{-1}(z)$. Table 3 provides some examples of locally c -optimal designs for estimating θ_0, θ_1 or θ_2 separately. Note that $x_1 \equiv 1$, $w_1 = 1 - w_2 - w_3$, and that because the c -optimal designs are

supported on Tchebycheff points, the design points are the same for different c , while the weights depend on c . From Table 3 we observe that the weights are independent of θ_2 , if the upper bound is not a design point.

3. Optimal Designs for Weibull Model

Assume model (3). For locally D - and A -optimal designs, we cannot apply the Theory of Tchebycheff Systems to get results similar to those in the previous section. In addition, the left boundary of the design interval may not be one of the design points. However, the following theorem shows that we can restrict ourselves to minimally supported designs. The proof of the theorem uses the technique of Yang (2010).

Theorem 3. Under model (3), for any design ξ , there exists a design ξ^* with two support points such that $I(\theta, \xi) = I(\theta, \xi^*)$.

Proof. For an arbitrary design $\xi = \{(x_i, w_i), i = 1, \dots, m\}$, we can rewrite the Fisher information matrix $I(\theta, \xi)$ as (Yang, 2010),

$$I(\theta, \xi) = P(\theta) \left(\sum_{i=1}^m w_i C(\theta, z_i) \right) P^T(\theta), \quad (11)$$

with

$$P(\theta) = \begin{pmatrix} \frac{1}{\theta_1} & 0 \\ \log \theta_1 & \frac{1}{\theta_2} \end{pmatrix} \text{ and } C(\theta, z_i) = \begin{pmatrix} \frac{z_i^2 e^{-z_i}}{1 - e^{-z_i}} & \frac{z_i^2 e^{-z_i} \log z_i}{1 - e^{-z_i}} \\ \frac{z_i^2 e^{-z_i} \log z_i}{1 - e^{-z_i}} & \frac{z_i^2 e^{-z_i} \log^2 z_i}{1 - e^{-z_i}} \end{pmatrix},$$

where $z_i = \theta_1 x_i^{\theta_2} > 0$. Denote $C(\theta, \xi)$ as the summation term in ((11)).

Now let $\psi_1(z) = \int_0^z \frac{t^2 e^{-t}}{1 - e^{-t}} dt$, $\psi_2(z) = \frac{z^2 e^{-z}}{1 - e^{-z}}$, $\psi_3(z) = \frac{z^2 e^{-z} \log z}{1 - e^{-z}}$, $\psi_4(z) = \frac{z^2 e^{-z} \log^2 z}{1 - e^{-z}}$. Note that

$\psi_1(z)$ is chosen for computational convenience. For any $z > 0$, it has been verified

(Appendix C) that $\psi_1'(z) > 0$, $\left(\frac{\psi_2'(z)}{\psi_1'(z)} \right)' < 0$, and

$$\left(\left(\frac{\Psi'_3(z)}{\Psi'_1(z)} \right)' \left(\frac{\Psi'_2(z)}{\Psi'_1(z)} \right)' \right)' > 0,$$

$$\left(\left(\left(\frac{\Psi'_4(z)}{\Psi'_1(z)} \right)' \left(\frac{\Psi'_2(z)}{\Psi'_1(z)} \right)' \right)' \left(\left(\frac{\Psi'_3(z)}{\Psi'_1(z)} \right)' \left(\frac{\Psi'_2(z)}{\Psi'_1(z)} \right)' \right)' \right)' > 0.$$

These conditions, along with Theorems 1 and 2 of Yang (2010), indicate that there exists a design, say ξ^* , with two support points such that one diagonal element of $C(\theta, \xi^*)$ is larger than the corresponding element of $C(\theta, \xi)$, and the remaining elements remain unchanged. This proves the theorem.

Given Theorem 3, for any information matrix based optimality criterion, we can obtain the locally optimal design by investigating only two-point designs. Consequently, analytical solutions are available for some criteria.

3.1. Locally D -optimal designs

As in the proof of Theorem 3, after a transformation $z = \theta_1 x^{\theta_2}$, the information matrix can be expressed in terms of z . Thus, we obtain the optimal design in the induced z -design interval, $\mathcal{Z} = [z_L, z_U]$, and then transform it back to the design in $\mathcal{X}$.

Corollary 3.1—Let (z_1^*, z_2^*) be the unique solution of the equation system

$$2 \left(1 - \frac{1}{\log z_2 - \log z_1} \right) - \frac{z_1}{1 - e^{-z_1}} = 0, \quad (12)$$

$$2 \left(1 + \frac{1}{\log z_2 - \log z_1} \right) - \frac{z_2}{1 - e^{-z_2}} = 0. \quad (13)$$

Then, under model (3), the locally D -optimal design, ξ_D , for θ is determined as follows.

1. If $z_L \leq z_1^* < z_2^* \leq z_U$, then ξ_D puts equal weight on support points x_1^* and x_2^* , where

$$x_i^* = \left(\frac{z_i^*}{\theta_1} \right)^{\frac{1}{\theta_2}}, \quad i = 1, 2.$$

2. If $z_L > z_1^*$, then ξ_D puts equal weight on support points x_L and $\min(x_2^*, x_U)$, where

$$x_2^* = \left(\frac{z_2}{\theta_1} \right)^{\frac{1}{\theta_2}} \text{ with } z_2 \text{ being the solution to}$$

$$2 \left(1 + \frac{1}{\log z_2 - \log z_L} \right) - \frac{z_2}{1 - e^{-z_2}} = 0.$$

3. If $z_U < z_2^*$, then ξ_D puts equal weight on support points on $\max(x_1^*, x_L)$ and x_U ,

where $x_1^* = \left(\frac{z_1}{\theta_1} \right)^{\frac{1}{\theta_2}}$ with z_1 being the solution to

$$2 \left(1 - \frac{1}{\log z_U - \log z_1} \right) - \frac{z_1}{1 - e^{-z_1}} = 0.$$

Proof. By Theorem 3, a locally D -optimal design which maximizes the determinant of $I(\theta, \xi)$ has two support points. Thus, it has equal weight on the support points. Let $z_1 < z_2$ be the two support points in $\mathcal{Z}$. Then ξ_D maximizes $\log h(z_1, z_2)$, where

$$h(z_1, z_2) = \frac{z_1^2 z_2^2 e^{-z_1 - z_2}}{(1 - e^{-z_1})(1 - e^{-z_2})} (\log z_2 - \log z_1)^2.$$

Equating the derivatives of $\log h(z_1, z_2)$, with respect to z_1 and z_2 separately, to zero, we obtain equations (12) and (13), respectively.

The statements in the theorem are true since given z_2 , the left hand side of ((12) as a function of z_1 is monotone decreasing with opposite signs at 0 and z_2 , and, given z_1 , the left hand side of ((13) as a function of z_2 is monotone decreasing with opposite signs at z_1 and ∞ .

Note that numerically we can solve ((13) and express z_2 as a function of z_1 , and substitute it into ((12) to get an equation involving z_1 only. Actually, $z_1^* \approx 0.2624$ and $z_2^* \approx 2.6635$. In Table 4, we present some examples of x_1^* and x_2^* for pre-selected values of θ_1 and θ_2 . It can be seen that given θ_2 , as θ_1 increases, both x_1^* and x_2^* tend to become smaller, and that given θ_1 , if θ_2 increases, x_1^* will increase but x_2^* will decrease. If the design interval is $\mathcal{R}^+$, these are support points of the locally D -optimal designs. Otherwise, one needs to compare z_1^* , z_2^* and the boundaries of $\mathcal{Z}$, as stated in Corollary 3.1, to determine the design points.

3.2. Locally A-optimal designs

From Theorem 3, the locally A -optimal design has two support points, say x_1, x_2 , and thus the corresponding weights can be determined as described in (9) with

$$\frac{f^T(x, \theta)}{\sqrt{P(x, \theta)(1 - P(x, \theta))}} = \sqrt{\frac{e^{-\theta_1 x^{\theta_2}}}{1 - e^{-\theta_1 x^{\theta_2}}}} (x^{\theta_2}, \theta_1 x^{\theta_2} \log x)^T. \quad (14)$$

Thus, it is easy to construct a locally A -optimal design by minimizing a two-dimensional function, $tr(F^{-1}(\theta, \xi))$. Table 5 provides the optimal designs for $\chi = [0.1, 10]$ for some pre-selected values of θ . We choose a wide interval to demonstrate that the endpoints of the design interval may or may not be the support points. As θ_1 or θ_2 changes, the design points change in the same way as they do in locally D -optimal designs.

3.3. Locally c -optimal designs

Obviously, the element functions in the vector ((14) form a T -system in χ since

$$\sqrt{\frac{e^{-\theta_1 x^{\theta_2}}}{1 - e^{-\theta_1 x^{\theta_2}}}} > 0 \text{ and } \{x^{\theta_2}, \theta_1 x^{\theta_2} \log x\} \text{ is a } T\text{-system. Thus, we can find the Tchebycheff}$$

points, x_1, x_2 , using the method described in Section 2.3. Then by Lemma 2.3 in Dette, Melas & Pepelyshev (2004), the design with support points $\{x_1, x_2\}$, and the corresponding weights calculated similarly as in (10), is c -optimal if $c = (0, 1)^T$ for any χ , and c -optimal if $c = (1, 0)^T$ when χ does not include 1.

4. Optimal Designs for Beta-Poisson Model

In this section, we discuss locally D -, A -, and c -optimal designs for model (4). Similarly, we only consider designs with at least two support points in order to have an invertible information matrix. Like in Section 2, the theory of T -systems is used to determine the number of support points.

4.1. Locally D -optimal designs

Theorem 4. For the Beta-Poisson model (4), the locally D -optimal design in any closed positive interval X puts equal weight on two support points.

Proof. Let $z = \phi(x) = \log\left(1 + \frac{x}{\theta_2}\right)$ and use the same notation as in Theorem 1. Denote ξ_D as the locally D -optimal design and m_{ij} as the $(i, j)^{th}$ element of $F^{-1}(\xi_D, \theta)$. Then, with

$$\frac{f^T(z, \theta)}{\sqrt{P(z, \theta)(1 - P(z, \theta))}} = \sqrt{\frac{e^{-\theta_1 z}}{1 - e^{-\theta_1 z}}} \left(z, -\frac{\theta_1}{\theta_2}(1 - e^{-z}) \right), \quad (15)$$

we have the variance function

$$d(z, \xi_D, \theta) = \frac{e^{-\theta_1 z}}{1 - e^{-\theta_1 z}} \left(m_{11} z^2 - 2m_{12} \frac{\theta_1}{\theta_2} z(1 - e^{-z}) + m_{22} \frac{\theta_1^2}{\theta_2^2} (1 - e^{-z})^2 \right).$$

Now, let $H_D(z) = (1 - e^{-\theta_1 z})[d(z, \xi_D, \theta) - 2]$ and $u(z) = m_{11} z^2 e^{-\theta_1 z} + m_{22} \frac{\theta_1^2}{\theta_2^2} e^{-2z - \theta_1 z}$, then

we claim that in $[0, \infty)$, $H_D(z)$ is a linear combination of T -system

$$\{-1, ze^{-\theta_1 z}, e^{-\theta_1 z}, ze^{-z - \theta_1 z}, e^{-z - \theta_1 z}, u(z)\}. \quad (16)$$

Hence, $H_D(z)$ has at most five zeros in $[0, \infty)$ (non-nodal zero counted twice). Since $H_D(0) = 0$, $H_D(z)$ has at most four zeros in $\mathbb{R}^+$. The strategy of determining the number of supports is similar to that of model (2). First, $H_D(z)$ should have at least two distinct zeros in $\mathcal{Z}$ so that the information matrix is invertible. Second, assume that $H_D(z)$ has more than two distinct zeros in $\mathcal{Z}$. This, along with the fact that $H_D(z) = 0$, $z \in \mathcal{Z}$, implies that there is one non-nodal zero inside $\mathcal{Z}$. Then $H_D(z)$ has both boundaries as nodal zeros, and thus has no zeros outside of $\mathcal{Z}$. This contradicts the fact that $\lim_{z \rightarrow \infty} H_D(z) = -2 < 0$. Hence $H_D(z)$ has two distinct zeros in $\mathcal{Z}$. That is, the locally D -optimal design has two support points.

The claim can be proven as follows. For any coefficients $\{a_i\}_{i=0}^5$, the linear combination of the system

$$\{-1, ze^{-\theta_1 z}, e^{-\theta_1 z}, ze^{-z - \theta_1 z}, e^{-z - \theta_1 z}, z^2 e^{-\theta_1 z}\}, \quad (17)$$

is actually an exponential polynomial, $a_0 + (a_1 + a_2 z + a_3 z^2)e^{-\theta_1 z} + (a_4 + a_5 z)e^{-(\theta_1 + 1)z}$, which has at most five zeros (p. 10, Karlin & Studden, 1966). Hence, the sequence of functions in ((17)) is a T -system. Similarly, the sequence of functions in

$$\{-1, ze^{-\theta_1 z}, e^{-\theta_1 z}, ze^{-z - \theta_1 z}, e^{-z - \theta_1 z}, e^{-2z - \theta_1 z}\},$$

is another T -system. Since $u(z)$ is a positively linear combination of $z^2 e^{-\theta_1 z}$ and $e^{-2z - \theta_1 z}$, the sequence of functions in ((16)) is a T -system, by Lemma 3.2 in Li & Balakrishnan (2011). This proves the claim and the theorem.

Constraining ourselves with two-point designs with equal weight, we find the supports of the locally D -optimal design by maximizing the determinant of the information matrix. In terms of z , setting the derivatives of the log of the determinant equal to zero, we first seek the solution (z_1^*, z_2^*) of the following equations,

$$-\frac{\theta_1}{1-e^{-\theta_1 z_1}} + 2 \frac{z_2 e^{-z_1} - (1-e^{-z_2})}{z_2(1-e^{-z_1}) - z_1(1-e^{-z_2})} = 0,$$

$$-\frac{\theta_1}{1-e^{-\theta_1 z_2}} + 2 \frac{(1-e^{-z_1}) - z_1 e^{-z_2}}{z_2(1-e^{-z_1}) - z_1(1-e^{-z_2})} = 0,$$

and then determine the design points as discussed in Corollary 3.1. Finally the design points in χ were obtained via $x = \phi^{-1}(z)$. For some initial values of θ_1 and θ_2 , Table 6 presents the support points of locally D -optimal designs in the design interval, $\chi = [0.2, 500]$, which is selected to show that the endpoints of χ may or may not be the design points. The table demonstrates that, unlike the Weibull model, the design points become larger as θ_2 increases, and become smaller as θ_1 increases.

4.2. Locally A -optimal designs

Similar to D -optimality, we can show that the locally A -optimal design for the Beta-Poisson model is minimally supported by using the Kiefer-Wolfowitz's Equivalence Theorem and the Theory of T -systems.

Theorem 5. For model (4) and design interval χ , the locally A -optimal design is supported on two points.

Proof. Let $z = \log\left(1 + \frac{x}{\theta_2}\right)$. Thus the information matrix of a single observation is given by,

$$\mu(z, \theta) = \frac{e^{-\theta_1 z}}{1-e^{-\theta_1 z}} \begin{pmatrix} z^2 & -\frac{\theta_1}{\theta_2} z(1-e^{-z}) \\ -\frac{\theta_1}{\theta_2} z(1-e^{-z}) & \frac{\theta_1^2}{\theta_2^2} (1-e^{-z})^2 \end{pmatrix}.$$

The if and only if condition (8) of a design ξ_A being locally A -optimal can then be rewritten as

$$e^{-\theta_1 z} \operatorname{tr} \left(\begin{pmatrix} z^2 & -\frac{\theta_1}{\theta_2} z(1-e^{-z}) \\ -\frac{\theta_1}{\theta_2} z(1-e^{-z}) & \frac{\theta_1^2}{\theta_2^2} (1-e^{-z})^2 \end{pmatrix} I^{-2}(\xi_A, \theta) \right) - (1-e^{-\theta_1 z}) \operatorname{tr}(I^{-1}(\xi_A, \theta)) \leq 0.$$

Now since the diagonal elements of $I^2(\xi_A, \theta)$ are positive and $\operatorname{tr}(I^{-1}(\xi_A, \theta)) > 0$, by some matrix operations, we simplify and denote the left-hand side of the above equation as

$$H_A(z) = -k_1 + k_2 z^2 e^{-\theta_1 z} + k_3 z e^{-\theta_1 z} + k_4 e^{-\theta_1 z} + k_3 z e^{-z-\theta_1 z} - 2k_5 e^{-z-\theta_1 z} + k_5 e^{-2z-\theta_1 z},$$

with $H_A(0) = 0$, and $k_1, k_2, k_4, k_5 > 0$. Let $u(z) = k_2 z^2 e^{-\theta_1 z} + k_5 e^{-2z-\theta_1 z}$, by the same argument as that in the proof of the claim in Theorem 4, we then have that $H_A(z)$ is a linear combination of the T -system

$$\{-1, z e^{-\theta_1 z}, e^{-\theta_1 z}, z e^{-z-\theta_1 z}, e^{-z-\theta_1 z}, u(z)\}.$$

Hence, $H_A(z)$ has at most five zeros in $[0, \infty)$. Following the same arguments (after equation (16)) in the proof of Theorem 4, we conclude that ξ_A has two support points.

Given two design points, x_1, x_2 , and the corresponding weights, calculated similarly to (9), we carry out a two-dimension numerical search for the locally A -optimal design points. Table 7 presents some locally A -optimal designs for model (4). Again, the design interval $[0.2, 500]$ is chosen to demonstrate that the endpoints may or may not be the design points. The design points change in the same way relative to θ_1 and θ_2 as they do for locally D -optimal designs.

4.3. Locally c -optimal designs

For model (4), in terms of $z = \log(1 + \frac{x}{\theta_2})$, the element functions in the vector ((15)

comprise a T -system since $\sqrt{\frac{e^{-\theta_1 z}}{1 - e^{-\theta_1 z}}} > 0$ and $\{z, -\frac{\theta_1}{\theta_2}(1 - e^{-z})\}$ is a T -system. The

Tchebycheff points $\{z_1, z_2\}$ are unique and determined using the method described in Section 2.3. Thus, the design supported on $\{z_1, z_2\}$ with weights calculated similarly to ((10) is c -optimal for any $\mathcal{L}$ if $c = (0, 1)^T$, or $c = (1, 0)^T$. Applying $x = \phi^{-1}(z)$, we have c -optimal designs in $\mathcal{X}$.

5. Discussion

In this paper we have shown that locally D -, A - and c -optimal designs for three dose-response models are minimally supported. Methods and examples are provided for constructing these locally optimal designs. These designs depend on the initial guess of the unknown parameters. We investigate the efficiency of such designs, when the initial values are different from the true values of the parameters, using efficiency indexes.

The D -, A - and c -efficiency of design ξ_1 relative to design ξ_2 , are given by (see, for example, Wong, 1994),

$$\text{eff}_D = \left(\frac{|I^{-1}(\xi_2, \theta)|}{|I^{-1}(\xi_1, \theta)|} \right)^{\frac{1}{p}}, \text{eff}_A = \frac{\text{tr}(I^{-1}(\xi_2, \theta))}{\text{tr}(I^{-1}(\xi_1, \theta))}, \text{eff}_c = \frac{c^T I^{-1}(\xi_2, \theta) c}{c^T I^{-1}(\xi_1, \theta) c},$$

where p is the number of model parameters. To assess the efficiency of optimal design under parameter misspecification, ξ_1 is the optimal design obtained from misspecified initial value of θ , while ξ_2 is the optimal design obtained from the true θ .

For illustration, we present the D -efficiencies of the locally D -optimal designs for the three models. For simplicity, we use $\chi = \mathcal{R}^+$ as the design space. Table 8 shows the dependency of D -efficiency on θ_0 in model (2), with θ_0 varying from 0.01 to 100, and the true parameters being $\theta_0 = \theta_1 = \theta_2 = 1$. It is clear that the D -efficiency is extremely high when θ_0 is misspecified. We skip the discussion of θ_1 and θ_2 for this model since the misspecification of these parameters has a similar effect to that of model (3). Tables 9 and 10 present the D -efficiencies with different, misspecified initial values of (θ_1, θ_2) for the two-parameter models (3) and (4), respectively, with $\theta_1 = \theta_2 = 1$ being the true value. We can see that for model (3), the efficiency declines more slowly when θ_1 and θ_2 are misspecified in opposite directions than in the same direction, and the loss of D -efficiency is dramatic if both θ_1 and θ_2 are underestimated. One partial explanation is that when θ_1 and θ_2 depart from the true values in the same direction, the difference in the probabilities of success, $P(x, \theta)$, is bigger than when they depart in opposite directions. The reverse happens for model (4).

As expected, the tables show that the loss in efficiency is greater when the estimates of the parameters deviate more from the true values. Should we know the possible range of the parameter θ and a prior distribution $\delta(\theta)$ becomes available, Bayesian optimal designs may improve efficiency. Other approaches for safeguarding against parameter misspecification include those based on minimax and sequential principles. Minimax designs minimize the maximum, over the parameter space, of the loss function of the design criterion. Sequential designs are proposed by Box & Lucas (1959), in which an initial design is chosen to obtain preliminary parameter estimates, and then design points are added to provide the biggest improvement with respect to the criterion and current parameter estimates. The process is repeated until all resources are used. Constructing these parameter robust designs for the dose-response models will be a topic of future research.

Acknowledgments

The research of Zhide Fang is supported by grants from the National Institute of General Medical Sciences of the National Institutes of Health, USA, which fund the Louisiana Clinical and Translational Science Center of Pennington Biomedical Research Center. The authors thank the Editor, Associate Editor, and two anonymous reviewers for their careful reading of our manuscript and their comments and suggestions.

Appendix

A. Proof that the system in (7) is a T -system

Definitions of T -system, ET -system and ECT -system can be found in Karlin & Studden (1966). Obviously, an ECT -system is an ET -system, and thus a T -system. Theorem 1.1 in Chapter XI of Karlin & Studden (1966) states that a sequence of functions, $u_0, u_1, \dots, u_m$ is an ECT -system on interval χ if and only if the Wronskian of functions $\{u_0\}_{i=0}^k$ is positive for any $k = 0, 1, \dots, n$. That is,

$$W(u_0, u_1, \dots, u_k)(z) = \begin{vmatrix} u_0(z) & u_0^{(1)}(z) & \dots & u_0^{(k)}(z) \\ u_1(z) & u_1^{(1)}(z) & \dots & u_1^{(k)}(z) \\ \vdots & \vdots & \ddots & \vdots \\ u_k(z) & u_k^{(1)}(z) & \dots & u_k^{(k)}(z) \end{vmatrix} > 0$$

where $z \in \chi$.

Now for the functions in (7) and $\chi \subset \mathbb{R}^+$, we have

$$\begin{aligned} W(u_0)(z) &= 1 > 0, \\ W(u_0, u_1)(z) &= 1 > 0, \\ W(u_0, u_1, u_2)(z) &= 2 > 0, \\ W(u_0, u_1, \dots, u_3)(z) &= \frac{4}{z} > 0, \\ W(u_0, u_1, \dots, u_4)(z) &= \frac{4}{z^4} > 0, \\ W(u_0, u_1, \dots, u_5)(z) &= \frac{16}{z^7} > 0, \\ W(u_0, u_1, \dots, u_6)(z) &= \frac{16e^z(z^3 + 7z^2 + 10z + 2)}{z^{10}} > 0. \end{aligned}$$

Thus, the sequence of functions in (7) is an ECT -system.

B. Proof that $\delta(z_1, z_2, z_3)$ is decreasing in z_1

Given z_2, z_3 , we have

$$\frac{\partial \delta(z_1, z_2, z_3)}{\partial z_1} = -1 - \frac{e^{-\theta_0 - z_1}}{1 - e^{-\theta_0 - z_1}} - 2 \frac{z_3 \left(\log \frac{z_3}{z_1} - 1 \right) - z_2 \left(\log \frac{z_2}{z_1} - 1 \right)}{z_2 z_3 \log \frac{z_3}{z_2} - z_1 z_3 \log \frac{z_3}{z_1} + z_1 z_2 \log \frac{z_2}{z_1}} < 0,$$

for $z_1 \in [z_L, z_2]$, because both the numerator and denominator of the ratio in the partial derivative are positive. Hence, $\delta(z_1, z_2, z_3)$ is a decreasing function of $z_1 \in [z_L, z_2]$.

C. Details of the Proof of Theorem 3

With the notation in the proof of Theorem 3, we have

$$\begin{aligned}\psi'_1(z) &= \frac{z^2 e^{-z}}{1 - e^{-z}} > 0, \\ \psi'_2(z) &= \frac{(2 - z - 2e^{-z})ze^{-z}}{(1 - e^{-z})^2}.\end{aligned}$$

Thus,

$$\left(\frac{\psi'_2(z)}{\psi'_1(z)}\right)' = \left(\frac{2 - z - 2e^{-z}}{z(1 - e^{-z})}\right)' = -\frac{2}{z^2} + \frac{e^{-z}}{(1 - e^{-z})^2} < 0.$$

Now, since $\psi'_3(z) = \frac{(2 - z - 2e^{-z})ze^{-z}}{(1 - e^{-z})^2} \log z + \frac{ze^{-z}}{1 - e^{-z}}$, and

$$\left(\frac{\psi'_3(z)}{\psi'_1(z)}\right)' = \left(\frac{(2 - z - 2e^{-z}) \log z}{z(1 - e^{-z})} + \frac{1}{z}\right)' = \left(-\frac{2}{z^2} + \frac{e^{-z}}{(1 - e^{-z})^2}\right) \log z + \frac{2 - z - 2e^{-z}}{z^2(1 - e^{-z})} - \frac{1}{z^2},$$

we have

$$\begin{aligned}& \left(\left(\frac{\psi'_3(z)}{\psi'_1(z)}\right)'\right) \left(\left(\frac{\psi'_2(z)}{\psi'_1(z)}\right)'\right) \\ &= \left(\log z + \frac{(1 - e^{-z})(1 - e^{-z} - z)}{-2(1 - e^{-z})^2 + z^2 e^{-z}}\right)' = \frac{1}{z} + \frac{n(z)}{(-2(1 - e^{-z})^2 + z^2 e^{-z})^2} > 0,\end{aligned}$$

where $n(z) = 6e^{-2z} - 6e^{-z} - 2e^{-3z} + 2z^2 e^{-z} - z^2 e^{-2z} - z^3 e^{-z} - z^2 e^{-3z} - 4ze^{-z} + 8ze^{-2z} - 4ze^{-3z} + 2 > 0$ for all $z > 0$.

The expression for

$$\left(\left(\left(\frac{\psi'_4(z)}{\psi'_1(z)}\right)'\right) \left(\left(\frac{\psi'_2(z)}{\psi'_1(z)}\right)'\right)\right) \left(\left(\left(\frac{\psi'_3(z)}{\psi'_1(z)}\right)'\right) \left(\left(\frac{\psi'_2(z)}{\psi'_1(z)}\right)'\right)\right)'$$

is too long to present. We have verified that it is greater than 0 for all $z > 0$ using symbolic computational software.

Bibliography

- Armitage P, Doll R. The age distribution of cancer and a multi-stage theory of carcinogenesis. *British Journal of Cancer*. 1954; 8(1):1–12. [PubMed: 13172380]
- Biedermann S, Dette H, Zhu W. Optimal designs for doseresponse models with restricted design spaces. *Journal of the American Statistical Association*. 2006; 101(474):747–759.
- Box GEP, Lucas HL. Design of experiments in non-linear situations. *Biometrika*. 1959; 46(1/2):77–90.
- Buchanan RL, Smith JL, Long W. Microbial risk assessment: Dose-response relations and risk characterization. *International Journal of Food Microbiology*. 2000; 58(3):159–172. [PubMed: 10939266]
- Chernoff H. Locally optimal designs for estimating parameters. *The Annals of Mathematical Statistics*. 1953; 24(4):586–602.
- Dette H, Haines LM. E-optimal designs for linear and nonlinear models with two parameters. *Biometrika*. 1994; 81(4):739–754.
- Dette H, Melas VB, Pepelyshev A. Optimal designs for a class of nonlinear regression models. *The Annals of Statistics*. 2004; 32(5):2142–2167.
- Fang Z, Wiens DP, Wu Z. Locally D-optimal designs for multistage models and heteroscedastic polynomial regression models. *Journal of Statistical Planning and Inference*. 2006; 136(11):4059–4070.
- Fedorov VV. *Theory of Optimal Experiments*. Academic Press; New York: 1972.
- Fedorov VV, Leonov SL. Optimal design of dose response experiments: A model-oriented approach. *Drug Information Journal*. 2001; 35(4):1373–1383.
- Fedorov VV, Leonov SL. *Optimal Design for Nonlinear Response Models*. CRC Press; Boca Raton, Florida, USA: 2014.
- Ford I, Titterton DM, Kitsos CP. Recent advances in nonlinear experimental design. *Technometrics*. 1989; 31(1):49–60.
- Ford I, Torsney B, Wu CJ. The use of a canonical form in the construction of locally optimal designs for non-linear problems. *Journal of the Royal Statistical Society Series B (Methodological)*. 1992; 54(2):569–583.
- Haas CN. Estimation of risk due to low doses of microorganisms: a comparison of alternative methodologies. *American Journal of Epidemiology*. 1983; 118(4):573–582. [PubMed: 6637984]
- Imhof LA, Studden WJ. E-optimal designs for rational models. *The Annals of Statistics*. 2001; 29(3):763–783.
- Karlin S, Studden WJ. *Tchebycheff Systems: With Applications in Analysis and Statistics*. John Wiley & Sons, Inc; New York: 1966.
- Krewski D, Van Ryzin J. Dose response models for quantal response toxicity data. In: Csorgo M, Dawson DA, Rao JNK, Saleh AK, editors *Statistics and Related Topics*. Amsterdam: North-Holland; 1981. 201–231.
- Li G, Balakrishnan N. Optimal designs for tumor regrowth models. *Journal of Statistical Planning and Inference*. 2011; 141(2):644–654.
- Portier C, Hoel D. Low-dose-rate extrapolation using the multistage model. *Biometrics*. 1983; 39(4):897–906. [PubMed: 6671125]
- Portier CJ, Edler L. Two-stage models of carcinogenesis, classification of agents, and design of experiments. *Fundamental and Applied Toxicology*. 1990; 14(3):444–460. [PubMed: 2340975]
- Pukelsheim F. *Optimal Design of Experiments*. Wiley; New York: 1993.
- Pukelsheim F, Studden WJ. E-optimal designs for polynomial regression. *The Annals of Statistics*. 1993; 21(1):402–415.
- Silvey SD. *Optimal Design: an Introduction to the Theory for Parameter Estimation*. Chapman and Hall; London: 1980.

- Sitter RR, Fainaru I. Optimal designs for the logit and probit models for binary data. *The Canadian Journal of Statistics*. 1997; 25(2):175–190.
- Sitter RR, Wu CFJ. Optimal designs for binary response experiments: Fieller, D, and A criteria. *Scandinavian Journal of Statistics*. 1993; 20(4):329–341.
- Studden WJ. Optimal designs on Tchebycheff points. *The Annals of Mathematical Statistics*. 1968; 39(5):1435–1447.
- Stufken J, Yang M. On locally optimal designs for generalized linear models with group effects. *Statistica Sinica*. 2012; 22(4):1765–1786.
- Wong WK. Comparing robust properties of A, D, E and G-optimal designs. *Computational Statistics & Data Analysis*. 1994; 18(4):441–448.
- Yang M. On the de la Garza phenomenon. *The Annals of Statistics*. 2010; 38(4):2499–2524.
- Yang M, Stufken J. Support points of locally optimal designs for nonlinear models with two parameters. *The Annals of Statistics*. 2009; 37(1):518–541.

Table 1

Locally D -optimal designs and corresponding $\log |\mathcal{I}(\boldsymbol{\theta}, \xi_D)|$ when $x \in [1, 4]$.

θ_0	θ_1	θ_2	x_1	x_2	x_3	$\log \mathcal{I}(\boldsymbol{\theta}, \xi_D) $
0.1	1	1	1	1.820	4	-9.346
0.1	1	2	1	1.388	2.204	-10.600
0.1	2	1	1	1.524	3.048	-14.709
0.1	2	2	1	1.235	1.746	-16.095
10	1	1	1	1.876	4	-39.621
10	1	2	1	1.415	2.221	-40.847
10	2	1	1	1.537	3.060	-44.585
10	2	2	1	1.240	1.749	-45.971

Table 2

Locally A -optimal designs and corresponding $\text{tr}(F^{-1}(\theta, \xi_A))$ when $x \in [1, 4]$.

θ_0	θ_1	θ_2	$\{x_1, x_2, x_3\}$	$\{w_1, w_2, w_3\}$	$\text{tr}(F^{-1}(\theta, \xi_A))$
0.1	1	1	{1, 1.787, 4}	{0.249, 0.457, 0.295}	1208.205
0.1	1	2	{1, 1.375, 2.448}	{0.209, 0.371, 0.420}	1206.652
0.1	2	1	{1, 1.504, 3.639}	{0.203, 0.370, 0.427}	12773.38
0.1	2	2	{1, 1.227, 1.907}	{0.196, 0.363, 0.440}	14785.71
1	1	1	{1, 1.822, 4}	{0.259, 0.452, 0.288}	3329.526
1	1	2	{1, 1.392, 2.460}	{0.218, 0.366, 0.416}	3241.368
1	2	1	{1, 1.511, 3.648}	{0.206, 0.369, 0.426}	32260.11
1	2	2	{1, 1.230, 1.909}	{0.199, 0.362, 0.439}	37314.78

Table 3

c-optimal designs in $\chi = [1, 4]$ for estimating individual parameters.

θ_0	θ_1	θ_2	$\{x_2, x_3\}$	$c = (1, 0, 0)^T$ $\{w_2, w_3\}$	$c = (0, 1, 0)^T$ $\{w_2, w_3\}$	$c = (0, 0, 1)^T$ $\{w_2, w_3\}$
0.1	1	1	{1.787, 4}	{0.448, 0.273}	{0.477, 0.291}	{0.436, 0.372}
0.1	1	2	{1.374, 2.454}	{0.380, 0.358}	{0.409, 0.385}	{0.337, 0.506}
0.1	2	1	{1.503, 3.642}	{0.367, 0.415}	{0.380, 0.429}	{0.320, 0.530}
0.1	2	2	{1.226, 1.908}	{0.367, 0.415}	{0.380, 0.429}	{0.320, 0.530}
1	1	1	{1.823, 4}	{0.443, 0.267}	{0.473, 0.285}	{0.435, 0.364}
1	1	2	{1.391, 2.466}	{0.373, 0.354}	{0.403, 0.383}	{0.334, 0.502}
1	2	1	{1.511, 3.651}	{0.365, 0.413}	{0.378, 0.428}	{0.320, 0.528}
1	2	2	{1.229, 1.911}	{0.365, 0.413}	{0.378, 0.428}	{0.320, 0.528}

Table 4

Some examples of x_1^*, x_2^* .

θ_1	θ_2	x_1^*	x_2^*	θ_1	θ_2	x_1^*	x_2^*
0.5	0.5	0.2755	28.3765	1	0.5	0.0689	7.0941
0.5	1	0.5249	5.3270	1	1	0.2624	2.6635
0.5	2	0.7245	2.3080	1	2	0.5123	1.6320
				2	0.5	0.0172	1.7735
				2	1	0.1312	1.3317
				2	2	0.3622	1.1540

Table 5

Locally A -optimal designs when $\chi = [0.1, 10]$.

θ_1	θ_2	$\mathbf{x}$	$\mathbf{x}_2$	$\mathbf{w}_1$	$\mathbf{w}_2$	$\text{tr}(\mathbf{I}^{-1}(\boldsymbol{\theta}, \boldsymbol{\xi}_A))$
0.5	0.5	0.1386	10	0.6796	0.3204	1.623895
0.5	1	0.3880	6.0595	0.6645	0.3355	3.089195
0.5	2	0.5479	2.5863	0.6451	0.3549	8.790303
1	0.5	0.1623	4.6038	0.5457	0.4543	3.056249
1	1	0.2825	2.5738	0.5608	0.4392	5.045563
1	2	0.4426	1.7376	0.5900	0.4100	11.30739
2	0.5	0.1	1.5270	0.2366	0.7634	9.101445
2	1	0.1504	1.2492	0.2915	0.7085	12.20599
2	2	0.3449	1.1795	0.4107	0.5893	20.63385

Table 6Locally D -optimal designs when $\chi = [0.2, 500]$.

θ_1	θ_2	x_1	x_2	$\log I(\theta, \xi_D) $
0.5	0.5	0.4801	73.8930	-1.730486
0.5	1	0.9595	147.9557	-3.116781
0.5	2	1.9196	295.6540	-4.503075
1	0.5	0.3087	9.2405	-3.213589
1	1	0.6177	18.4877	-4.599883
1	2	1.2348	36.9652	-5.986177
2	0.5	0.2	2.3477	-5.052318
2	1	0.3626	4.5950	-6.434784
2	2	0.7252	9.1883	-7.821079

Table 7

Locally A -optimal designs when $\chi = [0.2, 500]$.

θ_1	θ_2	x_1	x_2	w_1	w_2	$\text{tr}(I^{-1}(\theta, \xi_1))$
0.5	0.5	0.3045	213.2470	0.5412	0.4588	10.44669
0.5	1	0.5890	463.3173	0.5817	0.4183	35.52225
0.5	2	1.1565	500	0.6096	0.3904	136.953
1	0.5	0.2007	17.2813	0.3972	0.6028	21.76941
1	1	0.3940	35.5361	0.4736	0.5264	50.67058
1	2	0.7699	73.677	0.5142	0.4858	163.0536
2	0.5	0.2	3.7142	0.2882	0.7118	125.5789
2	1	0.2342	7.0989	0.3614	0.6386	165.1839
2	2	0.4663	14.2444	0.4186	0.5814	339.9664

Table 8

D -efficiencies when θ_0 deviates from the true value 1 for $\theta_1 = \theta_2 = 1$.

θ_0	Locally D -optimal design points	D -efficiency
0.01	{1, 1.918023, 4.847935}	0.9989300
0.1	{1, 1.926270, 4.855829}	0.9992295
0.5	{1, 1.953572, 4.883147}	0.9998669
1	{1, 1.973414, 4.903081}	1
2	{1, 1.991445, 4.921852}	0.9998903
10	{1, 2.001427, 4.932420}	0.9997379
100	{1, 2.002044, 4.930684}	0.9997282

Table 9

D -efficiencies when θ_1 and θ_2 deviate from the true values for model (3).

		θ_2				
		0.5	0.75	1	1.5	2
θ_1	0.5	0.0001	0.1976	0.6732	0.8821	0.7714
	0.75	0.0447	0.6436	0.9446	0.9044	0.7520
	1	0.3018	0.8850	1.0000	0.8747	0.7214
	1.5	0.6621	0.9095	0.9186	0.7913	0.6628
	2	0.5999	0.7785	0.8003	0.7163	0.6148

Table 10

D -efficiencies when θ_1 and θ_2 deviate from the true values for model (4).

		θ_2				
		0.5	0.75	1	1.5	2
θ_1	0.5	0.8482	0.8022	0.7459	0.6425	0.5599
	0.75	0.9605	0.9783	0.9544	0.8737	0.7911
	1	0.9034	0.9827	1.0000	0.9662	0.9053
	1.5	0.6944	0.8419	0.9193	0.9726	0.9657
	2	0.5178	0.6827	0.7884	0.8975	0.9348