Ann Inst Stat Math (2012) 64:945–957
DOI 10.1007/s10463-011-0340-y

Optimal and efficient designs for Gompertz regression models

Gang Li

Received: 13 July 2010 / Revised: 11 August 2011 / Published online: 20 November 2011

Abstract Gompertz functions have been widely used in characterizing biological growth curves. In this paper we consider D-optimal designs for Gompertz regression models. For homoscedastic Gompertz regression models with two or three parameters, we prove that D-optimal designs are minimally supported. Considering that minimally supported designs might not be applicable in practice, alternative designs are proposed. Using the D-optimal designs as benchmark designs, these alternative designs are found to be efficient in general.

Keywords D-optimality · Local optimality · Minimally supported designs · Sigmoid growth curve · Tchebycheff system

1 Introduction

Sigmoid growth curves are found in a wide range of disciplines, such as agriculture, biology, and microbiology. Sigmoid growth consists of three distinct phases: the initial exponential phase, the linear phase and the final plateau. Among other sigmoid functions, the Gompertz function has gained wide acceptance as an applicable function in a number of biological systems.

The Gompertz function (Gompertz 1825), $\mu_1(x) = \beta e^{-e^{-\gamma(x-\tau)}}$, has three parameters, where β is the upper asymptote, γ is the growth rate. The parameter τ is the point of inflection at which point the maximum growth rate, $\beta\gamma/e$, occurs; see Fig. 1. Compared with other sigmoid models such as logistic models, the Gompertz function provides a better approximation to tumor growth curves (Laird 1965; Marusic and Vuk-Pavlovic 1993). When the upper asymptote β is known, the Gompertz function can be simplified to a two-parameter function $\mu_2(x) = e^{-e^{-\gamma(x-\tau)}}$.

G. Li (✉)
GlaxoSmithKline Pharmaceuticals, 1250 S. Collegeville Road, Collegeville, PA 19426, USA
e-mail: gang.6.li@gsk.com

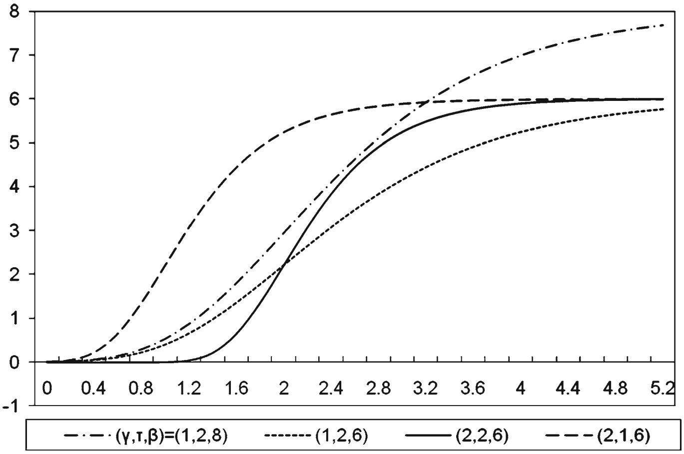

Fig. 1 Plot of Gompertz function

The choice of the experimental design is very important in order to accurately estimate the unknown model parameters and efficiently improve the quality of statistical inferences. The methodology based on the design of experiments is a useful tool that could be employed for choosing the best experimental design. Consider a vector of observations Y whose joint probability density function, $p(Y|x, \theta)$, depends on a vector of unknown parameters θ and the design variable x, whose values can be controlled by researchers in the design stage of the experiments. Consider a design region χ and let $\mathcal{H}$ be the set of all probability measures on χ. Let $M(x, \theta)$ be the information matrix of a single observation at point x, $M(x, \theta) = E\left[\frac{\partial \log p(Y|x,\theta)}{\partial \theta} \frac{\partial \log p(Y|x,\theta)}{\partial \theta^T}\right]$, where the expectation is taken with respect to the distribution of Y. The per observation information matrix for the design measure $\xi \in \mathcal{H}$ is $\mathcal{M}(\xi, \theta) = \int_\chi M(x, \theta)\, \mathrm{d}\xi(x)$.

Unlike linear models, the Fisher information matrix for nonlinear models depends on at least one of the unknown parameters. A common approach is locally optimal designs (Chernoff 1953). The locally optimal design maximizes the criterion function evaluated at the best guess of the unknown parameters. Thus it depends on provisional values for the unknown parameters. To account for the uncertainty about the parameter values in the local optimal approach, other approaches such as minimax designs (Fedorov and Hackle 1997) and Bayesian designs (Chaloner and Larntz 1989; Chaloner and Verdinelli 1995) can be considered. These approaches use different strategies to take the uncertainty of the unknown parameters into account and consequently the optimization problems are much more computationally intensive than the local optimality approach. Nevertheless, locally optimal designs remain valuable and they often serve as benchmark designs to investigate the efficiencies of all other designs (Ford et al. 1992).

In optimal design theory, optimality criteria are concave functions of the information matrix. D-optimality is one of the most popular criteria and it has been studied by many authors, including Ford et al. (1992), Sitter and Wu (1993), Hedayat et al. (1997), Han and Chaloner (2003), Dette et al. (2006), Melas (2006) and Li and Majumdar (2008);

Li and Majumdar (2009). The D-optimal criterion function is defined as the logarithm of $|\mathcal{M}(\xi, \theta)|$, the determinant of the information matrix, if $\mathcal{M}(\xi, \theta)$ is nonsingular, and $-\infty$ if $\mathcal{M}(\xi, \theta)$ is singular (Atkinson et al. 2007). An approximate D-optimal design maximizes this criterion function over $\mathcal{H}$. An important property of a D-optimal design is that it minimizes the volume of the asymptotic confidence region for θ. It is also notable that D-optimal designs often perform well under other optimality criteria (Atkinson et al. 2007).

In this paper we consider D-optimal designs for the following two Gompertz regression models with homoscedastic normal error,

$$Y = \exp\left(-e^{-\theta_1 x+\theta_2}\right) + \epsilon, \tag{1}$$

$$Y = \theta_3 \exp\left(-e^{-\theta_1 x+\theta_2}\right) + \epsilon. \tag{2}$$

The mean function of these two models are re-parameterized from the previously discussed two Gompertz functions. A design is locally D-optimal with respect to $\theta' s$ is also locally D-optimal with respect to the previous parametrization by the transformation-invariance property of D-optimal designs (Atkinson et al. 2007).

Some preliminary results are given in Sect. 2. D-optimal designs for the homoscedastic Gompertz regression models (1) and (2) are studied in Sects. 3 and 4, respectively. We prove theoretically that locally D-optimal designs for the two Gompertz regression models are minimally supported. In parallel, we propose alternative designs which are found to be highly efficient.

2 Preliminaries

Let us consider a general homoscedastic regression model

$$y = f(x, \theta) + \epsilon, \tag{3}$$

with independent $\epsilon \sim N(0, \sigma^2)$. If $(\theta^T, \sigma^2)^{\mathrm{T}}$ is the parameter vector of interest, the per observation Fisher information matrix for a design measure ξ is $\tilde{\mathcal{M}}(\xi, \theta, \sigma^2) = \begin{pmatrix} \mathcal{M}(\xi, \theta) & 0 \\ 0 & \frac{1}{2\sigma^2} \end{pmatrix}$, where $\mathcal{M}(\xi, \theta) = \int_\chi \frac{\partial f(x,\theta)}{\partial \theta} \frac{\partial f(x,\theta)}{\partial \theta^T} \mathrm{d}\xi(x)$ is the information matrix for θ. Since $|\tilde{\mathcal{M}}(\xi, \theta, \sigma^2)| = 1/(2\sigma^2)|\mathcal{M}(\xi, \theta)|$, the D-optimal design for $(\theta^T, \sigma^2)^{\mathrm{T}}$ is the same as that for θ.

The equivalence theorem provides an important tool in the theory of optimum design. Originally established for linear models (Kiefer and Wolfowitz 1960), it was extended to nonlinear models by White (1973). For model (3), the design ξ^* is locally D-optimal at $\theta = \theta_0$ if and only if $d(\xi^*, x) \le k$ for all points $x \in \chi$ with equality holding at the support points of ξ^*, where k is the number of the unknown model parameters and the function $d(\xi, x) = \frac{\partial f(x,\theta_0)}{\partial \theta^T} M^{-1}(\xi, \theta_0) \frac{\partial f(x,\theta_0)}{\partial \theta}$ is the standardized variance of the model-based predicted response at x.

Assume the design region χ has one of the following forms: $\chi_0 = (-\infty, \infty)$, $\chi_1 = [a, \infty)$, $\chi_2 = (-\infty, b]$ or $\chi_3 = [a, b]$ with known a and b. It follows from the definition of the D-optimal criterion that a D-optimal design over $\mathcal{H}$ must be a nonsingular design, a design with a nonsingular information matrix. Let H be the set of nonsingular designs in $\mathcal{H}$.

For many nonlinear models it is known that the number of support points of the D-optimal design is equal to the number of the model parameters (Ford et al. 1992; Han and Chaloner 2003; Li and Majumdar 2008; Li and Majumdar 2009; Yang and Stufken 2009; Yang 2010). These designs are called minimally supported designs or saturated designs. However, there exist models for which D-optimal designs are not minimally supported (Sitter and Wu 1993). If a D-optimal design is minimally supported, then it has uniform weights on all support points. As a result, we only need to determine the k support points and the computation is greatly reduced. Once a D-optimal design is determined, one may either use it in the experiment, or use it as a benchmark to evaluate the D-efficiency of any other designs. The D-efficiency of a design ξ is defined as $D_{\text{eff}} = \left[\frac{|\mathcal{M}(\xi,\theta)|}{|\mathcal{M}(\xi^*,\theta)|}\right]^{1/k}$, where ξ^* is the D-optimal design (Hedayat et al. 1997).

In general it is rather difficult to technically prove that local D-optimal designs for nonlinear models are minimally supported designs. However, some sufficient conditions were established in the literature and derived for many important nonlinear models. By examining the behavior of $d(x, \xi) - k$ in a vertical neighborhood of zero, Li and Majumdar (2008) provide the following sufficient conditions and applied them to logistic models.

Theorem 1 (i) *For* $\chi_0 = (-\infty, \infty)$*, if* $\forall \xi \in H$, $\exists \epsilon > 0$*, such that every function in* $\{d(\xi, x) - k + c : 0 < c < \epsilon\}$ *has at most* $2k + 1$ *roots in the design region and a D-optimal design over* $\mathcal{H}$ *exists, then the D-optimal design must be minimally supported and unique.*
(ii) *Let* χ *be one of the following two forms:* $\chi_1 = [a, \infty)$ *or* $\chi_2 = (-\infty, b]$*. If* $\forall \xi \in H$, $\exists \epsilon > 0$*, such that every function in* $\{d(\xi, x) - k + c : 0 < c < \epsilon\}$ *has at most* $2k$ *roots in the design region* χ_1 *or* χ_2 *and a D-optimal design over* $\mathcal{H}$ *exists, then the D-optimal design must be minimally supported and unique. In addition, if* $\forall \xi \in H$, $\exists \epsilon > 0$*, such that every function in* $\{d(\xi, x) - k + c : 0 < c < \epsilon\}$ *has at most* $2k - 1$ *roots in the design region and a D-optimal design over* $\mathcal{H}$ *exists, then a (for* χ_1*) or b (for* χ_2*) is one of the support points of the D-optimal design.*
(iii) *For* $\chi_3 = [a, b]$*, if* $\forall \xi \in H$, $\exists \epsilon > 0$*, such that every function in* $\{d(\xi, x) - k + c : 0 < c < \epsilon\}$ *has at most* $2k - 1$ *roots in* χ_3*, then the D-optimal design must be minimally supported and unique and at least one of the boundary points is a support point of the D-optimal design. In addition, if* $\forall \xi \in H$, $\exists \epsilon > 0$*, such that every function in* $\{d(\xi, x) - k + c : 0 < c < \epsilon\}$ *has at most* $2k - 2$ *roots in* χ_3*, then both a and b are support points of the D-optimal design.*

In another paper, Li and Majumdar (2009) derived another set of sufficient conditions and applied them to the one-compartment pharmacokinetic model and a Poisson regression model. Although these sufficient conditions were derived under either a homoscedastic regression model or a generalized linear model, they could be applied to other types of models as long as the Fisher information matrix for a single observation $M(x, \theta)$ has a rank of 1. Yang and Stufken (2009) provided a unified approach to characterizing the minimally supported D-optimal designs for two-parameter nonlinear models. Besides D-optimality, their approach can also be applied to other commonly used criteria. Yang (2010) extended this approach to nonlinear models with an arbitrary number of parameters.

In this paper we apply Theorem 1 to the Gompertz models. To verify the sufficient conditions presented in Theorem 1, it is most important to identify the maximum number of roots for a class of functions. In this aspect, the theory of Tchebycheff systems (Karlin and Studden 1966) plays a key role. Let $u_0, \ldots, u_n$ denote continuous real-valued functions defined on a closed finite interval $I = [a, b]$. These functions will be called a weak Tchebycheff system over I, provided the $n+1$st order determinants

$$U\begin{pmatrix} u_0, & u_1, & \ldots, & u_n \\ t_0, & t_1, & \ldots, & t_n \end{pmatrix} = \begin{vmatrix} u_0(t_0) & u_0(t_1) & \cdots & u_0(t_n) \\ u_1(t_0) & u_1(t_1) & \cdots & u_1(t_n) \\ \vdots & \vdots & & \vdots \\ u_n(t_0) & u_n(t_1) & \cdots & u_n(t_n) \end{vmatrix}$$

are nonnegative whenever $a \le t_0 < t_1 < \cdots \le b$. If the determinants are strictly positive, then $\{u_0, \ldots, u_n\}$ is called a Tchebycheff system over I, abbreviated T-system. If $\{u_0, \ldots, u_n\}$ is a T-system over every finite interval I, then it is a T-system on $(-\infty, \infty)$.

If $\{u_0, \ldots, u_n\}$ is a T-system then the maximum number of distinct roots of any nontrivial linear combination of u_i's is n; conversely, if the maximum number of distinct roots of any nontrivial linear combination of u_i's is n, then either $\{u_0, \ldots, u_n\}$ or $\{u_0, \ldots, -u_n\}$ is a T-system. Since the maximum number of the roots is directly related to the length of the T-system, it is of interest to shorten the length of the T-system. The following Lemma 1 provides a tool, which is a stronger version than that in Li and Majumdar (2008). The tool is applied to prove Lemma 2. The proofs are presented in an Appendix.

Lemma 1 *Let* $\{u_{ij}(t),\ j = 1, \ldots, l_i\}_{i=1}^{s}$ *be* s *sequences of functions. If* $\forall\, j_i \in \{1, 2, \ldots, l_i\}, i \in \{1, 2, \ldots, s\}$, $\{u_{1j_1}, u_{2j_2}, \ldots, u_{sj_s}\}$ *are weak* T*-systems over* I*, and* $c_{ij} > 0,\ i = 1, 2, \ldots, s,\ j = 1, 2, \ldots, l_i$. *Then*

$$\left\{ \sum_{j=1}^{l_1} c_{1j}u_{1j}, \sum_{j=1}^{l_2} c_{2j}u_{2j}, \ldots, \sum_{j=1}^{l_s} c_{sj}u_{sj} \right\} \tag{4}$$

is also a weak T*-system over* I. *In addition, if at least one of* $\{u_{1j_1}, u_{2j_2}, \ldots, u_{sj_s}\}$ *is a* T*-system over* I*, then it is also a* T*-system over* I.

Lemma 2 *For any finite interval* $I = [a, b]$ *and constant* α, $\{1, e^{\alpha t}, e^{2\alpha t}, te^{2\alpha t}, t^2e^{2\alpha t}, e^{2e^{\alpha t}}\}$ *is a* T*-system.*

In addition, the following variant of part (iii) in Theorem 1 will be used in later sections.

Remark 1 If $\forall\, \xi \in H\ \lim_{x\to\pm\infty} d(\xi, x) < k$, and $\forall\, \xi \in H$, $\exists\, \epsilon > 0$, such that every function in $\{d(\xi, x) - k + c : 0 < c < \epsilon\}$ has at most $2k$ roots, then the D-optimal design is minimally supported on all four types of design region.

3 Two-parameter Gompertz regression model

First we consider the two-parameter Gompertz regression model (1), where $\theta = (\theta_1, \theta_2)^{\mathrm{T}}$ is the parameter vector of interest. Let $\omega_1 = \exp(-e^{-\theta_1 x+\theta_2})$, $\omega_2 = e^{-\theta_1 x+\theta_2}$ and $h(x, \theta) = \omega_1\omega_2(x, -1)^{\mathrm{T}}$. Then the per observation information matrix for θ is $\mathcal{M}(\xi, \theta) = \int_\chi h(x, \theta)h(x, \theta)^{\mathrm{T}}\, \mathrm{d}\xi(x)$ and $d(\xi, x) = h(x, \theta)^{\mathrm{T}}\mathcal{M}^{-1}(\xi, \theta)h(x, \theta)$. Let m_{ij} denote the (i, j)th element of $\mathcal{M}^{-1}(\xi, \theta)$. Then $d(\xi, x) = \omega_1^2\omega_2^2(m_{11}x^2 - 2m_{12}x + m_{22})$. In the next section we will consider locally D-optimal designs for the model (1) under all four types of design region.

3.1 Locally D-optimal designs

It is easy to verify that $\lim_{x\to\pm\infty} d(\xi, x) = 0$. For any constant c, $\omega_1^{-2}[d(\xi, x)-2+c]$ is a linear combination of $\{\omega_2^2, x\omega_2^2, x^2\omega_2^2, \omega_1^{-2}\}$. It is noted that $\omega_1^{-2} = e^{2\omega_2} = \Sigma_{i=0}^{\infty}(2^i/i!)\omega_2^i$. Thus $\omega_1^{-2}[d(\xi, x) - 2 + c]$ is also a linear combination of

$$\{\omega_2, \omega_2^2, x\omega_2^2, x^2\omega_2^2, \Sigma_{i=0 or i\geq 2}(2^i/i!)\omega_2^i\}. \tag{5}$$

It follows by a similar argument to the proof of Lemma 2 that we can show that (5) is a T-system. Hence $d(\xi, x) - 2 + c$ has at most four roots. From Theorem 1, the locally D-optimal design is minimally supported for $(-\infty, \infty)$, $[a, \infty)$ and $(-\infty, b]$. Since $\lim_{x\to\pm\infty} d(\xi, x) = 0$, it follows from Remark 1 that the locally D-optimal design is also minimally supported for $[a, b]$.

Consider a two-point uniform design ξ with support x_1, x_2 and let $\lambda_i = -\theta_1 x_i + \theta_2$. The determinant of the information matrix is

$$|\mathcal{M}(\xi, \theta)| \propto \exp(2\lambda_1 + 2\lambda_2 - 2e^{\lambda_1} - 2e^{\lambda_2})(\lambda_1 - \lambda_2)^2. \tag{6}$$

From (6) we know that the locally D-optimal design depends on the parameters through the linear combination, λ_i's. Let $\Lambda = \{\lambda : \lambda = -\theta_1 x + \theta_2, x \in \chi\}$ be the induced design region spanned by λ. The support points of the D-optimal design, expressed in λ's, can be determined by maximizing the right hand of (6) in the corresponding induced design regions. To implement the optimal design, we plug in the initial guess $\{\theta_1^{(0)}, \theta_2^{(0)}\}$ for the unknown parameter and solve for the support points in the original design region χ, i.e. $x_i^* = [\lambda_i^* - \theta_2^{(0)}]/(-\theta_1^{(0)})$.

The following theorem summarizes D-optimal designs for the model (1) under different design regions and it establishes underlying relationships among support points of D-optimal designs under these design regions. It takes inspiration from identically structured results in Ford et al. (1992). Like those authors, for ease of presentation, the induced design regions are considered and support points are expressed in λ's. Selected examples of D-optimal designs are presented in Table 1 for illustration purposes.

Theorem 2 (i) *For $\Lambda_0 = (-\infty, \infty)$, the D-optimal design is supported on* $\{\lambda_1^* = -1.044, \lambda_2^* = 0.499\}$.
(ii) *Consider $\Lambda_1 = [a, \infty)$. If $a \leq \lambda_1^*$, the D-optimal design is supported on* $\{\lambda_1^*, \lambda_2^*\}$*;*

Table 1 D-optimal designs for two-parameter homoscedastic Gompertz regression models

Design region Λ	D-optimal design	Design region Λ	D-optimal design
$(-\infty, \infty)$	$\{-1.044, 0.499\}$	$[-3, 6]$	$\{-1.044, 0.499\}$
$(-\infty, 0]$	$\{-1.35, 0\}$	$[-3, 0]$	$\{-1.35, 0\}$
$[0, \infty)$	$\{0, 0.806\}$	$[0, 6]$	$\{0, 0.806\}$
$[-1, 2]$	$\{-1, 0.508\}$	$[-1, 0]$	$\{-1, 0\}$
$[-10, -3]$	$\{-4.015, -3\}$	$[1, 6]$	$\{1, 1.35\}$

If $a > \lambda_1^*$, the D-optimal design is supported on $\{a, \lambda_a^*\}$, where λ_a^* is the only solution to $\lambda = a + 1/(e^\lambda - 1)$.
(iii) *Consider $\Lambda_2 = (-\infty, b]$. If $b \geq \lambda_2^*$, the D-optimal design is supported on $\{\lambda_1^*, \lambda_2^*\}$; If $b < \lambda_2^*$, the D-optimal design is supported on $\{\lambda_b^*, b\}$, where λ_b^* is the only solution to $\lambda = b + 1/(e^\lambda - 1)$.*
(iv) *Consider $\Lambda_3 = [a, b]$. If $a \leq \lambda_1^*$ and $b \geq \lambda_2^*$, the D-optimal design is supported on $\{\lambda_1^*, \lambda_2^*\}$; if $a \leq \lambda_1^*$ and $b < \lambda_2^*$, the D-optimal design is supported on $\{\max(a, \lambda_b^*), b\}$; if $a > \lambda_1^*$ and $b \geq \lambda_2^*$, the D-optimal design is supported on $\{a, \min(\lambda_a^*, b)\}$; if $a > \lambda_1^*$ and $b < \lambda_2^*$, the D-optimal design is supported on $\{a, b\}$.*

Note that in part (i) of the theorem we consider $(-\infty, \infty)$ as a design interval. This makes sense since our model is equivalent to a weighted linear model for which the design space or locus is bounded for all λ. This is a feature of all the models considered by Ford et al. (1992). Moreover the support points of the D-optimal design are finite. In theory we do not need to impose limits on the design variable λ.

3.2 Efficient designs

Although the use of the D-optimal designs can achieve the best precision in the estimation of the model parameters, it may not be applicable in practice as we may not be so sure about the appropriateness of the chosen statistical model. However the D-optimal designs with the same number of support points as that of the model parameters can not be used to test the goodness of fit for the selected model. In addition, it may be inconvenient to implement the optimal design logistically. So we now propose some alternative efficient designs.

Equally spaced and uniformly weighted designs (ESUWDs) are widely adopted designs in practical situations. An m-point ESUWD has support points $\{\lambda + (i - 1)\delta_\lambda,\ i = 1, \ldots, m\}$ with uniform weight $1/m$. Choice of λ and δ_λ can be dictated by ease of implementation or for a desired efficiency. A D-optimal m-point ESUWD maximizes, over λ and δ_λ, the determinant of the Fisher information matrix among all the m-point ESUWDs.

It is noted that ESUWDs take equally spaced support points on the x-axis. By symmetry we can also take equally spaced points on the y-axis and then project points on the response curve to the x-axis to obtain the support points. We call such designs as equally spaced in response and uniformly weighted designs (ESRUWDs). An m-point ESRUWDs has support points $\{\ln[-\ln(y + (i - 1)\delta_y)],\ i = 1, \ldots, m\}$ with uniform

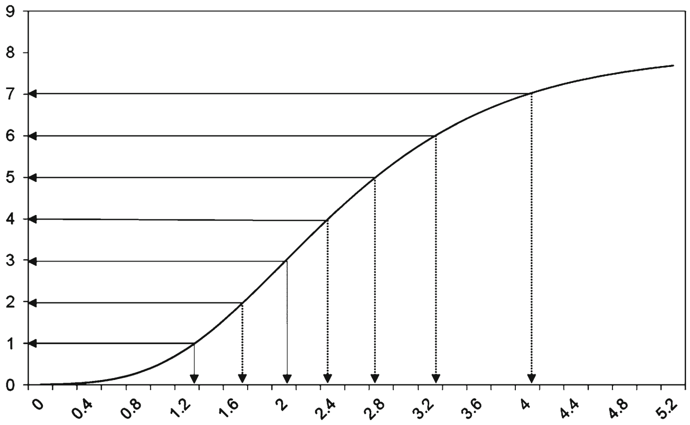

Fig. 2 Equally spaced in response and uniformly weighted designs (ESRUWD)

Table 2 D-efficiencies of alternative designs for two-parameter homoscedastic Gompertz models

m	D-optimal ESUWD		D-optimal ESRUWD		Naive ESRUWD	
	$(\lambda^*, \delta_\lambda^*)$	Efficiency (%)	(y^*, δ_y^*)	Efficiency (%)	(y^*, δ_y^*)	Efficiency (%)
3	(−1.26, 0.91)	91.8	(0.167, 0.288)	91.6	(0.250, 0.250)	87.1
4	(−1.31, 0.65)	90.0	(0.151, 0.205)	89.3	(0.200, 0.200)	87.5
5	(−1.34, 0.51)	89.4	(0.139, 0.160)	88.4	(0.167, 0.167)	86.9
6	(−1.37, 0.42)	89.1	(0.130, 0.132)	87.9	(0.143, 0.143)	86.1
7	(−1.40, 0.36)	88.9	(0.123, 0.112)	87.6	(0.125, 0.125)	85.3
8	(−1.45, 0.32)	88.8	(0.118, 0.098)	87.4	(0.111, 0.111)	84.6

weight $1/m$, where $y + (i - 1)\delta_y \in (0, 1)$. See Fig. 2 for an example of 7-point ESRUWD. A D-optimal m-point ESRUWD maximizes, over y and δ_y, the determinant of the Fisher information matrix among all the m-point ESRUWDs. For models such as the two-parameter Gompertz models whose mean function is nonlinear in the design variable, ESRUWDs do not have equally spaced support points. However, a desired property for ESRUWDs is that the distance of the support points reflects the steepness of the response curve. In the place of the response curve with sharper steepness the corresponding support points are closer.

To search for D-optimal ESUWDs and ESRUWDs numerical techniques, such as the Newton–Raphson method, are needed. In contrast, we can also consider a naive ESRUWD with support points $\{\ln[-\ln(i/(m+1))],\ i = 1, \ldots, m\}$. As m increase, this design provides a good coverage of the response range.

Table 2 provides the efficiencies of these designs for $3 < m < 8$. All three alternatives designs have high efficiencies and comparatively the D-optimal ESUWDs have the highest efficiencies. For $m > 3$, all three proposed alternative designs have the flexibility of enabling a check for the goodness of fit for the two-parameter Gompertz model.

4 Three-parameter Gompertz regression model

Consider the three-parameter Gompertz regression model (2), where $\theta = (\theta_1, \theta_2, \theta_3)^{\mathrm{T}}$ is the parameter vector of interest. Let $h(x, \theta) = \omega_1(\theta_3 x\omega_2, -\theta_3\omega_2, 1)^{\mathrm{T}}$, where $\omega_1 = \exp(-e^{-\theta_1 x+\theta_2})$, $\omega_2 = e^{-\theta_1 x+\theta_2}$. Then $\mathcal{M}(\xi, \theta) = \int_\chi h(x, \theta)h(x, \theta)^{\mathrm{T}}\, \mathrm{d}\xi(x)$ and $d(\xi, x) = h(x, \theta)^{\mathrm{T}}\mathcal{M}^{-1}(\xi, \theta)h(x, \theta)$. In the next two sections we will consider the locally D-optimal design and efficient designs for the model (2).

4.1 Locally D-optimal designs

Let m_{ij} denote the (i, j)th element of $\mathcal{M}^{-1}(\xi, \theta)$. Then, for any constant c, $\omega_1^{-2}[d(\xi, x) - 3 + c]$ is a linear combination of

$$\{1, \omega_2, \omega_2^2, x\omega_2^2, x^2\omega_2^2, u(x)\}, \tag{7}$$

where $u(x) = -2m_{13}\theta_3 x\omega_2 + (3 - c)\omega_1^{-2}$.

It follows from Lemma 2 that $\{1, \omega_2, \omega_2^2, x\omega_2^2, x^2\omega_2^2, \omega_1^{-2}\}$ is a T-system. For a design ξ with support $x_1, x_2, \ldots, x_t$ and corresponding weights $p_1, p_2, \ldots, p_t$, where $t \geq 3$, $p_i > 0$ and $\sum_{i=1}^{t} p_i = 1$, let M_{ij} be the minor corresponding to the (i, j)th element of $\mathcal{M}(\xi, \theta)$. Let $v(r, s) = \begin{vmatrix} \omega_1(r) & \omega_1(r)\omega_2(r) \\ \omega_1(s) & \omega_1(s)\omega_2(s) \end{vmatrix} \times \begin{vmatrix} \omega_1(r)\omega_2(r) & \omega_1(r)\omega_2(r)r \\ \omega_1(s)\omega_2(s) & \omega_1(s)\omega_2(s)s \end{vmatrix}$. It can be shown that $M_{13} = \theta_3^3 \sum_{1\leq i<j\leq t} p_i p_j v(x_i, x_j)$.

If $\theta_1 > 0$, then both $\{\omega_1, -\omega_1\omega_2\}$ and $\{\omega_1\omega_2, \omega_1\omega_2 x\}$ are T-systems. Thus $v(r, s) < 0\, \forall r \neq s$ and $M_{13}\theta_3 < 0$. Since $m_{13} = M_{13}/|M|, m_{13}\theta_3 < 0$ for $\xi \in H$. From Karlin and Studden (1966, p10), we know $\{1, \omega_2, \omega_2^2, x\omega_2^2, x^2\omega_2^2, x\omega_2\}$ is a T-system. So $u(x)$ is a linear combination of $\{\omega_1^{-2}, x\omega_2\}$ with positive coefficients for $c < 3$. It follows from Lemma 1 that (7) is a T-system, which implies that $(d(\xi, x) - 3 + c)$ has at most five roots. Similarly it can also be shown that $(d(\xi, x) - 3 + c)$ has at most five roots if $\theta_1 < 0$. From part (ii)&(iii) of Theorem 1, the locally D-optimal designs for $(-\infty, b]$ or $(-\infty, b]$ or $[a, b]$ are minimally supported if a D-optimal design exist in the corresponding design space.

From (8) we know that the locally D-optimal design does not depend on θ_3 and depends on θ_1 and θ_2 through the linear combination $-\theta_1 x_i + \theta_2$. Let $\Lambda = \{\lambda : \lambda = -\theta_1 x + \theta_2, x \in \chi\}$ be the induced design region spanned by λ. The support points of the D-optimal design, expressed in the λ's, can be determined by maximizing the right hand of (8) in the corresponding induced design region. Again we would like to summarize the D-optimal design under the induced design space Λ. Let us first consider the infinite design interval $\Lambda = (-\infty, \infty)$.

Consider a typical three-point uniformly weighted design ξ with support x_1, x_2, x_3 and let $\lambda_i = -\theta_1 x_i + \theta_2$. The determinant of the information matrix is

$$|\mathcal{M}(\xi, \theta)| = \frac{\theta_3^4[e^{\lambda_1+\lambda_2}(\lambda_1 - \lambda_2) + e^{\lambda_2+\lambda_3}(\lambda_2 - \lambda_3) + e^{\lambda_3+\lambda_1}(\lambda_3 - \lambda_1)]^2}{27\theta_1^2 \exp 2(e^{\lambda_1} + e^{\lambda_2} + e^{\lambda_3})} \tag{8}$$

For fixed $\lambda_2 < \lambda_3$, it can be shown that the numerator of (8) is decreasing and the denominator is increasing in $\lambda_1 \in (-\infty, \lambda_2)$. This implies that $-\infty$ is one of the support points of the locally D-optimal design in a limiting case when $-\infty$ is also included in the design space Λ. Plugging $\lambda_1 = -\infty$ into (8) will result in an equation equivalent to (6), which is the D-optimal criterion function for the two-parameter Gompertz regression model. Thus, in this limiting case, the other two support points of the locally D-optimal design for the three-parameter Gompertz regression model (2) are the same as those for the two-parameter Gompertz regression model (1), namely -1.044 and 0.499.

Hence a finite lower limit must be imposed on the induced design space Λ. Since the other two support points are finite, no finite upper limit on Λ is needed from a mathematical point of view. In the following theorem we summarize D-optimal designs for the three-parameter Gompertz regression model on $\Lambda = [a, \infty)$ and $[a, b]$.

Theorem 3 (i) *The D-optimal design $\xi^*_{a\infty}$ on $\Lambda = [a, \infty)$ is minimally supported with three support points, $a < \lambda^*_2(a) < \lambda^*_3(a)$.*

(ii) *The D-optimal design ξ^*_{ab} on $\Lambda = [a, b]$ is minimally supported with three support points, where the lowest support point is a and the largest support point is* $\min\{b, \lambda^*_3(a)\}$. *In the special case if $b \geq \lambda^*_3(a)$, then $\xi^*_{ab} = \xi^*_{a\infty}$.*

Proof We have shown that both $\xi^*_{a\infty}$ and ξ^*_{ab} are minimally supported and a is one of the support points. We only need to show the largest support point of ξ^*_{ab} is $\min\{b, \lambda^*_3(a)\}$.

If $b < \lambda^*_3(a)$ and the largest support point of ξ^*_{ab} is $\tilde{\lambda_3} < b$, then $d(\xi^*_{ab}, \lambda) < 3$ for $\lambda \in (\tilde{\lambda_3}, b]$. Given the fact that $(d(\xi^*_{ab}, \lambda) - 3 + c)$ has at most 5 roots for any $c < 3$, $d(\xi^*_{ab}, \lambda) < 3$ if $\lambda > b$. This implies that ξ^*_{ab} is a D-optimal design on $\Lambda = [a, \infty)$ different from $\xi^*_{a\infty}$, which contradicts with the uniqueness of the D-optimal design on $\Lambda = [a, \infty)$. Thus the largest support point of ξ^*_{ab} is b if $b < \lambda^*_3(a)$. Similarly it can shown that if $b \geq \lambda^*_3(a)$ then $\xi^*_{ab} = \xi^*_{a\infty}$. □

In practice, the induced design region Λ may be chosen by the extent to which researchers want to plan the experiment. For example, it is subject to ethical reasons that we would sacrifice the animals before the cancer tumor reaches a certain size in oncology animal studies. An induced design region $\Lambda = [\ln(-\ln(p)), \infty)$ corresponds to a design region with the size of the tumor up to $100p$ percent of the maximum tumor size. At the same time the experimenter may choose to make the first measurement after the tumor has grown to certain size. In this case, an induced design region of $\Lambda = [\ln(-\ln(p_1)), \ln(-\ln(p_2))]$ corresponds to a design region with the size of tumor between $100p_2$ and $100p_1$ percent of the maximum tumor size.

Table 3 provides the support points of the D-optimal designs for different design regions. It is noted that support points of the D-optimal design on the design space in the type of $\Lambda = [a, \infty)(p_2 = 0)$ increases with a, the lower limit of the design space. We can also observe that the D-optimal design on $\Lambda = [-2.25, \infty](\mathrm{p}_1 = 0.9, p_2 = 0)$ is the same as that on $\Lambda = [-2.25, 0.834](p_1 = 0.9, p_2 = 0.1)$ because $b > \lambda^*_3(a)$ (i.e. $0.834 > 0.643$).

Table 3 D-optimal designs for three-parameter homoscedastic Gompertz regression models; $\Lambda = [\ln(-\ln(p_1)), \ln(-\ln(p_2))]$

Design region		D-optimal design	Design region		D-optimal design
p_1	p_2	Support points	p_1	p_2	Support points
1/2	0	$\{-0.367, 0.191, 0.996\}$	1/2	1/10	$\{-0.367, 0.147, 0.834\}$
3/5	0	$\{-0.672, 0.014, 0.91\}$	3/5	1/10	$\{-0.672, -0.009, 0.834\}$
7/10	0	$\{-1.031, -0.181, 0.821\}$	7/10	1/5	$\{-1.031, -0.297, 0.476\}$
4/5	0	$\{-1.50, -0.389, 0.73\}$	4/5	1/5	$\{-1.50, -0.482, 0.476\}$
9/10	0	$\{-2.25, -0.633, 0.643\}$	9/10	1/10	$\{-2.25, -0.633, 0.643\}$

Table 4 D-efficiencies of alternative designs for three-parameter homoscedastic Gompertz models; $\Lambda = [\ln(-\ln(p)), \infty)$

p	m	D-optimal ESUWD		D-optimal ESRUWD	
		$(\lambda^*, \delta_\lambda^*)$	Efficiency (%)	(y^*, δ_y^*)	Efficiency (%)
1/2	4	(−0.367, 0.474)	92.6	(0.500, 0.146)	92.5
	5	(−0.367, 0.367)	88.8	(0.500, 0.111)	87.7
	6	(−0.367, 0.298)	86.0	(0.500, 0.090)	84.6
	7	(−0.367, 0.251)	83.8	(0.500, 0.075)	82.3
	8	(−0.367, 0.217)	82.2	(0.500, 0.065)	80.7
2/3	4	(−0.903, 0.605)	92.2	(0.667, 0.193)	92.7
	5	(−0.903, 0.467)	88.4	(0.667, 0.147)	88.0
	6	(−0.903, 0.380)	85.8	(0.667, 0.119)	84.9
	7	(−0.903, 0.320)	83.9	(0.667, 0.100)	82.6
	8	(−0.903, 0.277)	82.5	(0.667, 0.086)	80.9

Again, in order to implement the optimal designs, we plug in the initial guess $\{\theta_1^{(0)}, \theta_2^{(0)}\}$ for the unknown parameter and solve for the support points in the original design region χ, i.e. $x_i^* = [\lambda_i^* - \theta_2^{(0)}]/(-\theta_1^{(0)})$.

4.2 Efficient designs

In this section we present alternative designs for the three-parameter Gompertz models under the induced design region $\Lambda = [\ln(-\ln(p), \infty)$. Similar to the two-parameter Gompertz model, D-optimal ESUWDs and ESRUWDs can be found by numerical computations and D-optimal designs can be used as benchmarks to evaluate the performance of these designs. Table 4 presents the efficiency of these alternative designs. For $4 \le m \le 8$, the D-optimal ESUWDs and ESRUWDs have high efficiency although the efficiency decreases as m increases. It is worth pointing out that, $\lambda^* = \ln(-\ln(p))$ for all D-optimal ESUWD designs shown in the Table 4.

5 Conclusions

In this paper we have studied D-optimal designs for two Gompertz regression models with homoscedastic variance. Because of the nonlinear nature of the Gompertz functions, we considered local optimality criteria. For both Gompertz regression models, we provided theoretical proofs to show that D-optimal designs are minimally supported. In addition, we studied the efficiencies of equally weighted designs with support points equally spaced in either design space (i.e., ESUWD) or the response space (i.e, ESRUWD). Using D-optimal designs as benchmark designs these equally spaced designs have good efficiencies.

Although heteroscedastic regression models have been used in fitting growth curves more widely, we believe that the optimal and efficient designs obtained in this paper for homoscedastic models should still be valuable and applicable under certain heteroscedastic structures. We will continue to investigate this in future research.

6 Appendix

6.1 Proof of Lemma 1

Let $v_i = \sum_{j=1}^{l_i} c_{ij} u_{ij}$. The sth order discriminant is

$$U\begin{pmatrix} v_1, & v_2, & \dots, & v_s \\ t_1, & t_2, & \dots, & t_s \end{pmatrix} = \sum_{j_1=1}^{l_1} \cdots \sum_{j_s=1}^{l_s} \left[\left(\prod_{i=1}^{s} c_{i,j_i} \right) U \begin{pmatrix} u_{1j_1}, & u_{2j_2}, & \dots, & u_{sj_s} \\ t_1, & t_2, & \dots, & t_s \end{pmatrix} \right]$$

Since $\{u_{1j_1}, u_{2j_2}, \dots, u_{sj_s}\}$ are weak T-systems, $\forall\, j_i \in \{1, 2, \dots, l_i\}$, $i \in \{1, 2, \dots, s\}$, $U\begin{pmatrix} u_{1j_1}, & u_{2j_2}, & \dots, & u_{sj_s} \\ t_1, & t_2, & \dots, & t_s \end{pmatrix} \geq 0$ whenever $t_1 < t_2 < \cdots < t_s$. Thus $U\begin{pmatrix} v_1, & v_2, & \dots, & v_s \\ t_1, & t_2, & \dots, & t_s \end{pmatrix} \geq 0$ whenever $t_1 < t_2 < \cdots < t_s$. This means $\{v_1, v_2, \dots, v_s\}$, i.e. (4), is a weak T-system. If at least one of $\{u_{1j_1}, u_{2j_2}, \dots, u_{sj_s}\}$ is a T-system, then whenever $t_1 < t_2 < \cdots < t_s$. This means $\{v_1, v_2, \dots, v_s\}$ is a T-system.

6.2 Proof of Lemma 2

It is noted that $e^{2e^{\alpha t}} = \sum_{i=0}^{\infty} (2^i/i!) e^{i\alpha t}$. It follows from Karlin and Studden (1966, p10) that $\{1, e^{\alpha t}, e^{2\alpha t}, te^{2\alpha t}, t^2 e^{2\alpha t}, e^{i\alpha t}\}$ is a weak T-systems if $0 \leq i \leq 2$; Otherwise $\{1, e^{\alpha t}, e^{2\alpha t}, te^{2\alpha t}, t^2 e^{2\alpha t}, e^{i\alpha t}\}$ is a T-system. Since the coefficients for $e^{i\alpha t}$ are all positive, it follows from Lemma 1 that $\{1, e^{\alpha t}, e^{2\alpha t}, te^{2\alpha t}, t^2 e^{2\alpha t}, e^{e^{2\alpha t}}\}$ is a T-system.

Acknowledgments The author thanks the referees and editors for helpful suggestions that have improved the paper.

References

Atkinson, A. C., Donev, A. N. Tobias, R. (2007). *Optimum Experimental Designs, with SAS*. Oxford: Oxford University Press.

Chaloner, K., Larntz, K. (1989). Optimal Bayesian design applied to logistic regression experiments. *Journal of Statistical Planning and Inference, 21*, 191–208.

Chaloner, K., Verdinelli, I. (1995). Bayesian experimental design: a review. *Statistical Science, 10*, 273–304.

Chernoff, H. (1953). Locally optimal design for estimating parameters. *Annals of Mathematical Statistics, 24*, 586–602.

Dette, H., Melas, V. B., Wong, W.K. (2006). Locally D-optimal designs for exponential regression. *Statistica Sinica, 16*, 789–803.

Fedorov, V. V., Hackle, P. (1997). *Model-Oriented Design of Experiments*. New York: Springer.

Ford, I., Torsney, B., Wu, C. F. J. (1992). 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, 54*, 569–583.

Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contigencies. *Philosophical Transactions B, 115*, 513–585.

Han, C., Chaloner, K. (2003). D- and c-optimal designs for exponential regression models used in viral dynamics and other applications. *Journal of Statistical Planning and Inference, 115*, 585–601.

Hedayat, A., Yan, B., Pezzuto, J. M. (1997). Modeling and identifying optimum designs for fitting dose-response curves based on raw optical density data. *Journal of the American Statistical Association, 92*, 1132–1140.

Karlin, S., Studden, W. J. (1966). *Tchebyschev Systems: With Applications in Analysis and Statistics*. New York: Wiley.

Kiefer, J., Wolfowitz, J. (1960). The equivalence of two extremum problems. *Canadian Journal of Mathematics, 12*, 363–366.

Laird, K. A. (1965). Dynamics of tumour growth: comparison of growth rates and extrapolation of growth curve to one cell. *British Journal of Cancer, 19*, 278–291.

Li, G., Majumdar, D. (2008). D-optimal designs for logistic models with three and four parameters. *Journal of Statistical Planning and Inference, 138*, 1950–1959.

Li, G., Majumdar, D. (2009). Some results on D-optimal designs for nonlinear models with applications. *Biometrika, 96*, 487–493.

Marusic, M., Vuk-Pavlovic, S. (1993). Prediction power of mathematical models for tumor growth. *Journal of Biological System, 1*, 69–78.

Melas, V. B. (2006). Functional approach to optimal experimental design. In *Lecture notes in statistics* (Vol. 184). New York: Springer.

Sitter, R. R., Wu, C. F. J. (1993). Optimal designs for binary response experiments. *Scandinavian Journal of Statistics, 20*, 329–341.

White, L. (1973). An extension of the general equivalence theorem to nonlinear models. *Biometrika, 60*, 345–348.

Yang, M. (2010). On the de la Garza Phenomenon. *Annals of Statistics, 38*, 2499–2524.

Yang, M., Stufken, J. (2009). Support points of locally optimal designs for nonlinear models with two parameters. *Annals of Statistics, 37*, 518–541.