DESIGNS FOR TESTING LACK OF FIT FOR A CLASS OF SIGMOID CURVE MODELS

A Dissertation
Submitted
to the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy

By
Ying Su
January, 2013

Examining Committee Members:
Damaraju Raghavarao, Advisory Chair, Department of Statistics, Temple University
Boris Iglewicz, Department of Statistics, Temple University
Pallavi Chitturi, Department of Statistics, Temple University
Ivan S. F. Chan, External Member, Merck Research Laboratories

ABSTRACT

OPTIMAL DESIGNS FOR A CLASS OF SIGMOID CURVE MODELS

Ying Su

DOCTOR OF PHILOSOPHY

Temple University

Advisor Chair: Dr. Damaraju Raghavarao

Sigmoid curves have found broad applicability in biological sciences and biopharmaceutical research during the last decades. A well planned experiment design is essential to accurately estimate the parameters of the model. In contrast to a large literature and extensive results on optimal designs for linear models, research on the design for nonlinear, including sigmoid curve, models has not kept pace. Furthermore, most of the work in the optimal design literature for nonlinear models concerns the characterization of minimally supported designs. These minimal, optimal designs are frequently criticized for their inability to check goodness of fit, as there are no additional degrees of freedom for the testing. This design issue can be a serious problem, since checking the model adequacy is of particular importance when the model is selected without complete certainty. To assess for lack of fit, we must add at least one extra distinct design point to the experiment. The goal of this dissertation is to identify optimal or highly efficient designs capable of checking the fit for sigmoid curve models.

In this dissertation, we consider some commonly used sigmoid curves, including logistic, probit and Gompertz models with two, three, or four parameters. We use D-optimality as our design criterion. We first consider adding one extra point to the design, and consider five alternative designs and discuss their suitability to test for lack of fit. Then we extend the results to include one more additional point to better understand the compromise among the need of detecting lack of fit, maintaining high efficiency and the practical convenience for the practitioners. We then focus on the two-parameter Gompertz model, which is widely used in fitting growth curves yet less studied in literature, and explore three-point designs for testing lack of fit under various error variance structures.

One reason that nonlinear design problems are so challenging is that, with nonlinear models, information matrices and optimal designs depend on the unknown model parameters. We propose a strategy to bypass the obstacle of parameter dependence for the theoretical derivation. This dissertation also successfully characterizes many commonly studied sigmoid curves in a generalized way by imposing unified parameterization conditions, which can be generalized and applied in the studies of other sigmoid curves.

We also discuss Gompertz model with different error structures in finding an extra point for testing lack of fit.

ACKNOWLEDGMENTS

I would like to take this opportunity to express my deepest gratitude to my committee members, my colleagues and my family. Without their guidance, patience and support, the completion of this dissertation would not have been possible.

I am most grateful for my advisor, Dr. Damaraju Raghavarao, whose wisdom, compassion and encouragement guided me through the difficult times. I also would like to thank Dr. Boris Iglewicz, Dr. Pallavi Chitturi and Dr. Ivan SF Chan, for their insights and guidances. My work would not have been accomplished without their helps.

I am very thankful for having the constant support from my management and colleagues at Merck. Merck's philosophy to encourage employees to advance their knowledge offers me the opportunity to achieve this goal.

Finally, I would like to thank my family for their unconditional love and support. A special thank to my husband, Charles Tan, who is always there for me through the good times and bad. I also owe my gratitude to my parents and my brother, who are always supporting me with their best wishes.

DEDICATION

This dissertation is dedicated to my children, Nicole and Jake.

TABLE OF CONTENTS

PAGE

LIST OF TABLES

LIST OF FIGURES

Page

CHAPTER 1
INTRODUCTION

Sigmoid curves are widely used in biological sciences and biopharmaceutical research. For these sigmoid (or S shaped) curves, the dependent variable y stays almost flat at extremes of the independent variable, x; only changes noticeably when x is within a dynamic window. Further, y changes slowly, then rapidly, then slowly again within the window. Since the dependent variable y is nonlinear for the independent variable x, as well as some of the parameters, they are called nonlinear models.

Sigmoid models have found broad applicability for many biological phenomenon or reactions. One of the earliest examples is Archibald Hill's classical work more than a century ago, in 1910, on cooperative binding of oxygen by hemoglobin (Nelson and Cox, 2005). The Hill's equation is a logistic curve, which is a symmetric sigmoid curve. Following this tradition, four-parameter logistic curves are commonly used in bioanalytical methods, such as immunoassays and bioassay. Within the context of pharmaceutical or biotechnology development, another example is the concentration-effect curves. The Emax model (Meibohm and Darendorf, 1997) that is popular in pharmacology is essentially the same mathematical model as Hill's equation. Typical applications of sigmoid models can be found in biology (see Landaw and DiStefano,1984; Chervoneva, Li, Iglewicz, Waldman and Hyslop, 2007); chemistry, pharmacokinetics (see Liebig, 1988; Krug and Liebig, 1988); toxicology (see Becka, Bolt and Urfer, 1993; Becka and Urfer, 1996); or microbiology (see Coleman and Marks, 1998).

The usefulness and popularity of sigmoid models have spurred a large literature on their data analysis, but results on nonlinear optimal designs are sparse. There are extensive results on the theory of optimal designs for linear models (see Fedorov, 1972; Atkinson, 1982; Steinberg and Hunter, 1984; Pukelsheim, 1993), while the problem is much more difficult and not nearly as well understood for nonlinear models. One reason that nonlinear design problems are so challenging is that, with nonlinear models, information matrices and optimal designs depend on the unknown model parameters. A common approach to solve this dilemma is to use locally optimal designs based on the best guess of the parameters, as defined in Chernoff (1953). As in much of the past work, this dissertaion will focus on locally optimal designs. As pointed out by Ford, Torsney and Wu (1992), locally optimal designs are important if reliable initial parameters are available from previous experiments, but can also function as a benchmark for comparing designs.

With respect to experimental design criteria, our criterion will be D-optimality, introduced by Wald (1943) and Mood (1946). The D-optimal designs are comprised of those design points that minimize the determinant of the covariance matrix of the estimated parameters. In addition to minimizing the generalized variance, Box (1971) showed that the D-optimality criterion is desirable in that it tends to reduce the biases of the maximum likelihood estimates of the parameters in nonlinear models.

Local D-optimal designs are sometimes known for nonlinear models (e.g., Ford, Torsney and Wu, 1992; Hedayat, Yan and Pezzuto, 1997; Han and Chaloner, 2003; Hedayat, Zhong and Nie, 2004; Dette, Melas and Pepelyshev, 2004; Dette, Melas and Wong, 2005; Biedermann, Dette and Zhu, 2006; Wang, Myers, Smith, and Ye, 2006; Li and Majumdar, 2008, 2009, etc.). However, these designs are usually minimally supported, i.e., the number of design points is equal to the number of parameters in the model. Consequently there are no degrees of freedom available to perform a goodness-of-fit test; in other words, minimally supported designs provide no opportunity to test the lack of fit. This can be a serious problem, since checking the model adequacy is of particular importance when the model is selected without complete certainty. Optimal designs are model-dependent, therefore a misspecified model could lead to substantial loss of efficiency. Khuri, Mukherjee, Sinha and Ghosh (2006) reviewed nonlinear design issues.

It is not uncommon to observe that the theoretical optimal designs for many models used in practice typically do not have enough design points to enable the researcher to verify the model assumptions. However, we have found very limited statistical literature on optimal designs incorporating goodness-of-fit testing in nonlinear models (see Lupinacci and Raghavarao, 2000). In this dissertation we will develop designs that not only allow us to estimate the parameters with great precision, but also have the ability to test for lack of fit for commonly used sigmoid curves.

One reason that nonlinear design problems are so challenging is that, with nonlinear models, information matrices and optimal designs depend on the unknown model parameters. In order to deal with this obstacle, we propose a transformation technique for the information matrices so that the derivation of the optimal designs is completely independent from the parameters. This dissertaion also successfully characterizes many commonly studied sigmoid curves in a generalized way by imposing unified parameterization conditions. This approach yields very general results and can be applied in the studies of other sigmoid curves.

In Chapter 2 we will review the previous work that are relevant to this dissertation. There we will provide a brief overview on certain aspects of nonlinear models that we need, and summarize the work that has been done on designs for testing lack of fit.

In Chapter 3, we will develop minimal plus one point designs for testing lack of fit for sigmoid curves. We will focus on three most commonly used sigmoid models: logistic, probit and Gompertz. Each of these models will be discussed in two-, three- or four-parameter cases. The parameterization approach and transformation strategy are presented along with results of implementing these methodologies on sigmoid models that we study in this chapter.

In Chapter 4, we will extend the work of Chapter 3 to include one more point in the design for sigmoid models. We want to explore and see how the addition of one or two points to the minimal design impact the efficiency.

Among the class of sigmoid models we studied in this dissertation, the Gompertz model is somewhat different because of its asymmetry, and is less studied in literature. In Chapter 5, we will focus on the two-parameter Gompertz model, and explore three-point designs for testing lack of fit under various error variance structures. For each variation assumption, we will develop the minimal D-optimal design, then construct three-point designs for testing lack of fit. This work is expected to help further understanding of Gompertz models.

We will conclude this dissertation with Chapter 6 where we briefly summarize our results and outline possible future work on this topic.

CHAPTER 2
LITERATURE REVIEW

2.1 Overview of Nonlinear Models

The general form for nonlinear model with single independent variable is

$$y_i = f(\mathbf{x}_i;\boldsymbol{\theta}) + \varepsilon_i, \quad i = 1,2,\ldots n, \tag{2.1.1}$$

where $\boldsymbol{\theta}$ is a $k \times 1$ parameter vector. The nonlinearity of f with respect to $\boldsymbol{\theta}$ is the defining characteristic of a nonlinear model. It turns out that the degree of this nonlinearity drives the statistical and numerical properties of the model. For sigmoid curves, the nonlinearity of f with respect to $\mathbf{x}$ is more or less similar, but degree of the nonlinearity of f with respect to $\boldsymbol{\theta}$ can be very different.

As noted by Myers (1986), the Gauss-Newton algorithm is the most frequently used method in software computing algorithms to find the least squares estimator, $\hat{\boldsymbol{\theta}}$, of the parameter vector, $\boldsymbol{\theta}$, in nonlinear models. The Gauss-Newton algorithm basically reduces nonlinear least-squares problems to a sequence of linear least-squares problems and requires only first-order information about the model functions. When the prediction errors are independent, but not homogeneous, such that $Var(\varepsilon_i) = g[f(x_i;\boldsymbol{\theta})]$, the least squares objective is to find $\hat{\boldsymbol{\theta}}$ that minimizes the weighted residual sum of squares,

$$S(\boldsymbol{\theta}) = [\mathbf{y} - f(\boldsymbol{\theta})]' \mathbf{V}^{-1} [\mathbf{y} - f(\boldsymbol{\theta})], \qquad (2.1.2)$$

where $\mathbf{V}$ is a diagonal matrix whose i^{th} diagonal element is $g[f(x_i;\boldsymbol{\theta})]$, and we denote $f(\boldsymbol{\theta}) = f(x_i;\boldsymbol{\theta})$ for convenience. There could be different error structures in different situations. Due to the complexity of three- and four-parameter sigmoid models that we study, in this work we will concentrate on the constant variation, and assume that $\mathbf{V} = \sigma^2 \mathbf{I}_n$, where $\mathbf{I}_n$ is the identity matrix of order n. When ε_i's are independent and identically distributed as $N(0, \sigma^2)$, which we assume in Chapter 3 and 4, minimizing least square objective function, $S(\boldsymbol{\theta})$, is the same as maximizing the likelihood function. Without loss of generality, we further assume $\sigma^2 = 1$, so that (2.1.2) can be written as

$$S(\boldsymbol{\theta}) = [\mathbf{y} - f(\boldsymbol{\theta})]' [\mathbf{y} - f(\boldsymbol{\theta})]. \qquad (2.1.3)$$

To make the discussion consistent throughout the dissertation, we keep $\mathbf{V}$ in all the derivations even though we assume $\mathbf{V} = \mathbf{I}_n$ in Chapters 3 and 4. For nonlinear models, the closed form solutions of $\hat{\boldsymbol{\theta}}$ minimizing $S(\boldsymbol{\theta})$ are rarely available, hence iterative algorithms are needed to find $\hat{\boldsymbol{\theta}}$ due to complexity of the geometry of $S(\boldsymbol{\theta})$.

Let $\hat{\boldsymbol{\theta}}$ denote the point in $\Re^k$ where $S(\boldsymbol{\theta})$ reaches its global minimum. Suppose $\boldsymbol{\theta}^{(a)}$ is the current estimate of $\boldsymbol{\theta}$, an approximation to $\hat{\boldsymbol{\theta}}$. For $\boldsymbol{\theta}$ close to $\boldsymbol{\theta}^{(a)}$, we begin with approximating the model $f(\boldsymbol{\theta})$ by the first order Taylor series expansion

$$f(\boldsymbol{\theta}) \approx f(\boldsymbol{\theta}^{(a)}) + \mathbf{J}^{(a)}(\boldsymbol{\theta} - \boldsymbol{\theta}^{(a)}), \tag{2.1.4}$$

where

$$\mathbf{J}^{(a)} = \mathbf{J}(\boldsymbol{\theta}^{(a)}) = \left(\left. \frac{\partial f(x_i;\boldsymbol{\theta})}{\partial \theta_j} \right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(a)}} \right), i = 1,2,\ldots n; j = 1,2,\ldots k.$$

and $\mathbf{J}^{(a)}$ is the $n \times k$ matrix of the partial derivatives of $f(\boldsymbol{\theta})$ evaluated at $\boldsymbol{\theta}^{(a)}$. Replacing $f(\boldsymbol{\theta})$ in (2.1.2) with (2.1.4), we have

$$S(\boldsymbol{\theta}) \approx \left[\mathbf{y} - f(\boldsymbol{\theta}^{(a)}) - \mathbf{J}^{(a)}(\boldsymbol{\theta} - \boldsymbol{\theta}^{(a)})\right]' \mathbf{V}^{-1} \left[\mathbf{y} - f(\boldsymbol{\theta}^{(a)}) - \mathbf{J}^{(a)}(\boldsymbol{\theta} - \boldsymbol{\theta}^{(a)})\right].$$

This approximation produces a linear generalized least-squares problem and is thus minimized by

$$\boldsymbol{\theta} - \boldsymbol{\theta}^{(a)} \approx (\mathbf{J}^{(a)\prime} \mathbf{V}^{-1} \mathbf{J}^{(a)})^{-1} \mathbf{J}^{(a)\prime} \mathbf{V}^{-1} \left[\mathbf{y} - f(\boldsymbol{\theta}^{(a)})\right].$$

This leads to the iterative scheme,

$$\boldsymbol{\theta}^{(a+1)} = \boldsymbol{\theta}^{(a)} + (\mathbf{J}^{(a)\prime}\mathbf{V}^{-1}\mathbf{J}^{(a)})^{-1}\mathbf{J}^{(a)\prime}\mathbf{V}^{-1}\left[\mathbf{y} - f(\boldsymbol{\theta}^{(a)})\right].$$

The Gauss-Newton schemes should converge to the value of the least squares estimate minimizing (2.1.2). In practical implementation, updating continues until the difference between two successive iterates of the objective function (2.1.2) is sufficiently small. For details, see Seber and Wild (1989).

The Gauss-Newton-type procedure described above is commonly referred to as iteratively reweighted least squares (IRLS). Charnes, Frome and Yu (1976) showed that if the responses have the distribution from the regular exponential family, then under mild conditions the IRLS estimates are identical to those that would be obtained by using the maximum likelihood principles. More generally, Jorgensen (1983) and Green (1984) discussed the correspondence between maximum likelihood and IRLS.

An additional practical consideration for fitting a nonlinear model is the importance of the parameterization adopted. As Davidian and Giltinan (1995) indicated, the choice of parameterization can speed convergence and improve the quality of the linear approximation used to form tests and confidence intervals. Moreover, an appropriate parameterization may be used to enforce constraints on the values of the parameters, which are critical for the model to make physical sense. It is also worthwhile to try to find a representation of the model where the parameter estimates are approximately independent in order to improve stability and accuracy of estimation. An excellent discussion on this issue can be found in Bates and Watts (1988).

2.2 Local D-Optimality

If we are confident that a given nonlinear model as presented in (2.1.1) is appropriate for a certain experimental situation, then there is the question of designing the experiment, i.e. choosing the $\mathbf{x}_i$ so that the unknown parameters are estimated in an optimal fashion. The experimenter is allowed to choose n points, $x_{1,} \cdots, x_n$ in the design space, χ, and make observations using these points. The optimal design problem is to seek an n-observation design such that the estimator of $\boldsymbol{\theta}$ is optimal in some sense. In general, the search of the optimal design is focused on the minimization of certain criterion functions of the dispersion matrix of the estimated parameters. Various functionals of a variance-covariance matrix have been suggested as criteria for maximization, such as D-optimality, A-optimality, c-optimality, E-optimality, etc. In this dissertaion, we will focus on D-optimality, which is one of the widely used design criteria.

One of the early works that establishes the mathematical properties of optimal experiment design is Chernoff (1953). Fedorov (1972) provides an in-depth mathematical discussion of the theory of optimal experimental designs. Silvey (1980) and Atkinson and Donev (1992) explain most of the common optimality criteria and how they relate to practical design of experiments.

The first real advance in design methodology was the proposal of optimal criterion made in the pioneering paper of Box and Lucas (1959), in which the set of design points are chosen to minimize the determinant of the asymptotic variance-covariance matrix for the parameter vector. This is equivalent to the D-optimal criterion for the linear experiment design. Basically D-optimal designs are constructed to maximize the determinant of the inverse of the dispersion matrix, the Fisher information matrix, namely $\left|\mathbf{J}'(\boldsymbol{\theta})\mathbf{V}^{-1}\mathbf{J}(\boldsymbol{\theta})\right|$. Kiefer and Wolfowitz (1960) established the celebrated Equivalence Theorem in the context of linear models, which has been used as the necessary and sufficient conditions to develop the D-optimal designs. This was further extended by Whittle (1973) to General Equivalence Theorem for linear models to cover more criteria, which has been the key theorem in the optimal design theory since it provides a tool to verify the optimality of a given design measure. We refer to St. John and Draper (1975) for a review of D-optimality.

While there are extensive results on the theory of optimal designs for linear models (see Atkinson(1982), Steinberg and Hunter (1984), Shah and Sinha (1989) for general reviews), the problem is much more difficult and not nearly as well understood for nonlinear models. One reason for this significant research gap is that, for nonlinear models, the optimality criteria involve the unknown parameter $\boldsymbol{\theta}$ as well as the design points $\mathbf{x}_i$. A common approach to solve this dilemma is to use locally optimal designs based on the best guess of the parameters, as defined in Chernoff (1953) and Silvey (1980). This best guess may come from previous experiments or merely a guess. Because the resulted design is optimal under the specific criteria only when the true parameter value is exactly the same as the guess, this approach is called "local" optimality.

Since then many authors have considered locally D-optimal designs, including Ford, Torsney and Wu (1992), Hedayat, Yan and Pezzuto (1997), Antonello and Raghavarao (2000), Han and Chaloner (2003), Hedayat, Zhong and Nie (2004), Dette, Melas and Pepelyshev (2004) and more recently, Dette, Melas and Wong (2005), Biedermann, Dette and Zhu (2006), Wang, Myers, Smith, and Ye (2006), and Li and Majumdar (2008,2009). This dissertation will also focus on locally D-optimal designs. While a good guess may not always be available, locally optimal designs are still invaluable as they provide benchmarks to calibrate other available designs, as noted in Ford, Torsney and Wu (1992). As long as the true parameter values are not too far from the assumed parameters, the locally optimal designs will serve practical needs satisfactorily.

It is well known that for many linear and nonlinear models, the D-optimal designs are usually minimally supported, i.e., the number of support points is equal to the number of parameters in the model. Box and Lucas (1959) restricted their investigation to the case where the number of design points equals the number of parameters. Ford, Torsney and Wu (1992) considered a class of generalized linear models and obtained locally D-optimal designs for nine cases. Among these cases, the designs for two – parameter logistic regression model for the design spaces bounded from one end were only guaranteed to be D-optimal among two-point designs. Hedayat, Yan and Pezzuto (1997) analyzed the raw optical density data by using a family of nonlinear dose-response functions and demonstrated that the locally D-optimal designs for the homogeneous nonlinear dose-response models are minimally supported. Later, Han and Chaloner (2003) derived locally D-optimal designs that are minimally supported in exponential regression models used in viral dynamics. Hedayat, Zhong and Nie (2004) identified classes of two-parameter nonlinear models for which a locally D-optimal design is exactly supported on two-points. Li and Majumdar (2008, 2009) showed locally D-optimal designs for logistic models with three and four parameters, a one-compartment pharmacokinetic model and a Poisson regression model are minimally supported.

2.3 Optimal Design for Goodness-of-Fit

A major disadvantage of having minimally supported designs, in which the number of design points coincides with the number of parameters in the regression model, is that the adequacy of the model cannot be investigated as there are no additional degrees of freedom available for testing the goodness-of-fit of the assumed model. Checking the model assumptions is of particular importance when the model is selected without complete certainty, because optimal designs are model-dependent and a misspecified model could lead to substantial loss of efficiency. Refer to Khuri, Mukherjee, Sinha and Ghosh (2006) for a recent review on nonlinear design issues.

While the design problem for goodness-of-fit test has been discussed by some authors for linear models, there has been much less attention paid on this design issue for nonlinear models. As indicated in Dette, Melas and Wong (2005), Lupinacci and Raghavarao (2000, 2003) appear to be among the first to address such design issue for nonlinear models. Lupinacci and Raghavarao (2000) proposed adding an extra point of 0.5 for the D-optimal design for the two-parameter logistic model so that the symmetry in the variance function could be ensured. Using similar rationale, Lupinacci and

Raghavarao (2003) proposed that a third point be added midway between the minimally supported D-optimal design for the two-parameter Michaelis-Menten model.

Since then, in the context of nonlinear models, only a modest amount of attention has been focused on this design issue. Recently, Dette, Melas and Wong (2005) and Dette and Pepelyshev (2008) discuss an alternative approach to derive designs to perform a lack-of-fit test for nonlinear models. In order to find designs which are good for both estimation and goodness-of-fit testing, these two papers used a larger class of plausible nested models to embed the model of interest. The idea of embedding the postulated model in a larger class is not new. Some suggestions can be found for linear models in Stigler (1971) or Studden (1982) for a polynomial regression model, in which they suggested to embed the postulated model in an extended model, usually a polynomial of larger degree, and to construct an optimal design for testing the postulated model against the extended model. Atkinson (1972) proposed embedding the model in a larger class of models and then designing to estimate the additional parameters as precisely as possible (see also Atkinson and Cox, 1974). Pukelsheim and Rosenberger (1993) also computed D-optimal designs for the largest model for discrimination purposes. These D-optimal designs usually have more support points than the number of parameters in the original models, therefore a lack of fit test can be performed.

Dette, Melas and Wong (2005) constructed designs for the Michaelis-Menten enzyme kinetic model by embedding the model into the more general EMAX model and computing maximin D- and maximin c-optimal designs for the EMAX model. They used a maximin approach, so that they can have designs to be used for testing the postulated model assumptions with reasonable efficiency, and also be robust with respect to misspecification of the unknown parameters in the assumed model.

Furthermore, Dette and Pepelyshev (2008) applied the same approach to two models: the exponential regression model (where the extension is Weibull model) and the three-parameter logistic regression model (where the extension is the Richards regression model). They constructed the D- and D_1-optimal designs that both can estimate the parameters efficiently and detect a lack of fit for misspecification of the models.

Despite the attempt of these two papers to establish a general theoretical framework for constructing optimal designs capable of performing a lack-of-fit test for misspecification in a nonlinear model, it is still not clear how good a D-optimal design is in general for discrimination purposes. In this dissertation, we will extend the work of Lupinacci and Raghavarao (2000, 2003) to most commonly used nonlinear models, with the focus on sigmoid curves, which includes logistic, probit and Gompertz models. Each of these models will be discussed with two, three or four parameters.

CHAPTER 3
TESTING LACK OF FIT IN SIGMOID CURVE MODELS: ONE EXTRA POINT

3.1 Parameterization of Sigmoid Curves

In this section, we introduce specific parameterizations of several familiar sigmoid curves. Seber and Wild (1989) relate “sigmoidal growth models” to their growth mechanism, which is a differential equation that links the “growth rate”, i.e., the first derivative (with respect to the independent variable), to “current size”. We choose to describe sigmoid curves on geometric terms, which accommodates both increasing and decreasing functions and allows flexible asymptotes.

For each of the sigmoid curves we describe in this dissertaion, there are many different algebraic forms to express the same mathematical model. Choosing the proper form to study is important since parameterization can impact the statistical inference by inducing or reducing parameter-effect curvature (See Bates and Watts ,1980; Ratkowsky, 1983; Seber and Wild, 1989). However, without specification of further details, it is impossible to determine which parameterization is desirable. We choose to parameterize all the sigmoid curves in such a way that certain geometric interpretations of the parameters are maintained across the models.

Let x denote the independent variable, $\boldsymbol{\theta}$ denote the four parameters vector (a,b,c,d), and $y = f(x;\boldsymbol{\theta})$. We impose the following conditions on $x, y,$ and $\boldsymbol{\theta}$:

A. The curve is sigmoid when y is plotted against x;

B. When $x = c$, $y = (a+d)/2$;

C. When $b > 0$, d is the left asymptote ($y \to d$, as $x \to -\infty$) and a is the right asymptote ($y \to a$, as $x \to +\infty$);

D. When $b < 0$, a is the left asymptote ($y \to a$, as $x \to -\infty$) and d is the right asymptote ($y \to d$, as $x \to +\infty$);

E. y is a function of x through $b(x-c)$.

In practice, most of the time the curve is sigmoid when the response or signal y is plotted against log dose, or log concentration, or log dilution, which we denote as the independent variable x. The slope, b, governs the slope of the curve. The parameters d and a represent the minimum and maximum response levels, respectively. The parameter c is commonly referred to as log ED_{50}, log EC_{50}, or log IC_{50}, which gives a response midway between a and d. By calling c mid point, a purely geometric term, we avoid confusion when y is in the scale of percent, say, percent inhibition, and $(a+d)/2$ is not 50%.

We have not imposed any conditions on the relationship between a and d. In criteria C and D, a and d are described as left or right asymptotes, but not lower or upper asymptotes. As we'll see that for some sigmoid curves, the condition of $a > d$ (or $a < d$) is needed to resolve the identifiability problem; for others, however, either condition could arbitrarily eliminate half of the model. All the models discussed below can be re-written by replacing $b(x-c)$ with $(x-c)/b$. We expect that they produce very different parameter-effect curvature when b is near zero as compared to $|b| >> 1$. However, when b is close to -1 or 1, we expect them to behave similarly.

A sigmoid curve is symmetric if and only if its first derivative (or "growth rate") is an even function centered at the mid point c:

$$\left.\frac{\partial f}{\partial x}\right|_{x=c-u} = \left.\frac{\partial f}{\partial x}\right|_{x=c+u}, \quad \text{for any } u \in \Re .$$

In other words, the rates of change are the same when x is equi-distant from either side of the mid point. A necessary, but not sufficient, condition for symmetry is that the inflection point, where $\partial f / \partial x$ reaches a (local) minimum or maximum, is unique and coincides with the mid point c. The growth mechanism interpretation of inflection point is the time where the growth rate is the greatest.

For all the sigmoid models discussed below, we will first provide the general form with the full set of four parameters, b, c, a and d. The reduced form with fewer parameters can be easily derived. The two-parameter models only have the parameters b and c, which can be easily produced by taking $d = 0$ and $a = 1$ from the four-parameter form. If we assume $d = 0$, then the four-parameter model will be reduced to the three-parameter form, which contains the parameters b, c and a.

Optimal design theory is relevant to maximization of certain concave criterion function of information matrix over the set of the information matrices. When the information matrix is independent of the unknown parameters, it is straightforward to maximize the criterion function over the set of the information matrices to find the optimal design. One reason that nonlinear design problems are so challenging is that, with nonlinear models, information matrices and optimal designs depend on the unknown model parameters. By doing the parameter transformation, we found a strategy to bypass the obstacle of parameter dependence for the theoretical derivation.

Let us denote P as the standardized y, which is $(y-d)/(a-d)$; and let $t = b(x-c)$, the linearized x. For all the sigmoid models in Chapter 3 and 4, the relationship between P and t can be established by a linearizing function, $G^{-1}(.)$, expressed as

$$P_i = G(t_i), \qquad (3.1.1)$$

or

$$t_i = G^{-1}(P_i), \qquad (3.1.2)$$

where $G(.)$ or $G^{-1}(.)$ has different expression for different models. In the cases we will discuss in Sections 3.3 through 3.5, $G^{-1}(.)$ is a logit function for the logistic model, a probit function for the probit model, and a log-log function for the Gompertz model. This transformation makes the determinant of the information matrix a function of

(a,b,c,d) only through P and t, which enables the search for the optimal points of P_i's completely independent of these parameters.

With the transformations indicated in (3.1.1) and (3.1.2), the nonlinear model in (2.1.1) can be written as

$$y_i = d + (a-d)G(t_i) + \varepsilon_i \qquad (i = 1,2,\ldots n).$$

By taking the derivative of y_i with respect to a,b,c,d, we get the design matrix X with its i^{th} row as

$$\left[G_i \quad (a-d)g_i(x_i - c) \quad -(a-d)g_i b \quad 1-G_i\right],$$

where $g_i = g(t_i) = \dfrac{\partial G(t_i)}{\partial t_i}$. $Det(X'X)$ is then given by

$$\det(X'X) \propto \begin{vmatrix} \sum_{i=1}^{n} G_i^{\,2} & \sum_{i=1}^{n} t_i G_i g_i & -\sum_{i=1}^{n} G_i g_i & \sum_{i=1}^{n} G_i \\ \sum_{i=1}^{n} t_i G_i g_i & \sum_{i=1}^{n} t_i^{\,2} g_i^{\,2} & -\sum_{i=1}^{n} t_i g_i^{\,2} & \sum_{i=1}^{n} t_i g_i \\ -\sum_{i=1}^{n} G_i g_i & -\sum_{i=1}^{n} t_i g_i^{\,2} & \sum_{i=1}^{n} g_i^2 & -\sum_{i=1}^{n} g_i \\ \sum_{i=1}^{n} G_i & \sum_{i=1}^{n} t_i g_i & -\sum_{i=1}^{n} g_i & n \end{vmatrix}, \qquad (3.1.3)$$

where n denotes the number of points in the design. The determinant in (3.1.3) is the determinant of the Fisher information matrix for the four-parameter sigmoid curve

($k = 4$). The determinant of the information matrix for the three-parameter models ($k = 3$) can be obtained by crossing out the fourth column and the fourth row in (3.1.3). Deleting the first and the fourth row, and the first and the fourth column from (3.1.3) leads to the determinant of the information matrix for the two-parameter models ($k = 2$), which is similar to what was presented in Hedayat, Zhong and Nie (2004), as well as Yang and Stufken (2009).

Optimal designs for sigmoid curves are often specified on the scale of P_i or t_i. In that case, x_i can be determined based on the best available information on the parameters. This produces locally optimal designs, as defined in Chernoff (1953). As long as the true parameter values are not too far from the assumed parameters, the locally optimal designs will satisfactorily serve practical needs.

3.2 Designs for Testing Lack of Fit: Minimal Plus One Point

In recent years, optimal designs for nonlinear models have been studied by many authors. Most optimal nonlinear designs are often minimally supported in that the number of distinct support points is equal to the number of unknown parameters in the model. These optimal designs have the shortcoming of not checking the model assumptions as there are no additional degrees of freedom available to test lack of fit of the model. When the model function is not known with complete knowledge, a near-optimal design for the assumed model but which contains extra design points that can be used to test for the lack of fit would be more desirable. In this section, we will introduce one extra point to the minimal D-optimal design so that the lack-of-fit test can be carried out after comparing different choices for the extra design point.

The framework of this section is organized as follows. For each k - parameter sigmoid model, we will first present the structure of its Fisher information matrix, then derive the minimally supported (k -point) D-optimal design at which $\det(X'X)$ is maximized. In order to test the lack of fit, we will add an extra distinct point to the design. For each kind of sigmoid curves, five different approaches will be discussed for the k -parameter, $(k+1)$ -point design:

I. $(k+1)$ -point, D-optimal design without any restriction;

II. D-optimal design with the extra $(k+1)$ th point, P_{k+1}, restricted at 0.5;

III. Take the minimally supported D-optimal designs of the k - parameter model for the first k support points, then add the extra $(k+1)$ th point, P_{k+1}, as the average of the first k points;

IV. Similar to III by taking the minimally supported D-optimal designs of the k - parameter model for the first k support points, but simply take $P_{k+1} = 0.5$;

V. D-optimal design where the design points are equally spaced: to find the best $(k+1)$ -point equally spaced design, we define

$t_1 = s_1$, $t_2 = s_1 + \frac{1}{k}(s_{k+1} - s_1)$, $t_3 = s_1 + \frac{2}{k}(s_{k+1} - s_1)$, …,

$t_n = s_{k+1}$, and $s_1 < s_{k+1}$. The corresponding $\det(X^{'}X)$ is then a function of s_1 and s_{k+1}, which can be used to derive P_1 and P_{k+1}, as well as the rest of the design points.

Apparently, among all these five designs, Design I will have the highest value of $\det(X^{'}X)$, since it is obtained through the D-optimal criterion with no restriction. As shown in the results below, Design I also gives distinct design points for the four-parameter models to allow a test for lack of fit; however, for two and three parameter families, this design usually ends up as the minimally supported D-optimal design plus one of the design points replicated. Hence Design I is not useful to test the lack of fit of the model for two and three parameter families. To serve the purpose of testing the model adequacy, the extra design point should be different from the minimally supported D-optimal design points.

In order to choose this extra point for the models we study in this chapter, we start with the graphical display of the functions of $\det(X'X)$ on P_{k+1}, when the first k support points take the values from the minimally supported D-optimal designs of the k-parameter model. The plots are given in Figure 3.1, Figure3.2, Figure 3.3 and Figure 3.4. Except for the slight asymmetry of the Gompertz models, the three plots are somewhat similar as indicated below:

- Saddle-shaped curves;

- The three-parameter curve is identical to the two-parameter curve before reaching the first peak where the curve attains the maximum; it is also identical to the four-parameter curve after reaching the second peak. The part of the curve between these two peaks is sandwiched by the two-parameter curve and the four-parameter curve, with the two-parameter curve at the bottom.

From the plots, a value in the middle of these known design points that are not 0 and 1 seems to be a natural choice for the extra point, since it is away from the established design points yet still gives relatively high value for $\det(X'X)$. Hence the middle point will be taken as 0.5 or at the midway of other design points, which is adopted in Design IV and III, respectively.

Design III and IV are not D-optimal designs. The question now is how they measure up to the D-optimal design I. To answer this question, we add Design II, a modification of Design IV where the D-optimal design is derived by taking the $(k+1)^{th}$ point as 0.5. Biological and chemical experiments frequently use equally spaced designs due to the simplicity, therefore we investigate Design V to best understand its use in obtaining a good lack-of-fit test. The rationales of choosing and testing these five designs will be discussed in more detail in the following sections.

To compare the designs listed above, we will calculate the D-efficiency by using Design I, the D-optimal design without any restriction, as the benchmark. D- efficiency is the ratio of the determinant of $(X'X)$ of any design to that of Design I.

3.3 Logistic Models

Lupinacci and Raghavarao (2000) identified the design issue related to goodness-of-fit for nonlinear models, and introduced the approach to estimate the lack of fit for a two-parameter logistic function with binomial variation. In this section, we will use constant variance at all design points, and expand the study to two-, three- and four-parameter logistic models.

We will start with the four-parameter logistic (4PL) model, which is the most commonly used model in curve fitting analysis in bioassays or immunoassays such as ELISAs or dose-response curves. The following is a form of 4PL curve:

$$y_i = f(x_i; a, b, c, d) = d + \frac{a-d}{1+e^{-b(x_i-c)}}. \qquad (3.3.1)$$

Since $f(x; a, b, c, d)$ is the same curve as $f(x; a, -b, c, d)$, the condition of $a > d$ (or $a < d$) is needed to resolve the indentifiability problem. 4PL of (3.3.1) is symmetric. It is called logistic curve because y, or more precisely, P, the standardized y, as defined in Section 3.1, can be linearized through the logit function:

$$\text{logit}\left(\frac{y-d}{a-d}\right) = b(x-c),$$

which is equivalent to

$$t_i = G^{-1}(P_i) = \text{logit}(P_i) = \log\left(\frac{P_i}{1-P_i}\right).$$

For the 4PL model,

$$P_i = G(t_i) = \frac{1}{1+e^{-t_i}}.$$

By taking the partial derivatives of P_i with respect to a, b, c and d, we get the design matrix

$$X = \begin{bmatrix} X_1' \\ \vdots \\ X_n' \end{bmatrix} = \begin{bmatrix} G_1 & \frac{a-d}{b} t_1 g_1 & -(a-d) b g_1 & 1-G_1 \\ \vdots & \vdots & \vdots & \vdots \\ G_n & \frac{a-d}{b} t_n g_n & -(a-d) b g_n & 1-G_n \end{bmatrix},$$

where $t_i = b(x_i - c)$, $G_i = P_i = \dfrac{1}{1+e^{-t_i}}$ and $g_i = g(t_i) = \dfrac{e^{-t_i}}{(1+e^{-t_i})^2}$.

The Fisher information matrix is

$$X'X \propto \begin{bmatrix} \sum_{i=1}^{n} G_i^2 & \frac{a-d}{b}\sum_{i=1}^{n} t_i G_i g_i & -(a-d)b\sum_{i=1}^{n} G_i g_i & \sum_{i=1}^{n} G_i \\ \frac{a-d}{b}\sum_{i=1}^{n} t_i G_i g_i & \frac{(a-d)^2}{b^2}\sum_{i=1}^{n} t_i^{\ 2} g_i^{\ 2} & -(a-d)^2\sum_{i=1}^{n} t_i g_i^{\ 2} & \frac{a-d}{b}\sum_{i=1}^{n} t_i g_i \\ -(a-d)b\sum_{i=1}^{n} G_i g_i & -(a-d)^2\sum_{i=1}^{n} t_i g_i^{\ 2} & (a-d)^2 b^2 \sum_{i=1}^{n} g_i^2 & -(a-d)b\sum_{i=1}^{n} g_i \\ \sum_{i=1}^{n} G_i & \frac{a-d}{b}\sum_{i=1}^{n} t_i g_i & -(a-d)b\sum_{i=1}^{n} g_i & n \end{bmatrix}.$$

Therefore, the determinant of the Fisher information matrix can be transformed to

$$\det(X'X) \propto (a-d)^4 \begin{vmatrix} \sum_{i=1}^{n} G_i^2 & \sum_{i=1}^{n} t_i G_i g_i & -\sum_{i=1}^{n} G_i g_i & \sum_{i=1}^{n} G_i \\ \sum_{i=1}^{n} t_i G_i g_i & \sum_{i=1}^{n} t_i^{\ 2} g_i^{\ 2} & -\sum_{i=1}^{n} t_i g_i^{\ 2} & \sum_{i=1}^{n} t_i g_i \\ -\sum_{i=1}^{n} G_i g_i & -\sum_{i=1}^{n} t_i g_i^{\ 2} & \sum_{i=1}^{n} g_i^2 & -\sum_{i=1}^{n} g_i \\ \sum_{i=1}^{n} G_i & \sum_{i=1}^{n} t_i g_i & -\sum_{i=1}^{n} g_i & n \end{vmatrix},$$

which can be futher simplified as (3.1.3), and the D-optimal design could be obtained by maximizing this determinant.

The same basic approach can be applied to the two- (2PL) and three-parameter logistic (3PL) models, as described in Section 3.1. Table 3.1 is the D-efficiency summary for the k-parameter, n-point logistic designs, where k =2, 3 or 4, $n = k$ for the minimally supported design, and $n = k + 1$ for the design with an extra point to test the lack of fit. We will discuss the results for the 2PL in more details, since it is the model that has received the most attention in the literature.

Table 3.1. Designs and D-Efficiencies for Two-, Three- and Four-Parameter Logistic Models: Minimal Plus One Point

		Two-parameter Logistic Model ($k=2$)			Three-parameter Logistic Model ($k=3$)			Four-parameter Logistic Model ($k=4$)		
Design	Support Points, n	Design Points	$\det(X'X)$	Efficiency (%)	Design Points	$\det(X'X)$	Efficiency (%)	Design Points	$\det(X'X)$	Efficiency (%)
Minimal D-optimal Design	k	0.2605, 0.7396	0.0060		0.2605, 0.7396, 1	0.0060		0, 0.2605, 0.7396, 1	0.0060	
<u>Designs for lack-of-fit test</u>										
I	$k+1$	0.2605, 0.7396, 0.2605 or 0.7396	0.0120	100	0.2551, 0.6773, 0.7859, 1	0.0120	100	0, 0.2427, 0.6267, 0.8006, 1	0.0121	100
II	$k+1$	0.2271, 0.5, 0.7729	0.0115	95.83	0.2243, 0.5, 0.7830, 1	0.0117	97.5	0, 0.2136, 0.5, 0.7864, 1	0.0120	99.17
III	$k+1$	0.2605, 0.5, 0.7396	0.0110	91.67	0.2605, 0.6667, 0.7396, 1	0.0118	98.33	0, 0.2605, 0.5, 0.7396, 1	0.0113	93.39
IV	$k+1$	0.2605, 0.5, 0.7396	0.0110	91.67	0.2605, 0.5, 0.7396, 1	0.0112	93.33	0, 0.2605, 0.5, 0.7396, 1	0.0113	93.39
V	$k+1$	0.2271, 0.5, 0.7729	0.0115	95.83	0.2352, 0.7605, 0.9704, 0.9971	0.0095	79.17	0.0010, 0.0161, 0.2083, 0.8091, 0.9856	0.0064	52.89

Antonello and Raghavarao (2000) found the two-parameter, two-point D-optimal design under constant variation, which confirmed the result in Kalish and Rosenberger (1978). They showed that the optimal design points were 0.260 and 0.740 by maximizing the determinant of $X^{'}X$. Later, in his unpublished Ph.D. dissertation, Lupinacci (2001) proved that the three-point D-optimal design is the two-point D-optimal design derived by Antonello and Raghavarao (2000) with one of the design points replicated. This is the same result we obtained for Design I defined in Section 3.2., by maximizing (3.1.3) when $a = 1$, $d = 0$. As noted before, this design will undoubtedly have the highest value of $\det(X^{'}X)$, but lacks a distinct point besides the D-optimal points to test the lack of fit.

To choose this extra point, P_3, we first plug in the values from the minimally supported D-optimal designs (in this case, $P_1 = 0.2605, P_2 = 0.7396$) in (3.3.1), then plot this reduced functions of $\det(X'X)$ on P_3 (see Figures 3.3.1). The curve is saddle-shaped, with the two peaks sitting at the minimally support D-optimal design points, 0.2605 and 0.7396.

Figure 3.1 illustrates the choice of the extra point in Designs III and IV. Designs I, II and V are also evaluated for the purpose of comparison in Table 3.1. As shown in Table 3.1, Design II leads in all the comparisons, except in the three-paramenter model. The equally spaced design V ties with Design II in terms of D-efficiency at the first place in the two-parameter model; but it has relatively smaller efficiency in the three-parameter model, and rather poor efficiency in the four-parameter model. Design III has the highest efficiency in the three-parameter model, where the efficiency of Design II and IV are also high. For the two- and four-parameter models, Design III and IV have the same efficiencies, where the drop in efficiency from Design I is negligible.

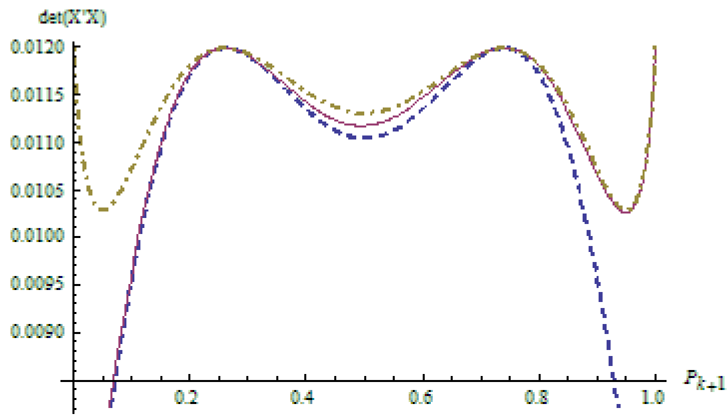

Figure 3.1. Determinant of $(X'X)$ of k-parameter, $(k+1)$-point logistic function on P_{k+1}, when the first k support points take the values from the minimally supported D-optimal designs of the k- parameter model: two-parameter model(k =2,dashes), three-parameter model(k =3, solid), and four-parameter model(k =4,dot-dash).

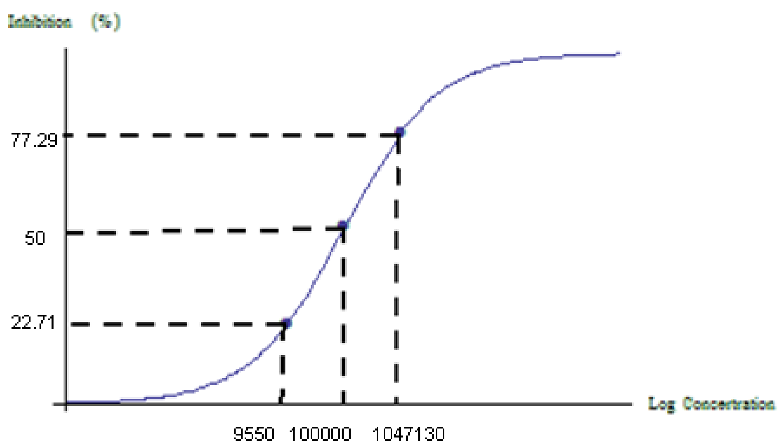

Figure 3.2. Example of Design II in a two-parameter logistic (2PL) model ($b=1.2, c=5$).

An example in vaccine clinical assay is used to illustrate the application of Table 3.1 for two-parameter logistic models. One way to evaluate vaccine efficacy in humans is to measure the functional activity of the antibody induced by vaccine via competitive inhibition. In such assays, specific amount of target antigen is coated on all the wells on a 96-well microplate. Then fixed amount of free and tagged antibody are added to all the wells; and the florescence light is emitted when the antigen binds to the antibody. When sera from vaccinated subjects are added to the wells, the active antibody will competitively bind to the coated antigen, hence reduces the light intensity. In the experiment, a highly concentrated reference standard with known amount of active antibody is first prepared. Then a serial dilution of the reference will generate a two-parameter logistic standard curve with percentage of inhibition versus log concentration. When unknown samples are tested along the standard curve, the concentrations of active antibody in the sample can be estimated by comparing the percentage of inhibition of the sample to that of the standard curve. For example, in practice we know that b is near 1.2 and c is near 5 from previous experiments, we would like to get the best concentrations (or dilution) of the reference for estimating the standard curve. If we decide to use Design II, we will know from Table 3.1 that, the optimal concentration are 9550 ng/ml, 100000 ng/ml and 1047130 ng/ml which correspond to 22.71%, 50% and 77.29% of inhibition, respectively. This example is illustrated in Figure 3.2.

3.4 Probit Models

In this section we will consider the designs for probit models. Similar to logistic models, probit models are popular specifications to analyze many kinds of dose-response or binomial response experiments in a variety of fields.

The form of four-parameter probit (4PP) curve is:

$$y_i = f(x_i; a, b, c, d) = d + (a-d)\Phi[b(x_i - c)] \tag{3.4.1}$$

where Φ is the cumulative distribution function of the standard normal distribution. 4PP of (3.4.1) is symmetric. Once again, $f(x; a, b, c, d)$ is the same curve as $f(x; a, -b, c, d)$, the condition of $a > d$ (or $a < d$) is needed to resolve the identifiability problem. It is called probit curve because the probit function, i.e., Φ^{-1}, can linerize P, the standardized y:

$$\Phi^{-1}\left(\frac{y-d}{a-d}\right) = b(x-c),$$

which brings

$$t_i = G^{-1}(P_i) = \Phi^{-1}(P_i).$$

when we use the probit function as the linearizing function $G^{-1}(\cdot)$, and $t_i = b(x_i - c)$.

For the 4PP model,

$$P_i = G(t_i) = \Phi[b(x_i - c)].$$

The design matrix can be obtained by taking the partial derivatives of P_i with respect to a, b, c and d:

$$X = \begin{bmatrix} X_1' \\ \vdots \\ X_n' \end{bmatrix} = \begin{bmatrix} \Phi_1 & (a-d)(x_1 - c)\phi_1 & -(a-d)b\phi_1 & 1-\Phi_1 \\ \vdots & \vdots & \vdots & \vdots \\ \Phi_n & (a-d)(x_n - c)\phi_n & -(a-d)b\phi_n & 1-\Phi_n \end{bmatrix} = \begin{bmatrix} G_1 & \frac{a-d}{b} t_1 g_1 & -(a-d)bg_1 & 1-G_1 \\ \vdots & \vdots & \vdots & \vdots \\ G_n & \frac{a-d}{b} t_n g_n & -(a-d)bg_n & 1-G_n \end{bmatrix},$$

where $t_i = b(x_i - c)$, $G_i = P_i = \Phi(t_i)$, $g_i = \phi_i = \phi(t_i)$, and ϕ is the probability density function of standard Normal.

The determinant of the information matrix for the four-parameter probit (4PP) model is (3.1.3) with the specifications on t_i, G_i and g_i defined above. The minimally supported D-optimal designs for the 4PP can be derived from maximizing the above determinant. The designs for the two- (2PP) and three-parameter probit (3PP) models can be found numerically in the same way, after simplifying (3.1.3) for the corresponding models by following the method discussed in Section 3.1. The minimally supported D-optimal designs for two-, three- and four-parameter probit models are listed in Table 3.2, along with the comparisons of other designs.

Table 3.2. Designs and D-Efficiencies for Two-, Three- and Four-Parameter Probit Models: Minimal Plus One Point

Design	Support Points, n	Two-parameter Probit Model ($k=2$)			Three-parameter Probit Model ($k=3$)			Four-parameter Probit Model ($k=4$)		
		Design Points	$\det(X'X)$	Efficiency (%)	Design Points	$\det(X'X)$	Efficiency (%)	Design Points	$\det(X'X)$	Efficiency (%)
Minimal D-optimal Design	k	0.2398, 0.7603	0.0186		0.2398, 0.7603, 1	0.0186		0, 0.2398, 0.7603, 1	0.0186	
<u>Designs for lack-of-fit test</u>										
I	$k+1$	0.2398, 0.7603, 0.2398 or 0.7603	0.0373	100	0.2398, 0.7603, 0.2398 or 0.7603, 1	0.0373	100	0, 0.2357, 0.7087, 0.7997, 1	0.0373	100
II	$k+1$	0.2077, 0.5, 0.7923	0.0351	94.10	0.2054, 0.5, 0.8014, 1	0.0359	96.25	0, 0.1959, 0.5, 0.8041, 1	0.0367	98.39
III	$k+1$	0.2398, 0.5, 0.7603	0.0340	91.15	0.2398, 0.6667, 0.7603, 1	0.0348	93.30	0, 0.2398, 0.5, 0.7603, 1	0.0348	93.30
IV	$k+1$	0.2398, 0.5, 0.7603	0.0340	91.15	0.2398, 0.5, 0.7603, 1	0.0344	92.23	0, 0.2398, 0.5, 0.7603, 1	0.0348	93.30
V	$k+1$	0.2077, 0.5, 0.7923	0.0351	94.10	0.2259, 0.7857, 0.9902, 1	0.0328	87.94	0.0036, 0.1795, 0.8035, 0.9957, 1	0.0272	72.92

The results show a generally consistent pattern: in terms of D-efficiency, across two-, three- and four-parameter probit designs, Design II > Design III ≥ Design IV > Design V. The only exception is that, in the 2PP model, the D-efficiency of Design V is identical to Design II, therefore also has the highest efficiency in that case.

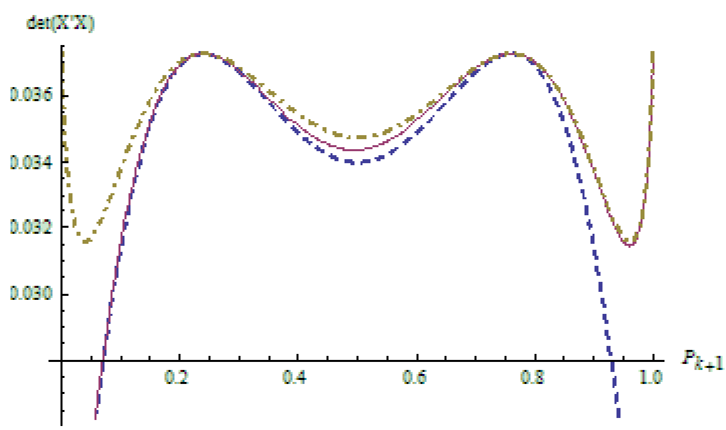

Figure Figure 3.3. Determinant of $(X'X)$ of k-parameter, $(k+1)$-point probit function on P_{k+1}, when the first k support points take the values from the minimally supported D-optimal designs of the k- parameter model. --- for $k=2$, — for $k=3$, and –·–·– for $k=4$.

The choice of the extra point in Design III and IV is displayed in Figure Figure 3.3. It is clear, from Table 3.2, that although not D-optimal designs, the drop in efficiency for Design III and IV, which are calculation-free and easily implemented, is negligible.

Note that, for two-parameter Probit model (2PP), the three-point D-optimal design (Design I) is simply the minimal D-optimal design plus one of the design points replicated. The proof of this fact under constant variation is provided in the Appendix.

3.5 Gompertz Models

One sigmoid curve which is an asymmetric extension of the logistic/probit model is Gompertz model. The Gompertz function has been widely used in fitting growth curve, especially in tumor growth. This function was defined by Gompertz (1825), and was applied as a growth function by Winsor (1932). Gompertz function has gained wide acceptance as an applicable function in a number of biological systems, as indicated in Kidwell et al. (1969), Laird (1964), Rygaard and Spang-Thomsem (1997).

The form of the four-parameter Gompertz (4PG) function satisfying conditions of A-E in Section 3.1 is

$$y_i = f(x_i; a, b, c, d) = d + \frac{a-d}{2^{\exp[-b(x_i - c)]}}. \tag{3.5.1}$$

We have found that (3.5.1) has significantly better numerical properties, in terms of both convergence proportion and rate, than other forms in which the mid point is not captured by a particular parameter.

A log-log function linearizes the standardized y:

$$-\log\left(-\log_2\left(\frac{y-d}{a-d}\right)\right) = b(x-c)$$

which leads to

$$t_i = G^{-1}(P_i) = -\log(-\log_2(P_i)).$$

For the 4PG model,

$$P_i = G(t_i) = \frac{1}{2^{\exp[-b(x_i-c)]}} = \frac{1}{2^{\exp(-t_i)}}.$$

The design matrix, by taking the partial derivatives of P_i with respect to a,b,c,d, is

$$X = \begin{bmatrix} X_1' \\ \vdots \\ X_n' \end{bmatrix} = \begin{bmatrix} \frac{1}{2^{\exp[-b(x_1-c)]}} & (a-d)(x_1-c)\frac{e^{-b(x_1-c)}\log 2}{2^{\exp[-b(x_1-c)]}} & -(a-d)b\frac{e^{-b(x_1-c)}\log 2}{2^{\exp[-b(x_1-c)]}} & 1-\frac{1}{2^{\exp[-b(x_1-c)]}} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{1}{2^{\exp[-b(x_n-c)]}} & (a-d)(x_n-c)\frac{e^{-b(x_n-c)}\log 2}{2^{\exp[-b(x_n-c)]}} & -(a-d)b\frac{e^{-b(x_n-c)}\log 2}{2^{\exp[-b(x_n-c)]}} & 1-\frac{1}{2^{\exp[-b(x_n-c)]}} \end{bmatrix}$$

$$= \begin{bmatrix} G_1 & \frac{a-d}{b}t_1 g_1 & -(a-d)bg_1 & 1-G_1 \\ \vdots & \vdots & \vdots & \vdots \\ G_n & \frac{a-d}{b}t_n g_n & -(a-d)bg_n & 1-G_n \end{bmatrix}, \qquad (3.5.2)$$

where $t_i = b(x_i - c)$, $G_i = \frac{1}{2^{e^{-t_i}}}$, and $g_i = \frac{e^{-t_i}\log 2}{2^{e^{-t_i}}}$.

With these specifications on t_i, G_i and g_i for the 4PG model, the determinant of the information matrix is (3.1.3). The determinants of the information matrix for the 3PG model and the 2PG model are also (3.1.3) without corresponding row(s) and column(s) as described in Section 3.1. D-optimal designs are obtained by maximizing the determinant of the Fisher information matrix.

Figure 3.4. shows the aymmetry of the Gompertz model, and presents evidence for choosing an extra point at the midway, or 0.5, for Design III and IV, respectively.

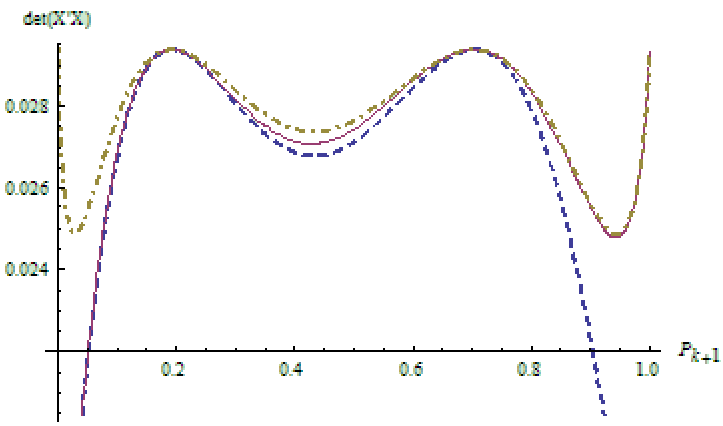

Figure 3.4. Determinant of $(X'X)$ of k-parameter, $(k+1)$-point Gompertz function on P_{k+1}, when the first k support points take the values from the minimally supported D-optimal designs of the k- parameter model: two-parameter model(k =2,dashes), three-parameter model(k =3, solid), and four-parameter model(k =4,dot-dash).

The design comparison results are given in Table 3.3, which has a lot of similarity with the findings in the logistic models: Design II has the highest efficiency, except for the three-parameter model where Design III is the best; the equally spaced design, V, is fairly efficient in the two-parameter model, but ranks the last in the efficiency contest for the three- and four-parameter models.

In the logistic and probit models, Design III and IV are identical in terms of the design points and efficiency for the two- and four-parameter cases. This is because both logistic and probit models are symmetric. Gompertz models are different in this aspect. Design II always has higher efficiency than Design IV, no matter how many parameters the model has. However, Design II needs specific calculation to obtain, not as easily implemented as Design IV.

Similar to the two-parameter logistic (2PL) and probit (2PP) model under the assumption of constant variation, the three-point D-optimal design (Design I) for the two-parameter Gompertz model is the minimally supported D-optimal design with one of the design points replicated. We will give the proof to this observation in the Appendix.

Table 3.3. Designs and D-Efficiencies for Two-, Three- and Four-Parameter Gompertz Models: Minimal Plus One Point

		Two-parameter Gompertz Model ($k=2$)			Three-parameter Gompertz Model ($k=3$)			Four-parameter Gompertz Model ($k=4$)		
Design	Support Points, n	Design Points	$\det(X'X)$	Efficiency (%)	Design Points	$\det(X'X)$	Efficiency (%)	Design Points	$\det(X'X)$	Efficiency (%)
Minimal D-optimal Design	k	0.1925, 0.7032	0.0147		0.1925, 0.7032, 1	0.0147		0, 0.1925, 0.7032, 1	0.0147	
<u>Designs for lack-of-fit test</u>										
I	$k+1$	0.1925, 0.7032, 0.1925 or 0.7032	0.0294	100	0.1925, 0.7032, 0.1925 or 0.7032, 1	0.0294	100	0, 0.1589, 0.2371, 0.7076, 1	0.0294	100
II	$k+1$	0.1710, 0.5, 0.7430	0.0279	94.90	0.1699, 0.5, 0.7591, 1	0.0286	97.28	0, 0.1653, 0.5, 0.7629, 1	0.0290	94.83
III	$k+1$	0.1925, 0.4478, 0.7032	0.0271	92.18	0.1925, 0.6319, 0.7032, 1	0.0290	98.64	0, 0.1925, 0.4739, 0.7032, 1	0.0275	93.54
IV	$k+1$	0.1925, 0.5, 0.7032	0.0268	91.16	0.1925, 0.5, 0.7032, 1	0.0274	93.20	0, 0.1925, 0.5, 0.7032, 1	0.0276	93.88
V	$k+1$	0.1743, 0.4942, 0.7525	0.0278	94.56	0.1770, 0.7591, 0.9571, 0.9930	0.0197	67.01	0, 0.1751, 0.7598, 0.9576, 0.9932	0.0197	67.01

For the sigmoid models we discussed in this chapter, we found that equally spaced designs have high efficiency for the two-parameter models, but should be avoided for the three- and four-parameter models. In general, the $(k+1)$-point D-optimal design without any restriction or with one point restricted at 0.5 performs the best for the k-parameter models; but these design points can only be obtained through specific calculations. Design IV and III can be easily implemented in practice by adding an extra point of 0.5 or the average to the points obtained from the minimally supported D-optimal design, respectively. Although these two designs are not always the most efficient, the loss in efficiency is minimal. Therefore Design III or IV is recommended over the other choices for testing lack of fit due to their practical implementation.

As an alternative to Design III, we also tested another design which takes the minimally supported D-optimal design points plus the median of the first k points for P_{k+1}. For two-parameter or four-parameter models, this alternative designs are completely identical to Design III. And for three-parameter models, this alternative design has the design points of 0.2605, 0.7396, 07396, 1 for the logistic model at the D-efficiency of 99.95%, those of 0.2398, 0.7603, 0.7603, 1 for the probit model with 99.93% D-efficiency, and those of 0.1925, 0.7032, 0.7032, 1 for the Gompertz model with D-efficiency at 99.91%. However, these alternative designs all have one design point duplicated, which do not serve the purpose of testing lack of fit.

CHAPTER 4
TESTING LACK OF FIT IN SIGMOID CURVE MODELS: TWO EXTRA POINTS

In this chapter, we extend minimal plus one point designs for testing lack of fit in sigmoid curve models constructed in the previous chapter to include one more point so that the model adequacy could be better tested. For each model, we will develop Design I through IV that are proposed and defined in Section 3.2. From the conclusions in Chapter 3, despite the convenience of implementation and the relative high D-efficiency for the two-parameter models, Design V should be avoided for the three- and four-parameter models; therefore it is excluded from the discussion in this chapter.

As shown in Chapter 3, the two- and three-parameter sigmoid models can be easily derived from the four-parameter models; therefore we will only present the discussion on the four-parameter sigmoid models in this chapter. We will continue to use the criterion of D-optimality, and compare the designs for their D-efficiency.

4.1 Logistic Models

The determinant of the Fisher information matrix, $\det(X'X)$, for the four-parameter logistic (4PL) model is (3.1.3), with $t_i = b(x_i - c)$, $G_i = P_i = \frac{1}{1+e^{-t_i}}$, and $g_i = g(t_i) = \frac{e^{-t_i}}{(1+e^{-t_i})^2}$. Design I, which is a unrestricted D-optimal design, now has $n = k+2$ support points for a k-parameter model; it can be derived by maximizing this determinant. For the other designs (Design II through IV), we plug the values of the first $k+1$ design points obtained in Chapter 3 into $\det(X'X)$, respectively, then obtain the value for the second extra support point by maximizing $\det(X'X)$.

Table 4.1 tabulates the results of this exercise. Note that, despite being unrestricted (Design I) or restricted at 0.5 (Design II), these two D-optimal designs both tend to replicate the design points identified before ($k+1$) for this additional second point. On the contrary, Design III and IV move away for the previously identified $k+1$ design points and find a distinct value for this second extra point, which better serves the purpose of testing lack of fit; and the D-efficiency for these two designs is always above 80%. Note that Design III has the best efficiency as high as 97.67% for a three-parameter logistic model. Therefore Design III and IV give us good coverage of the design space, good efficiency in estimating the model parameters as well as the ability to test for lack of fit.

Table 4.1. Designs and D-Efficiencies for Two-, Three- and Four-Parameter Logistic Models: Minimal Plus Two Extra Points

		Two-parameter Logistic Model ($k=2$)			Three-parameter Logistic Model ($k=3$)			Four-parameter Logistic Model ($k=4$)		
Design	Support Points, n	Design Points	$\det(X'X)$	Efficiency (%)	Design Points	$\det(X'X)$	Efficiency (%)	Design Points	$\det(X'X)$	Efficiency (%)
Minimal D-optimal Design	k	0.2605, 0.7396	0.0060		0.2605, 0.7396, 1	0.0060		0, 0.2605, 0.7396, 1	0.0060	
Designs for lack-of-fit test										
I	$k+2$	0.2605, 0.2605, 0.7396, 0.7396	0.02399	100	0.2605, 0.2605, 0.7396, 0.7396, 1	0.02399	100	0, 0.2052, 0.3454, 0.7498, 1, 1	0.02409	100
II	$k+1$	0.2271, 0.5, 0.7729	0.02006	83.62	0.2243, 0.5, 0.7830, 1	0.02058	85.79	0, 0.2136, 0.5, 0.7864, 1	0.0211	87.59
	+1	0.2271 or 0.7729			0.7830			0.2136 or 0.7864		
III	$k+1$	0.2605, 0.5, 0.7396	0.01988	82.87	0.2605, 0.6667, 0.7396, 1	0.02343	97.67	0, 0.2605, 0.5, 0.7396, 1	0.0207	86.09
	+1	0.2147 or 0.7853			0.2431			0.1846 or 0.8154		
IV	$k+1$	Same as Design III			0.2605, 0.5, 0.7396, 1	0.02039	84.99	Same as Design III		
	+1				0.8108					

4.2 Probit Models

Recall that the determinant of the Fisher information matrix for the four-parameter probit (4PP) model is (3.1.3) with $t_i = b(x_i - c)$, $G_i = P_i = \Phi(t_i)$, $g_i = \phi_i = \phi(t_i)$. Still we use this determinant to find the design points of Design I through IV as defined before.

As indicated in Table 4.2., similar to our findings for the logistic models, Design I and II do not add much new coverage in terms of model assumption testing; they simply put the second extra point on one of the $k+1$ support points for the k-parameter, $(k+1)$-point models that we have already found in Chapter 3. On the other hand, through Design III and IV, we are able to find the second distinct point for testing lack of fit, with only a rather moderate loss in efficiency. Also similar to logistic models, when we have a three-parameter probit model, Design III has much higher efficiency than Design II and IV.

Table 4.2. Designs and D-Efficiencies for Two-, Three- and Four-Parameter Probit Models: Minimal Plus Two Extra Points

Design	Support Points, n	Two-parameter Probit Model ($k=2$) Design Points	$\det(X'X)$	Efficiency (%)	Three-parameter Probit Model ($k=3$) Design Points	$\det(X'X)$	Efficiency (%)	Four-parameter Probit Model ($k=4$) Design Points	$\det(X'X)$	Efficiency (%)
Minimal D-optimal Design	k	0.2398, 0.7603	0.0186		0.2398, 0.7603, 1	0.0186		0, 0.2398, 0.7603, 1	0.0186	
<u>Designs for lack-of-fit test</u>										
I	$k+2$	0.2398, 0.2398 , 0.7603, 0.7603	0.07455	100	0.2398, 0.2398, 0.7603, 0.7603, 1	0.07455	100	0, 0.2398,0.2398, 0.7603, 0.7603, 1	0.07455	100
II	$k+1$	0.2077, 0.5, 0.7923	0.06164	82.68	0.2054, 0.5, 0.8014, 1	0.06305	84.57	0, 0.1959, 0.5, 0.8041, 1	0.06441	86.40
	+1	0.2077 or 0.7923			0.8014			0.1959 or 0.8041		
III	$k+1$	0.2398, 0.5, 0.7603	0.06113	82.00	0.2398, 0.6667, 0.7603, 1	0.07188	96.42	0, 0.2398, 0.5, 0.7603, 1	0.06341	85.06
	+1	0.1971 or 0.8029			0.2207			0.1719 or 0.8281		
IV	$k+1$	Same as Design III			0.2398, 0.5, 0.7603, 1	0.06247	83.80	Same as Design III		
	+1				0.8245					

4.3 Gompertz Models

We now consider the scenario for testing the lack of fit for Gompertz models when we add two more points to the minimally supported designs. For the four-parameter Gompertz (4PG) model, (3.1.3) is the determinant of the information matrix with the specifications of $t_i = b(x_i - c)$, $G_i = \frac{1}{2^{e^{-t_i}}}$, and $g_i = \frac{e^{-t_i} \log 2}{2^{e^{-t_i}}}$. The solutions to the design results presented in Table 4.3. are obtained by maximizing this determinant, $\det(X'X)$.

Consistent with the conclusions we got for the logistic and probit models, the unrestricted and restricted D-optimal designs (Design I and II respectively) now only repeat one of the D-optimal points as the additional support point, which makes it less useful for testing lack of fit. However, Designs III and IV introduce an extra distinct point that better characterize the middle part of the curve. These designs allow good coverage of the design space, provide the opportunity to test the lack of fit, and have relatively high efficiency.

Table 4.3. Designs and D-Efficiencies for Two-, Three- and Four-Parameter Gompertz Models: Minimal Plus Two Extra Points

		Two-parameter Gompertz Model ($k=2$)			Three-parameter Gompertz Model ($k=3$)			Four-parameter Gompertz Model ($k=4$)		
Design	Support Points, n	Design Points	det($X'X$)	Efficiency (%)	Design Points	det($X'X$)	Efficiency (%)	Design Points	det($X'X$)	Efficiency (%)
Minimal D-optimal Design	k	0.1925, 0.7032	0.0147		0.1925, 0.7032, 1	0.0147		0, 0.1925, 0.7032, 1	0.0147	
Designs for lack-of-fit test										
I	$k+2$	0.1925, 0.1925, 0.7032, 0.7032	0.05875	100	0.1925, 0.1925, 0.7032, 0.7032, 1	0.05875	100	0, 0, 0.1921, 0.6862, 0.7191, 1	0.05875	100
II	$k+1$	0.1710, 0.5, 0.7430	0.05160	87.83	0.1699, 0.5, 0.7591, 1	0.05265	89.62	0, 0.1653, 0.5, 0.7629, 1	0.05347	91.01
	+1	0.1710			0.1699			0.1653		
III	$k+1$	0.1925, 0.4478, 0.7032	0.04882	83.10	0.1925, 0.6319, 0.7032, 1	0.05765	98.13	0, 0.1925, 0.4739, 0.7032, 1	0.05159	87.81
	+1	0.1571			0.1808			0.1437		
IV	$k+1$	0.1925, 0.5, 0.7032	0.05133	87.37	0.1925, 0.5, 0.7032, 1	0.05200	88.51	0, 0.1925, 0.5, 0.7032, 1	0.05269	89.69
	+1	0.1628			0.1611			0.1494		

CHAPTER 5
TESTING LACK OF FIT FOR TWO-PARAMETER GOMPERTZ MODEL UNDER VARIOUS ERROR VARIATIONS: THREE-POINT DESIGNS

Among the sigmoid models we studied in this dissertation, the Gompertz model is somewhat differrent from logistic and probit curves, separated by its property of asymmetry. As mentioned earlier, the Gompertz model has found wide application in fitting growth curve, especially in tumor growth; however it is less studied in literature. In order to further understanding of Gompertz models, in this chapter, we will focus on two-parameter Gompertz (2PG) models with various error variance structures, and discuss the designs for testing lack of fit.

Throughout this chapter, we simply assume the parameters of a and d in the four-parameter Gompertz (4PG) model are known, therefore be replaced with $y_{\max}$ and $y_{\min}$ in the discussion. We will continue to apply the parameterization conditions, as described in Section 3.1., to the Gompertz models. Since we will focus on two-parameter models, the Gompertz model we study in this chapter is

$$y_i = f(x_i; b, c) = y_{\min} + \frac{y_{\max} - y_{\min}}{2^{\exp[-b(x_i - c)]}} \qquad (5.1.1)$$

Different from the objective function (2.1.3) we have used in Chapter 3 and 4, the objective function is (2.1.2) reproduced in the following display:

$$S(\boldsymbol{\theta}) = [\mathbf{y} - f(\boldsymbol{\theta})]' \mathbf{V}^{-1} [\mathbf{y} - f(\boldsymbol{\theta})],$$

where $\mathbf{V}$ is a diagonal matrix whose i^{th} diagonal element is $g[f(x_i;\boldsymbol{\theta})]$, and we define $f(\boldsymbol{\theta}) = f(x_i;\boldsymbol{\theta})$ for easy reference. In Sections 3.5 and 4.3, we discussed the designs of Gompertz models under constant variation, assuming that $\mathbf{V} = \sigma^2 \mathbf{I}_n$, where $\mathbf{I}_n$ is the identity matrix of order n. However, alternative error structures may exist. We will first develop minimal D-optimal designs using the model in (5.1.1) under bionomial, exponential and Poisson variations. Then, for testing lack of fit, we will follow the methods of Chapter 3, and discuss different three-point designs under the alternative error structures.

5.1. Bionomial Variation: Minimal D-Optimal Design

Typically with proportion response, the error variances are binomially distributed, that is

$$Var(\varepsilon_i) \propto \eta_i(1-\eta_i),$$

where $\eta_i = y_{\min} + P_i(y_{\max} - y_{\min})$.

As indicated in Section 2.2, D-optimal designs are basically constructed to maximize the determinant of the Fisher information matrix, namely $|X'V^{-1}X|$. In the particular case of the minimal design for two-parameter Gompertz (2PG) model, we can directly work with the matrix of $V^{-\frac{1}{2}}X$, and find the optimal design points that maximize $\left|V^{-\frac{1}{2}}X\right|$.

For the 2PG model given in (5.1.1), the design matrix is

$$X = (y_{\max} - y_{\min})\begin{bmatrix} \frac{1}{b}t_1 g_1 & -bg_1 \\ \frac{1}{b}t_2 g_2 & -bg_2 \end{bmatrix},$$

where $g_i = \frac{e^{-t_i}\log 2}{2^{e^{-t_i}}}$. This design matrix can be easily derived from that of the 4PG model, which is represented in (3.5.2).

For the Gompertz model, we have

$$t_i = G^{-1}(P_i) = -\log(-\log_2(P_i)),$$

and

$$P_i = G(t_i) = \frac{1}{2^{\exp[-b(x_i-c)]}} = \frac{1}{2^{\exp(-t_i)}}.$$

These expressions lead to the design matrix

$$X = (y_{\max} - y_{\min})\begin{bmatrix} P_1\log(P_1)[\log(-\log_2(P_1))] & P_1\log(P_1) \\ P_2\log(P_2)[\log(-\log_2(P_2))] & P_2\log(P_2) \end{bmatrix}. \quad (5.1.2)$$

expressed in terms of P_i.

The matrix of weights is

$$V \propto \begin{bmatrix} \eta_1(1-\eta_1) & 0 \\ 0 & \eta_2(1-\eta_2) \end{bmatrix}.$$

In this case,

$$\left|V^{-\frac{1}{2}}X\right| \propto \frac{(y_{\max}-y_{\min})^2 P_1 P_2 \log(P_1)\log(P_2)\log\left(\frac{\log P_1}{\log P_2}\right)}{\sqrt{[y_{\min}+P_1(y_{\max}-y_{\min})][y_{\min}+P_2(y_{\max}-y_{\min})]}}$$
$$* \frac{1}{\sqrt{[1-y_{\min}-P_1(y_{\max}-y_{\min})][1-y_{\min}-P_2(y_{\max}-y_{\min})]}}.$$

Here P_i that maximizes $\left|V^{-\frac{1}{2}}X\right|$ depends on the values of $y_{\max}$ and $y_{\min}$. Optimal P_i values are provided in Table 5.1 for $y_{\max}$ values ranging from 0.75 to 1 and $y_{\min}$ values ranging from 0 to 0.25. To illustrate, if $y_{\min}=0.10$ and $y_{\max}=0.75$, then the D-optimal design is obtained at the doses corresponding to response proportions of $0.10+0.17(0.75-0.10)=0.211$ and $0.10+0.71(0.75-0.10)=0.562$. For fixed value of $y_{\max}$ ($y_{\min}$), the optimal values of P_1 and P_2 increase monotonically with respect to $y_{\min}$ ($y_{\max}$).

For a special case that $y_{\max}=1$ and $y_{\min}=0$. In this case, then we have

$$Var(\varepsilon_i)=\frac{P_i(1-P_i)}{n_i},$$

where n_i is the number of individuals tested at x_i.

The matrix of weights with $y_{\max}=1$ and $y_{\min}=0$ is given by

$$V = \frac{1}{n}\begin{bmatrix} P_1(1-P_1) & 0 \\ 0 & P_2(1-P_2) \end{bmatrix},$$

where n denotes the number of replications at each design point.

Then we have

$$\left|V^{-\frac{1}{2}}X\right| = \left|V^{-\frac{1}{2}}\right|\left|X\right| = n\sqrt{\frac{P_1P_2}{(1-P_1)(1-P_2)}}\log(P_1)\log(P_2)\log\left(\frac{\log P_1}{\log P_2}\right).$$

This determinat is maximized at $P_1 = 0.0697$ and $P_2 = 0.7692$ with the determinant equals to 0.809403.

Table 5.1. Minimal D-Optimal Designs for Two-Parameter Gompertz Model Under Binomial Variation

y_{max}	y_{min}					
	0	0.05	0.10	0.15	0.20	0.25
0.75	0.051,	0.116,	0.141,	0.158,	0.170,	0.179,
	0.616	0.674	0.692	0.703	0.710	0.715
0.80	0.052,	0.117,	0.143,	0.159,	0.172,	0.181,
	0.633	0.687	0.704	0.715	0.722	0.726
0.85	0.055,	0.118,	0.145,	0.162,	0.175,	0.184,
	0.653	0.704	0.720	0.730	0.736	0.741
0.90	0.058,	0.121,	0.148,	0.166,	0.179,	0.189,
	0.678	0.725	0.741	0.750	0.757	0.761
0.95	0.062,	0.125,	0.153,	0.172,	0.186,	0.198,
	0.713	0.757	0.772	0.781	0.787	0.792
1.00	0.070,	0.135,	0.165,	0.187,	0.203,	0.217,
	0.769	0.813	0.830	0.841	0.849	0.856

55

5.2. Bionomial Variation: Three-Point Designs

Now we will further discuss three-point designs under bionomial variation. We will construct the designs to maximize the determinant of the Fisher information matrix, namely $\left|X'V^{-1}X\right|$; we cannot just consider maximizing $\left|V^{-\frac{1}{2}}X\right|$.

For the 2PG model given in (5.1.1), the three-point design matrix is

$$X = (y_{\max} - y_{\min})\begin{bmatrix} \frac{1}{b}t_1 g_1 & -bg_1 \\ \frac{1}{b}t_2 g_2 & -bg_2 \\ \frac{1}{b}t_3 g_3 & -bg_3 \end{bmatrix}, \qquad (5.2.1)$$

where $g_i = \dfrac{e^{-t_i}\log 2}{2^{e^{-t_i}}}$.

With the definition of t_i and g_i denoted before, we have the design matrix represented in P_i, which is

$$X = (y_{\max} - y_{\min})\begin{bmatrix} P_1\log(P_1)[\log(-\log_2(P_1))] & P_1\log(P_1) \\ P_2\log(P_2)[\log(-\log_2(P_2))] & P_2\log(P_2) \\ P_3\log(P_3)[\log(-\log_2(P_3))] & P_3\log(P_3) \end{bmatrix}. \quad (5.2.2)$$

Let us start with the simpler case with $y_{\max} = 1$ and $y_{\min} = 0$, the matrix of weights is

$$V = \frac{1}{n}\begin{bmatrix} P_1(1-P_1) & 0 & 0 \\ 0 & P_2(1-P_2) & 0 \\ 0 & 0 & P_3(1-P_3) \end{bmatrix}.$$

Therefore

$$X'V^{-1}X \propto \begin{bmatrix} \sum_{i=1}^{3}\frac{P_i}{1-P_i}\left(\log P_i\right)^2[\log(-\log_2 P_i)]^2 & \sum_{i=1}^{3}\frac{P_i}{1-P_i}\left(\log P_i\right)^2[\log(-\log_2 P_i)] \\ \sum_{i=1}^{3}\frac{P_i}{1-P_i}\left(\log P_i\right)^2[\log(-\log_2 P_i)] & \sum_{i=1}^{3}\frac{P_i}{1-P_i}\left(\log P_i\right)^2 \end{bmatrix},$$

This determinat is maximized at $P_1 = 0.0697$, $P_2 = 0.7692$ and $P_3 = 0.0697$ or 0.7692, with the determinant equals to 1.31027. This result is consistant with the results obtained in Section 3.5 for Gompertz models with constant variation: when we add one extra point, the D-optimal design is simply the minimal design with one of the design point replicated. This observation under binomial variation can be mathematically proved as following.

Let us rewrite $X'V^{-1}X$ as

$$X'V^{-1}X \propto \frac{1}{v_{11}}X_1X_1' + \frac{1}{v_{22}}X_2X_2' + \frac{1}{v_{33}}X_3X_3'$$

$$\propto A_{12} + \frac{1}{v_{33}}X_3X_3',$$

where $A_{12} = \frac{1}{v_{11}} X_1 X_1' + \frac{1}{v_{22}} X_2 X_2'$ and v_{ii} is the ith diagonal element of the matrix of weights, $P_i(1-P_i)$.

Then we have the determinant as

$$|X'V^{-1}X| \propto \left| A_{12} + \frac{1}{v_{33}} X_3 X_3' \right| = |A_{12}| \left(1 + \frac{1}{v_{33}} X_3' A_{12}^{-1} X_3 \right).$$

To maximize $|X'V^{-1}X|$ with respect to P_3, we have

$$\frac{\partial |X'V^{-1}X|}{\partial P_3} = |A_{12}| \left\{ -\frac{1-2P_3}{v_{33}^2} X_3{}' A_{12}^{-1} X_3 + \frac{2}{v_{33}} \left\{ \frac{\partial X_3}{\partial P_3} \right\}' A_{12}^{-1} X_3 \right\},$$

where

$$\left\{ \frac{\partial X_3}{\partial P_3} \right\}' = [1 + (1 + \log P_3) \log(-\log_2 P_3), 1 + \log P_3].$$

Therefore,

$$\frac{\partial |X'V^{-1}X|}{\partial P_3} = 0$$

$$\Rightarrow \frac{|A_{12}|}{v_{33}} \left[2 + \frac{2(1-P_3) + \log P_3}{1-P_3} \log(-\log_2 P_3), 2 + \frac{1}{1-P_3} \log P_3 \right] A_{12}^{-1} X_3 = 0.$$

Note

$$A_{12} = \begin{bmatrix} \sum_{i=1}^{2} \frac{1}{v_{ii}} P_i^2 (\log P_i)^2 t_i^2 & -\sum_{i=1}^{2} \frac{1}{v_{ii}} P_i^2 (\log P_i)^2 t_i \\ -\sum_{i=1}^{2} \frac{1}{v_{ii}} P_i^2 (\log P_i)^2 t_i & \sum_{i=1}^{2} \frac{1}{v_{ii}} P_i^2 (\log P_i)^2 \end{bmatrix},$$

where $t_i = -\log(-\log_2 P_i)$, and

$$X_3 = \begin{bmatrix} -P_3 \log P_3 t_3 \\ P_3 \log P_3 \end{bmatrix},$$

Then

$$A_{12}^{-1} X_3 = \frac{P_3 \log P_3}{|A_{12}|} \begin{bmatrix} \frac{1}{v_{11}} P_1^2 (\log P_1)^2 (t_1 - t_3) + \frac{1}{v_{22}} P_2^2 (\log P_2)^2 (t_2 - t_3) \\ \frac{1}{v_{11}} P_1^2 (\log P_1)^2 (t_1 - t_3) t_1 + \frac{1}{v_{22}} P_2^2 (\log P_2)^2 (t_2 - t_3) t_2 \end{bmatrix}.$$

It leads to

$$\frac{\partial |X'V^{-1}X|}{\partial P_3} = 0 \Rightarrow$$

$$\frac{P_3 \log P_3}{v_{33}} \left\{ \frac{1}{v_{11}} (2 + \frac{\log P_3}{1 - P_3}) P_1^2 (\log P_1)^2 (t_1 - t_3)^2 \right.$$

$$+ \frac{1}{v_{22}} (2 + \frac{\log P_3}{1 - P_3}) P_2^2 (\log P_2)^2 (t_2 - t_3)^2$$

$$\left. + \frac{2}{v_{11}} P_1^2 (\log P_1)^2 (t_1 - t_3) + \frac{2}{v_{22}} P_2^2 (\log P_2)^2 (t_2 - t_3) \right\} = 0 . \qquad (5.2.3)$$

When $P_3 = P_1$, (5.2.3) becomes

$$2 + (2 + \frac{\log P_1}{1 - P_1})(t_2 - t_1) = 0 .$$

When $P_3 = P_2$, (5.2.3) becomes

$$-2 + (2 + \frac{\log P_1}{1 - P_1})(t_2 - t_1) = 0 .$$

These are the exact equations that you would obtain if you took the derivative of $|X'V^{-1}X|$ when you have the two-point D-optimal design. So we have proved that, in the three-point D-optimal design, the third point is simply a replicate of one of minimal D-optimal design points.

Next we consider the situation that

$$Var(\varepsilon_i) \propto \eta_i(1 - \eta_i) ,$$

where $\eta_i = y_{min} + P_i(y_{max} - y_{min})$. The design matrix X is the same as in (5.2.2), however we now have the matrix of weights as

$$V \propto \begin{bmatrix} \eta_1(1-\eta_1) & 0 & 0 \\ 0 & \eta_2(1-\eta_2) & 0 \\ 0 & 0 & \eta_3(1-\eta_3) \end{bmatrix} = \begin{bmatrix} \Omega_1 & 0 & 0 \\ 0 & \Omega_2 & 0 \\ 0 & 0 & \Omega_3 \end{bmatrix},$$

where $\Omega_i = [1+P_i(r-1)] - y_{\min}[1+P_i(r-1)]^2$, and $r = \dfrac{y_{\max}}{y_{\min}}$.

The matrix of $X'V^{-1}X$ can be written as

$$X'V^{-1}X \propto \frac{1}{\Omega_1}X_1X_1' + \frac{1}{\Omega_2}X_2X_2' + \frac{1}{\Omega_3}X_3X_3' \propto A_{12} + \frac{1}{\Omega_3}X_3X_3',$$

where $A_{12} = \dfrac{1}{\Omega_1}X_1X_1' + \dfrac{1}{\Omega_2}X_2X_2'$.

Therefore,

$$\left|X'V^{-1}X\right| \propto \left|A_{12} + \frac{1}{\Omega_3}X_3X_3'\right| \propto \left|A_{12}\right|\left\{1 + \frac{1}{\Omega_3}X_3'A_{12}^{-1}X_3\right\}.$$

Taking the derivative of this with respect to P_3, and equating this derivative to zero, we get

$$\frac{\left|A_{12}\right|}{\Omega_3}\left\{-\frac{(r-1)[1-2y_{\min}(1+P_3(r-1))]}{\Omega_3}\begin{bmatrix} t_3P_3\log P_3 \\ P_3\log P_3 \end{bmatrix}' + 2\begin{bmatrix} 1+(1+\log P_3)t_3 \\ 1+\log P_3 \end{bmatrix}'\right\}A_{12}'X_3 = 0$$

which reduces to

$$-\frac{(r-1)[1-2y_{\min}(1+P_3(r-1))]P_3^2(\log P_3)^2P_1^2(\log P_1)^2}{\Omega_1\Omega_3}(t_3-t_1)^2$$

$$-\frac{(r-1)[1-2y_{\min}(1+P_3(r-1))]P_3^2(\log P_3)^2P_2^2(\log P_2)^2}{\Omega_2\Omega_3}(t_3-t_2)^2$$

$$+2[1+(1+\log P_3)t_3]\frac{P_3\log P_3P_1^2(\log P_1)^2}{\Omega_1}(t_3-t_1)$$

$$+2[1+(1+\log P_3)t_3]\frac{P_3\log P_3P_2^2(\log P_2)^2}{\Omega_2}(t_3-t_2)$$

$$-2(1+\log P_3)t_1\frac{P_3\log P_3P_1^2(\log P_1)^2}{\Omega_1}(t_3-t_1)$$

$$-2(1+\log P_3)t_2\frac{P_3\log P_3P_2^2(\log P_2)^2}{\Omega_2}(t_3-t_2)=0\,. \tag{5.2.4}$$

When $P_3 = P_1$, (5.2.4) becomes

$$-\frac{(r-1)[1-2y_{\min}(1+P_1(r-1))]P_1\log P_1}{\Omega_1}(t_1-t_2)+2+2(1+\log P_1)(t_1-t_2)=0.$$

When $P_3 = P_2$, (5.2.4) becomes

$$-\frac{(r-1)[1-2y_{\min}(1+P_1(r-1))]P_2\log P_2}{\Omega_2}(t_2-t_1)+2+2(1+\log P_2)(t_2-t_1)=0.$$

These are the exact expressions that you would get by taking the derivative of the determinant of the two-point design. This implies that the three-point design is the two-point design with one of the design point repeated.

By following the conclusions we obtained in Chapter 3 for minimal plus one point design, we are going to derive the design points under Design II, III and IV. As indicated before, Design I for this case is the 3-point D-optimal design without any restriction, with $y_{\max}$ varying from 0.75 to 1 and $y_{\min}$ varying from 0.05 to 0.25. The designs are basically the results shown in Table 5.1, with one of the design points replicated as the third point.

For Design II, we will restrict ourselves to the class of designs where $P_3 = 0.5$. With this restriction the determinant is proportional to

$$\left(\sum_{i=1}^{2} \frac{P_i^2(\log P_i)^2 t_i^2}{\Omega_i}\right)\left(\sum_{i=1}^{2} \frac{P_i^2(\log P_i)^2}{\Omega_i}\right) - \left(\sum_{i=1}^{2} \frac{P_i^2(\log P_i)^2 t_i^2}{\Omega_i}\right)^2 + \frac{0.120114}{\Omega_3}\left(\sum_{i=1}^{2} \frac{P_i^2(\log P_i)^2 t_i^2}{\Omega_i}\right),$$

of which the maximization leads to the D-optimal design points of P_1 and P_2 under varying minimum and maximum values when $P_3 = 0.5$, shown in Table 5.2. The corresponding D-efficiencies for each of the 25 scenarios are included in the parentheses in the table. It shows that each of the 25 D-efficiencies is above 93%, which means the drop in efficiencies under Design IV is minimal.

For Design III, we will take the designs from Table 5.1 as the first two points for each of the 25 scenarios, then add the mean of these two points as the third point. Table 5.3 displays the design points as well as the D-efficiencies in the parentheses. The D-efficiencies are all around 90%, therefore the efficiency that we lose to estimate the model parameters by going with Design III is not large.

Table 5.2. D-Optimal Design Points of P_1 and P_2 with the Restriction of $P_3 = 0.5$ (Design II) for Two-Parameter, Three-Point Gompertz Model Under Binomial Variation

y_{max}	y_{min}				
	0.05	0.10	0.15	0.20	0.25
0.75	0.106, 0.715 (95.88%)	0.127, 0.733 (95.16%)	0.141, 0.744 (94.80%)	0.151, 0.750 (94.60%)	0.158, 0.755 (94.49%)
0.80	0.106, 0.729 (95.46%)	0.128, 0.745 (94.77%)	0.142, 0.755 (94.44%)	0.152, 0.761 (94.27%)	0.159, 0.765 (94.18%)
0.85	0.106, 0.745 (94.95%)	0.128, 0.760 (94.31%)	0.143, 0.769 (94.02%)	0.153, 0.775 (93.88%)	0.161, 0.778 (93.82%)
0.90	0.107, 0.767 (94.36%)	0.130, 0.780 (93.78%)	0.145, 0.788 (93.54%)	0.156, 0.793 (93.45%)	0.164, 0.796 (93.43%)
0.95	0.109, 0.797 (93.69%)	0.132, 0.809 (93.22%)	0.148, 0.816 (93.07%)	0.160, 0.821 (93.04%)	0.169, 0.824 (93.08%)

Table 5.3. Minimal Plus One Point Restricted at Mean Design (Design III) for Two-Parameter, Three-Point Gompertz Model Under Binomial Variation

y_{max}	y_{min}				
	0.05	0.10	0.15	0.20	0.25
0.75	0.116, 0.395, 0.674 (90.75%)	0.141, 0.417, 0.692 (90.75%)	0.158, 0.430, 0.703 (90.79%)	0.170, 0.440, 0.710 (90.84%)	0.179, 0.447, 0.715 (90.89%)
0.80	0.117, 0.402, 0.687 (90.64%)	0.143, 0.423, 0.704 (90.63%)	0.159, 0.437, 0.715 (90.68%)	0.172, 0.447, 0.722 (90.73%)	0.181, 0.454, 0.726 (90.79%)
0.85	0.118, 0.411, 0.704 (90.52%)	0.145, 0.432, 0.720 (90.50%)	0.162, 0.446, 0.730 (90.55%)	0.175, 0.455, 0.736 (90.60%)	0.184, 0.463, 0.741 (90.66%)
0.90	0.121, 0.423, 0.725 (90.40%)	0.148, 0.444, 0.741 (90.37%)	0.166, 0.458, 0.750 (90.41%)	0.179, 0.468, 0.757 (90.46%)	0.189, 0.475, 0.761 (90.52%)
0.95	0.125, 0.441, 0.757 (90.31%)	0.153, 0.462, 0.772 (90.25%)	0.172, 0.476, 0.781 (90.29%)	0.196, 0.487, 0.787 (90.33%)	0.198, 0.495, 0.792 (90.38%)

For two-parameter, three-point Gompertz Model, Design IV is the minimal D-optimal design (shown in Table 5.1) for P_1 and P_2 plus $P_3 = 0.5$. Table 5.4 gives the corresponding D-efficiencies for the designs under each combination of $y_{\max}$ and $y_{\min}$ relative to the three-point Design IV. Again the drop in efficiency is relatively small. And interestingly, for fixed value of $y_{\max}$ ($y_{\min}$), the drops in efficiency increase monotonically with respect to $y_{\min}$ ($y_{\max}$).

Table 5.4. D-Efficiencies of Minimal Plus One Point Restricted at 0.5 Design (Design IV) for Two-Parameter, Three-Point Gompertz Model Under Binomial Variation

y_{max}			y_{min}		
	0.05	0.10	0.15	0.20	0.25
0.75	94.44%	93.31%	92.67%	92.28%	92.02%
0.80	93.91%	92.83%	92.24%	91.88%	91.66%
0.85	93.27%	92.26%	91.73%	91.42%	91.24%
0.90	92.51%	91.60%	91.15%	90.92%	90.79%
0.95	91.64%	90.88%	90.57%	90.44%	90.41%

5.3. Exponential Variation: Minimal D-Optimal Design

In this section we consider the situation where $Var(\varepsilon_i) \propto \eta_i^2$, and

$\eta_i = y_{min} + P_i(y_{max} - y_{min})$, which can be rewritten as $\eta_i = y_{min}\left[1 + P_i(r-1)\right]$, where

$r = \dfrac{y_{max}}{y_{min}}$.

The design matrix X is the same as in (5.1.2), however the matrix of weights is now

$$V \propto \begin{bmatrix} \eta_1^2 & 0 \\ 0 & \eta_2^2 \end{bmatrix}.$$

The minimal D-optimal designs which maximize $\left|V^{-\frac{1}{2}}X\right|$ depend on *r*, and

$$\left|V^{-\frac{1}{2}}X\right| \propto \frac{(r-1)^2 P_1 P_2 \log(P_1)\log(P_2)\log\left(\dfrac{\log P_1}{\log P_2}\right)}{[1+P_1(r-1)][1+P_2(r-1)]}.$$

Optimal designs for P_i's are provided in Table 5.5 for *r* values ranging from 2 to 10. As suggested by the table, the optimal P_i values are monotonically decreasing with respect to *r*.

Table 5.5. Minimal D-Optimal Designs for Two-Parameter Gompertz Model Under Exponential Variation

r	2	3	4	5	6	7	8	9	10
P_1	0.142	0.117	0.101	0.089	0.081	0.074	0.068	0.064	0.060
P_2	0.622	0.573	0.538	0.512	0.490	0.472	0.457	0.443	0.431

5.4. Exponential Variation: Three-Point Designs

For exponential variation where $Var(\varepsilon_i) \propto \eta_i^2$, the design matrix X is the same as in (5.2.2), and the matrix of weights is

$$V \propto \begin{bmatrix} \eta_1^2 & 0 & 0 \\ 0 & \eta_2^2 & 0 \\ 0 & 0 & \eta_3^2 \end{bmatrix},$$

where for simplicity we use η_i for $1+P_i(r-1)$, and $r = \dfrac{y_{\max}}{y_{\min}}$. Then we have

$$X'V^{-1}X = \sum_{i=1}^{3} \frac{(r-1)^2}{[1+P_i(r-1)]^2} X_i X_i{}' = A_{12} + \frac{(r-1)^2}{[1+P_3(r-1)]^2} X_3 X_3{}',$$

where $A_{12} = \sum_{i=1}^{2} \frac{(r-1)^2}{[1+P_i(r-1)]^2} X_i X_i{}'$. Therefore,

$$\left|X'V^{-1}X\right| = \left|A_{12}\right| \left\{ 1 + \frac{(r-1)^2}{[1+P_3(r-1)]^2} X_3{}' A_{12}^{-1} X_3 \right\}. \qquad (5.4.1)$$

Now we will follow what we defined in Chapter 3 and find the three-point designs to test lack of fit.

As before, we will first show that the D-optimal three-point design with no restriction (Design I) is the minimal D-optimal design with one of the design points replicated. First maximizing (5.4.1) with respect to P_3 leads to

$$\frac{\partial\left|X'V^{-1}X\right|}{\partial P_3} = 0 \Rightarrow \frac{2(r-1)^2}{[1+P_3(r-1)]^2}\left|A_{12}\right| \bullet$$

$$\left\{ \frac{-(r-1)}{1+P_3(r-1)} \begin{bmatrix} P_3 \log P_3 t_3 \\ P_3 \log P_3 \end{bmatrix}' + \begin{bmatrix} 1+(1+\log P_3)t_3 \\ 1+\log P_3 \end{bmatrix}' \right\} A_{12}^{-1} X_3 = 0$$

$$\Rightarrow -\frac{r-1}{v_{11}[1+P_3(r-1)]} P_3 \log P_3 P_1^2 (\log P_1)^2 (t_3 - t_1)^2$$

$$-\frac{r-1}{v_{22}[1+P_3(r-1)]} P_3 \log P_3 P_2^2 (\log P_2)^2 (t_3 - t_2)^2$$

$$+\frac{P_1^2 (\log P_1)^2}{v_{11}}(t_3 - t_1) + \frac{P_2^2 (\log P_2)^2}{v_{22}}(t_3 - t_2)$$

$$+\frac{1+\log P_3}{v_{11}} P_1^2 (\log P_1)^2 (t_3 - t_1)^2 + \frac{1+\log P_3}{v_{22}} P_2^2 (\log P_2)^2 (t_3 - t_2)^2 = 0. \qquad (5.4.2)$$

When $P_3 = P_1$, (5.4.2) becomes

$$(1+\log P_1)(t_1 - t_2) + 1 - \frac{(r-1)P_1 \log P_1}{1 + P_1(r-1)}(t_1 - t_2) = 0\,.$$

When $P_3 = P_2$, (5.4.2) becomes

$$(1+\log P_2)(t_1 - t_2) - 1 - \frac{(r-1)P_2 \log P_2}{1 + P_2(r-1)}(t_1 - t_2) = 0\,.$$

These are the exact equations that you would obtain if we took the derivative of $|X'V^{-1}X|$ when we have the two-point design. This implies that the D-optimal three-point design points are the same points that are used in the two-point D-optimal designs. Therefore, Design I for testing lack of fit is simply the two-point D-optimal design with one of the design points replicated.

For Design II, with the restriction of $P_3 = 0.5$, we have

$$|X'V^{-1}X| = \frac{(r-1)^4 P_1^2 P_2^2 (\log P_1)^2 (\log P_2)^2}{[1+P_1(r-1)]^2[1+P_2(r-1)]^2}(t_1 - t_2)^2$$

$$+ \frac{0.120114}{[1+0.5(r-1)]^2}\left\{\frac{P_1^2(\log P_1)^2(\log P_2)^2 t_1^2}{[1+P_1(r-1)]^2} + \frac{P_2^2(\log P_2)^2 t_2^2}{[1+P_2(r-1)]^2}\right\}.$$

Maximizing this gives us the results in Table 5.6, which show that the drop in efficiency is negligible.

Table 5.6. D-Optimal Design Points of P_1 and P_2 with the Restriction of $P_3 = 0.5$ (Design II) for Two-Parameter Gompertz Model Under Exponential Variation

r	2	3	4	5	6	7	8	9	10
P_1	0.134	0.113	0.09592	0.089	0.081	0.075	0.069	0.065	0.061
P_2	0.661	0.602	0.09584	0.517	0.486	0.460	0.438	0.419	0.403
D-Efficiency	97.62%	99.06%	98.93%	99.97%	99.98%	99.84%	99.62%	99.35%	99.05%

Table 5.7 lists the design points for Design III and the corresponding D-efficiencies. And Table 5.8 has the D-efficiencies of using Design IV. For both designs, the ratio *r* does affect the value of the design points. Adding one point at either the mean or simply 0.5 only loses very little efficiency, but gains the ability to test for lack of fit.

Table 5.7. Minimal Plus One Point Restricted at Mean Design (Design III) for Two-Parameter Gompertz Model Under Exponential Variation

r	2	3	4	5	6	7	8	9	10
P_1	0.142	0.117	0.101	0.089	0.081	0.074	0.068	0.064	0.060
P_2	0.382	0.345	0.320	0.300	0.285	0.273	0.263	0.254	0.246
P_3	0.622	0.573	0.538	0.512	0.490	0.472	0.457	0.443	0.431
D-Efficiency	91.69%	92.03%	92.31%	92.53%	92.72%	92.89%	93.04%	93.19%	93.28%

Table 5.8. D-Efficiencies of Minimal Plus One Point Restricted at 0.5 Design (Design IV) for Two-Parameter, Three Point Gompertz Model Under Exponential Variation

r	2	3	4	5	6	7	8	9	10
D-Efficiency	96.42%	98.63%	99.61%	99.96%	99.97%	99.79%	99.53%	99.21%	98.81%

5.5. Poisson Variation: Minimal D-Optimal Design

Next we discuss the situation where $Var(\varepsilon_i) \propto \eta_i$. Similar to the situation under exponential variation, we can denote $\eta_i = y_{\min}\left[1 + P_i(r-1)\right]$, where $r = \dfrac{y_{\max}}{y_{\min}}$.

The design matrix remains the same as (5.1.2), however the matrix of weights is now

$$V \propto \begin{bmatrix} \eta_1 & 0 \\ 0 & \eta_2 \end{bmatrix}.$$

Then we have

$$\left| V^{-\frac{1}{2}} X \right| \propto \frac{(r-1)^2 P_1 P_2 \log(P_1)\log(P_2)\log\left(\dfrac{\log P_1}{\log P_2}\right)}{\sqrt{[1+P_1(r-1)][1+P_2(r-1)]}}.$$

Table 5.9 lists the minimal D-optimal designs, derived from maximizing $\left| V^{-\frac{1}{2}} X \right|$, for various values of *r* from 2 to 10. Similar to the results under exponential variation, the optimal P_i values are monotonically decreasing with respect to *r*.

Table 5.9. Minimal D-Optimal Designs for Two-Parameter Gompertz Model Under Poisson Variation

r	2	3	4	5	6	7	8	9	10
P_1	0.165	0.149	0.138	0.130	0.123	0.118	0.114	0.110	0.106
P_2	0.665	0.644	0.629	0.619	0.611	0.604	0.598	0.593	0.589

5.6. Poisson Variation: Three-Point Designs

Under Poisson variation where $Var(\varepsilon_i) \propto \eta_i$, the matrix of weights is

$$V \propto \begin{bmatrix} \eta_1 & 0 & 0 \\ 0 & \eta_2 & 0 \\ 0 & 0 & \eta_3 \end{bmatrix},$$

where the design matrix is the same as (5.2.2).

Under a similar proof as demonstrated in Section 5.4 under exponential variation, we can show that the three-point D-optimal design (Design I) is the minimal D-optimal design with one of the design pints replicated.

Design II is the D-optimal design restricted at $P_3 = 0.5$. Table 5.10 lists the optimal points for P_1 and P_2, as well as the D-efficiency for using these designs in stead of those of Design I, for various values of r. It shows that, as r, the ratio of $y_{\max}$ to $y_{\min}$, goes up, the values of P_1 and P_2 as well as the drop in efficiency go down.

Table 5.10. D-Optimal Design Points of P_1 and P_2 with the Restriction of $P_3 = 0.5$ (Design II) for Two-Parameter, Three Point Gompertz Model Under Poisson Variation

r	2	3	4	5	6	7	8	9	10
P_1	0.151	0.138	0.129	0.122	0.117	0.112	0.108	0.105	0.102
P_2	0.707	0.685	0.669	0.657	0.647	0.639	0.632	0.626	0.620
D-Efficiency	96.15%	96.87%	97.36%	97.72%	97.99%	98.21%	98.39%	98.53%	98.65%

The results for Design III are displayed in Table 5.11. For *r* ranging from 2 to 10, the designs all have the D-efficiencies around 91%. Again the drop in efficiency by utilizing Design III is very small.

Table 5.11. Minimal Plus One Point Restricted at Mean Design (Design III) for Two-Parameter, Three Point Gompertz Model Under Poisson Variation

R	2	3	4	5	6	7	8	9	10
P_1	0.165	0.149	0.138	0.130	0.123	0.118	0.114	0.110	0.106
P_2	0.415	0.397	0.384	0.374	0.367	0.361	0.356	0.351	0.348
P_3	0.665	0.644	0.629	0.619	0.611	0.604	0.598	0.593	0.589
D-Efficiency	91.40%	91.47%	91.52%	91.55%	91.58%	91.60%	91.62%	91.64%	91.66%

Also we apply Design IV to the model under Poisson variation, and summarize the D-efficiencies for various values of *r* in Table 5.12. Similarly as Design II, as *r* increases, the values of P_1 and P_2 as well as the drop in efficiency decrease monotonically.

Table 5.12. D-Efficiencies of Minimal Plus One Point Restricted at 0.5 Design (Design IV) for Two-Parameter, Three Point Gompertz Model Under Poisson Variation

r	2	3	4	5	6	7	8	9	10
D-Efficiency	94.17%	95.33%	96.12%	96.68%	97.11%	97.44%	97.71%	97.93%	98.11%

CHAPTER 6
SUMMARY AND FUTURE WORK

In this dissertation we focused on obtaining nonlinear optimal or highly efficient designs that can also be used to test lack of fit of the model assumptions. We first proposed the use of Design III or IV for a k-parameter sigmoid curve, which consist of the minimally supported D-optimal points plus an additional point computed as a arithmetic mean of the first k points or simply 0.5, respectively. Then we extended these designs to include one more point by allocating the D-optimal point as this additional point when the first $k+1$ support points are restricted at either Design III or IV. We found that, Design III or IV as well as their extended versions not only provides a meaningful test for lack of fit, but also are robust with high efficiency for parameter estimation.

Among the class of sigmoid models we studied, the Gompertz model is somewhat different due to its asymmetry, therefore is not widely discussed in the literature. In this dissertation, we focused on two-parameter Gompertz (2PG) models with various error variance structures. We first developed the minimal D-optimal designs under binomial, exponential and Poisson variations, then further discussed Design I through IV under these variance structures for testing lack of fit. Similar to the result under constant variance, the conclusion follows the argument that Design III or IV gives reasonably good coverage of the design space, is convenient to implement, and has the ability to test for lack of fit.

In future research, to further evaluate the efficiency of Design III and IV, one may study the minimal plus one point designs for the two-parameter models in greater depth. Suppose one has $2(k+1)+1$ subjects to be allocated to $k+1$ design points (for the two-parameter models, $k=2$; therefore, in this case, 7 subjects are to be allocated to 3 design points), 3 sceniarios of experiementation is possible:

1) Allocate 1 subject to each of the two minimally supported D-optimal points, then allocate the rest of 5 subjects to the extra design point (i.e., the mean of the first two points for Design III or 0.5 for Design IV);

2) Allocate 1 subject to the extra point, 1 subject to one of the two minimally supported D-optimal points, then replicate the other minimally supported D-optimal points 5 times.

3) Allocate 1 subject to the extra point, and replicate 3 times on each of the minimally supported D-optimal points.

One can compare the efficiency under these three sceniarios for each of the two-parameter sigmoid models studied in this dissertaion. One can also compare the power of lack-of-fit test of these designs.

In Chapter 5 and the Appendix, we proved that, for two-parameter models, the three-point D-optimal design is a replication of the minimal D-optimal design for their specific model that they are using. However, there does not seem to be a general proof of this. One needs to extend the proof to general sigmoid models with more than two parameters.

BIBLIOGRAPHY

Antonello, J. M. and Raghavarao, D. (2000). Optimal designs for the individual and joint exposure general logistic regression models. *J. Biopharm. Statist.*, *10,* 351–367.

Atkinson, A. C. (1972). Planning experiments to detect inadequate regression models. *Biometrika*, *59*, 275-293.

Atkinson, A. C. (1982). Developments in the design of experiments. *Internat. Stat. Rev., 50,* 161-177.

Atkinson, A. C. and Cox, D. R. (1974). Planning experiments for discriminating between models. *J. Roy. Statist. Soc. B36,* 321–348.

Atkinson, A. C. and Donev, A. V. (1992). *Optimum Experimental Design.* Oxford: Clarendon Press.

Bates, D.M. and Watts, D.G. (1980). Relative curvature measures of nonlinearity. *J. Roy. Statist. Soc. B42*, 1–25.

Bates, D. M. and Watts, D. G. (1988). *Nonlinear Regression Analysis and its Applications.* New York: Wiley.

Becka, M., Bolt, H.M., Urfer, W. (1993). Statistical evaluation of toxicokinetic data. *Environmetrics*, *4,* 311-322.

Becka, M., Urfer, W. (1996). Statistical aspects of inhalation toxicokinetics. *Environ. Ecol. Statist. 3,* 51-64.

Biedermann, S., Dette, H., Zhu, W. (2006). Optimal designs for dose-response models with restricted design spaces. *J. Amer. Statist. Assoc., 101,* 747–759.

Box, G.E.P. and Lucas, H.L. (1959). Design of experiments in nonlinear situations. *Biometrika*, *46,* 77-90.

Box, M.J. (1971). Bias in nonlinear estimation. *J. Roy. Statist. Soc. B33,* 171–201.

Charnes, A., Frome, E.L., Yu, P.L. (1976). The equivalence of generalized least squares and maximum likelihood estimates in the exponential family. *J. Amer. Statist. Assoc.,71,* 169-171.

Chernoff, H. (1953). Locally optimal designs for estimating parameters. *Ann. Math. Statist., 24,* 586–602.

Chervoneva, I., Li, Y., Iglewicz, B., Waldman, S., Hyslop, T. (2007). Relative quantification based on logistic models for individual polymeraze chain reactions. *Statist. Med.*, *26,* 5596-5611.

Coleman, M. and Marks, H. (1998). Topics in dose-response modeling. *J. Food Protection*, *61(11),* 1550-1559.

Davidian M. and Giltinan D.M. (1995). *Nonlinear Models for Repeated Measurement Data*. London: Chapman & Hall.

Dette, H., Melas, V.B., Pepelyshev, A. (2004). Optimal designs for a class of nonlinear regression models. *Ann. Statist., 32,* 2142-2167.

Dette, H., Melas, V., Wong, W. K. (2005). Optimal design for goodness-of-fit of the Michaelis-Menton enzyme kinetic function. *J. Amer. Statist. Assoc., 100,* 1370–1381.

Dette, H. and Pepelyshev, A. (2008). Efficient experimental designs for sigmoidal growth models. *J. Statist. Plan. Inf*., *138,* 2-17.

Fedorov, V.V. (1972). *Theory of Optimal Experiments.* New York: Academic Press.

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. *J. Roy. Statist. Soc., B46,* 149-192.

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 contingencies. *Phil. Trans. Roy. Soc. Lond.*, *115,* 513-585.

Green, P.J. (1984). Iteratively reweighted least squares for maximum likelihood estimation, and some robust resistant alternatives (with discussion). *J. Roy. Statist. Soc., B42,* 1–25.

Han, C. and Chaloner, K. (2003). D- and c-optimal designs for exponential regression models used in viral dynamics and other applications. *J. Statist. Plan. Inf., 115,* 585-601.

Hedayat, A.S., Yan, B., Pezzuto, J.M. (1997). Modeling and identifying optimum designs for fitting dose-response curves based on raw optical density data. *J. Amer. Statist. Assoc*., *92,* 1132-1140.

Hedayat, A. S., Zhong, J., Nie, L. (2004). Optimal and efficient designs for 2-parameter nonlinear models. *J. Statist. Plan. Inf., 124,* 205–217.

Jorgensen, B. (1983). Maximum likelihood estimation and large sample inference for generalized linear and nonlinear regression models. *Biometrika*, *70,* 19-28.

Kalish, L.A. and Rosenberger, J.L. (1978). Optimal designs for the estimation of the logistic function. Tech. Rep. 33, Pennsylvania State University.

Khuri, A. I., Mukherjee, B., Sinha, B. K., Ghosh M. (2006). Design issues for generalized linear models: A review. *Statist. Science*, *21,* 376–399.

Kidwell, J. F., Howard, A., Laird, A. (1969). The inheritance of growth and form in the mouse. II. The Grompertz growth equation. *Growth, 33,* 339-352.

Kiefer, J. and Wolfowitz, J. (1960). The equivalence of two extremum problems. *Canadian J. Math.*, *12,* 363-366.

Krug, H. and Liebig, H.P. (1988). Static regression models for planning greenhouse production. *Acta Horticulturae*, *230,* 427-433.

Laird, A.K. (1964). Dynamics of tumor growth. *Br. J. of Cancer, 18,* 490-502.

Landaw, E.W. and DiStefano 3rd, J.J. (1984). Multiexponential, multicompartmental, and noncompartmental modeling. II. Data analysis and statistical considerations. *Amer. J. Physiol.*, *246,* 665-677.

Li, G. and Majumdar, D. (2008). D-optimal designs for logistic models with three and four Parameters. *J. Statist. Plan. Inf.*, *138,* 1950–1959.

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

Liebig, H.P. (1988). Temperature integration by kohlrabi growth. *Acta Horticulturae*, *230,* 427-433.

Lupinacci, P. J. (2001). D-optimal designs for a class of nonlinear models. Unpublished Ph.D. Thesis, Temple University, Philadelphia, PA.

Lupinacci, P. J. and Raghavarao, D. (2000). Designs for testing lack of fit for a nonlinear dose-response curve model. *J. Biopharm. Statist.*, *10,* 45–53.

Lupinacci, P.J. and Raghavarao, D. (2003). D-optimal designs for the Michaelis-Menten model under various error variance structures. Technical report, Villanova University, Dept. of Mathematical Sciences.

Meibohm, B. and Derendorf, H. (1997). Basic concepts of pharmacokinetic / pharmacodynamic (PK/PD) modeling. *Int. J. Clin. Pharmacol. Therapeut., 35,* 401-413.

Mood, A.M. (1946). On Hotelling's weighing problem. *Ann. Math. Statist., 17,* 432–446.

Myers, R.H. (1986). *Classical and Modern Regression with Applications*. Boston: PWS Publishers.

Nelson, D.L. and Cox, M.M. (2005). *Lehninger Principles of Biochemistry*. New York: Freeman.

Pukelsheim, F. (1993). Optimal Design of Experiments. New York: Wiley.

Pukelsheim, F. and Rosenberger, J.L. (1993). Experimental designs for model discrimination. *J. Amer. Statist. Assoc., 88,* 642–649.

Ratkowsky, D.A. (1983). *Nonlinear Regression Modelling: A Unified Practical Approach.* New York: Marcel Dekker.

Rygaard, K., Spang-Thomsen, M. (1997). Quantitation and Gompertzian analysis of tumor growth. *Breast Cancer Research and Treatment*, *46,* 303-312.

St. John, R.C. and Draper, N.R. (1975). D-optimality for regression designs: A review. *Technometrics*, *17,* 15-23.

Seber, G.A.F. and Wild, C.J. (1989). *Nonlinear Regression*. New York: Wiley.

Shah, K.R. and Sinha, B.K. (1989). *Theory of Optimal Designs* (Lecture Notes in Statistics). New York: Springer-Verlag.

Silvey, S.D. (1980). *Optimal Design.* London: Chapman and Hall.

Steinberg, D.M. and Hunter, W.G. (1984). Experiment design: Review and comment. *Technometrics, 26,* 71-97.

Stigler, S.M. (1971). Optimal experimental design for polynomial regression. *J. Amer. Statist. Assoc., 66,* 311-318.

Studden, W.J. (1982). Some robust type D-optimal designs in polynomial regression. *J. Amer. Statist. Assoc., 77,* 916-921.

Wald, A. (1943). On the efficient design of statistical investigations. *Ann. Math. Statist., 14,* 134–140.

Wang, Y., Myers, R. H., Smith, E. P., Ye, K. (2006). D-optimal designs for Poisson regression models. *J. Statist. Plan. Inf., 136,* 2831–2845.

Whittle, P. (1973). Some general points in the theory of optimal experimental design. *J. Roy. Statist. Soc., B35*, 123-130.

Winsor, C. (1932). The Gompertz curve as a growth curve. *Proc. Natl. Acad. Sci., 18,* 1-8.

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

APPENDIX

Lupinacci (2001) proved that, for two-parameter logistic model (2PL) under constant variation, the three-point D-optimal design is the minimal D-optimal design with one of the design points replicated. In this section, we will prove that the results are the same for two-parameter probit model (2PP) and two-parameter Gompertz model (2PG) under constant variation, which is the results we derived in Chapter 3 for Design I.

Assume that $\mathbf{V} = \sigma^2 \mathbf{I}_3$, where $\mathbf{I}_3$ is the 3×3 identity matrix. Therefore,

$$\begin{aligned} X'V^{-1}X &\propto \frac{1}{v_{11}} X_1 X_1' + \frac{1}{v_{22}} X_2 X_2' + \frac{1}{v_{33}} X_3 X_3' \\ &\propto A_{12} + X_3 X_3', \end{aligned}$$

where $A_{12} = X_1 X_1' + X_2 X_2'$. Then

$$\left| X'V^{-1}X \right| \propto \left| A_{12} + X_3 X_3' \right| = \left| A_{12} \right| \left(1 + X_3' A_{12}^{-1} X_3 \right).$$

To maximize $\left| X'V^{-1}X \right|$ with respect to P_3, we have

$$\frac{\partial \left| X'V^{-1}X \right|}{\partial P_3} = \left| A_{12} \right| \left\{ 2 \left\{ \frac{\partial X_3}{\partial P_3} \right\}' A_{12}^{-1} X_3 \right\} = 0 . \qquad \text{(A.1.1)}$$

A.1 Two-Parameter Probit Model Under Constant Variation: Three-Point D-Optimal Design (Design I)

For the 2PP model,

$$y_i = f(x_i; b, c) = y_{\min} + (y_{\max} - y_{\min})\Phi[b(x_i - c)].$$

The design matrix is the same as (5.2.1), except that now $g_i = \phi_i = \phi(t_i)$, and $t_i = b(x_i - c) = \Phi^{-1}(P_i)$. Here,

$$\left\{\frac{\partial X_3}{\partial P_3}\right\}' = [1 - t_3^2\phi_3, t_3].$$

and

$$A_{12}^{-1}X_3 = \frac{\phi_3}{|A_{12}|}\begin{bmatrix} \phi_1^2(t_3 - t_1) + \phi_2^2(t_3 - t_2) \\ \phi_1^2(t_1 - t_3)t_1 + \phi_2^2(t_2 - t_3)t_2 \end{bmatrix}.$$

Therefore

$$\frac{\partial |X'V^{-1}X|}{\partial P_3} = 0 \Rightarrow$$

$$\phi_1^2(1 - \phi_3 t_3^2 + t_1 t_3)(t_3 - t_1)$$

$$+ \phi_2^2(1 - \phi_3 t_3^2 + t_2 t_3)(t_3 - t_2) = 0. \qquad \text{(A.1.2)}$$

When $P_3 = P_1$, (A.1.2) becomes

$$(1-\phi_1 t_1^2 + t_1 t_2)(t_1 - t_2) = 0\,.$$

When $P_3 = P_2$, (A.1.2) becomes

$$(1-\phi_2 t_2^2 + t_1 t_2)(t_2 - t_1) = 0\,.$$

These equations are the exact equations that must be solved in the derivation of the two-point D-optimal design. This implies that, the three-point D-optimal design (Design I) for two-parameter Probit model with constant variation is the minimal D-optimal design plus one of the design points replicated.

A.2 Two-Parameter Gompertz Model Under Constant Variation: Three-Point D-Optimal Design (Design I)

As demonstrated in Chapter 3, under constant variation, Design I for two-parameter Gompertz model (2PG) is simply the minimal D-optimal design with one of the design point replicated. Here in this section we are going to prove it mathematically.

The design matrix is the same in (5.2.2). Now since

$$\left\{\frac{\partial X_3}{\partial P_3}\right\}' = \left[1+(1+\log P_3)\log(-\log_2 P_3), 1+\log P_3\right],$$

and

$$A_{12}^{-1}X_3 = \frac{P_3 \log P_3}{|A_{12}|}\begin{bmatrix} \frac{1}{v_{11}}P_1^2(\log P_1)^2(t_3 - t_1) + \frac{1}{v_{22}}P_2^2(\log P_2)^2(t_3 - t_2) \\ \frac{1}{v_{11}}P_1^2(\log P_1)^2(t_1 - t_3)t_1 + \frac{1}{v_{22}}P_2^2(\log P_2)^2(t_2 - t_3)t_2 \end{bmatrix},$$

(A.1.1) is equivalent to

$$\begin{aligned} &\frac{\partial |X'V^{-1}X|}{\partial P_3} = 0 \Rightarrow \\ &P_1^2(\log P_1)^2(1+\log P_3)(t_3 - t_1)^2 \\ &+ P_2^2(\log P_2)^2(1+\log P_3)(t_3 - t_2)^2 \\ &+ P_1^2(\log P_1)^2(t_3 - t_1) + P_2^2(\log P_2)^2(t_3 - t_2) = 0 . \end{aligned} \qquad \text{(A.2.1)}$$

When $P_3 = P_1$, (A.2.1) becomes

$$1 + (1 + \log P_1)(t_1 - t_2) = 0 .$$

When $P_3 = P_2$, (A.2.1) becomes

$$1 - (1 + \log P_1)(t_1 - t_2) = 0 .$$

These are the exact equations that you would obtain if you took the derivative of $|X'V^{-1}X|$ when you have the two-point D-optimal design. So we have proved that, in the three-point D-optimal design for two-parameter Gompertz model under constant variation, the third point is simply a replicate of one of minimal D-optimal design points.