DT-OPTIMAL DESIGNS FOR PROBIT MODELS IN CLINICAL TRIALS

A Thesis
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Andrew Michael Lexvold

In Partial Fulfillment of the Requirements
for the Degree of
MASTER OF SCIENCE

Major Department:
Statistics

May 2015

Fargo, North Dakota

North Dakota State University
Graduate School

Title

DT-Optimal Designs for Probit Models in Clinical Trials

By

Andrew Michael Lexvold

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

MASTER OF SCIENCE

SUPERVISORY COMMITTEE:

Dr. Seung Won Hyun
Chair

Dr. Wonwoo Byun

Dr. Gang Shen

Approved:

08/06/2015
Date

Dr. Rhonda Magel
Department Chair

ABSTRACT

The Optimal designs used in a clinical trial depends on the goals of the study. Common goals are estimating model parameters and choosing between models. D-optimal designs are used when the goal is to estimate the model parameters. This is achieved by maximizing the determinant of the information matrix. When the goal is model discrimination, T-optimal designs are used. The design is optimal when the minimum difference between the models is maximized. Generally, D-optimal designs are not efficient when the goal is model discrimination and T-optimal designs perform poorly when the goal is parameter estimation. However, because D-optimal and T-optimal designs have a common criterion structure, they can be combined into a new design called a DT-optimal design. DT-optimal designs provide a balance between parameter estimation and model discrimination. The efficiency of DT-optimal designs relative to D and T-optimal designs shows that they work for parameter estimation and model discrimination.

Keywords: Experimental design; Dose-response; Dual objective optimal design, efficiency.

ACKNOWLEDGEMENTS

I would like to thank my committee members Dr. Wonwoo Byun, Dr. Gang Shen, and especially my advisor, Dr. Seung Won Hyun, for all of his assistance and encouragement while working on my thesis. Without his continued support I might have never finished. I also want to thank my parents and undergraduate professors. They encouraged me to pursue a graduate degree and I wouldn't be in this position without them.

TABLE OF CONTENTS

ABSTRACT ……...iii

ACKNOWLEDGEMENTS ...……iv

LIST OF TABLES ……vi

LIST OF FIGURES ……...vii

CHAPTER 1. INTRODUCTION ……1

CHAPTER 2. BACKGROUND ……..4

2.1. Dose-response Studies ……...……...4

2.2. T-optimal Designs ……...……6

2.3. D-optimal Designs ……...……7

CHAPTER 3. DT-OPTIMAL DESIGNS ……8

3.1. DT-optimality ……...……8

3.2. Algorithm to Find Optimal Design ……...……9

CHAPTER 4. DESIGN RESULTS ……...11

4.1. Competing Models ……...……..11

4.2. Nominal Parameter Values and Dose Range ……..……11

4.3. T-optimal Designs ……...……...14

4.4. D-optimal Designs ……...……...14

4.5. DT-optimal Designs ……...……....15

CHAPTER 5. EFFICIENCY ……18

CHAPTER 6. CONCLUSIONS ……20

BIBLIOGRAPHY ……21

APPENDIX. R CODE ……...23

LIST OF TABLES

Table Page

4.1. Parameter sets for η_1 and estimates for η_2……..…….....….……12

5.1. Efficiency for parameter set 1…….…….....…….18

5.2. Efficiency for parameter set 2……..……...……..19

5.3. Efficiency for parameter set 3…….…….……...……..19

LIST OF FIGURES

Figure Page

2.1. Strong-downturn dose response……..…….…….…….4

2.2. Slight-downturn dose response…….....…….....…..5

2.3. No-downturn dose response……..……...……..5

4.1. The two competing models for the strong-downturn dose response……...…….12

4.2. The two competing models for the slight-downturn dose response……...….13

4.3. The two competing models for the no-downturn dose response…….…….13

4.4. Optimality check for parameter set 1……...…....16

4.5. Optimality check for parameter set 2……..……..……...…16

4.6. Optimality check for parameter set 3……..……..….....17

CHAPTER 1. INTRODUCTION

Dose response studies are a common type of clinical trial used to determine the appropriate dose of a drug to give to patients. Patients are given a range of dose levels that help explain possible reactions to the drug in question. Nonmonotone dose-response functions are common and three types will be considered: strong-downturn, slight-downturn, and no-downturn. The functions will be discussed in more detail in Chapter 2. For drugs that exhibit a downturn in the dose-response, a common question among researches is whether this downturn is significant or not. To assess this significance two competing models can be estimated. One model describing the full dose-response function with a downturn, and one that only describes the increasing component of the function.

In order to describe these functions, parameters must be estimated. This is done using an experimental design. Designs consist of design points (dose levels) and weights (subject allocations). In order for the experiment to be efficient, the optimal design must be used. Because there are many different objectives possible when conducting clinical trials, there are also many types of optimal designs. Here we consider T-optimal designs, D-optimal designs, and their combination, DT-optimal designs.

Welshons et al. (2003) gives motivation for studying dose-response functions with a downturn and contains useful information on obtaining optimal designs, such as nominal parameter values and the dose range. A Probit model will be used to describe the dose-response functions because it can easily describe the downturn by adding a quadratic term and it provides a good fit to Welshons' data, see Hyun (2013). We use T-optimal designs for discrimination between the rival two and three parameter models. In other words, we can use T-optimal designs to choose the most appropriate model. Atkinson and Fedorov (1975a,b) show that T-optimality

criterion maximizes the minimum difference between the two competing models, which in turn gives more power to an F-test of lack of fit, assuming the first model is true. More information about and uses of T-optimal designs can be found in Ucinski and Bogacka (2005), Wiens (2009), Tommasi and L´opez-Fidalgo (2010), Atkinson, Bogacka and Bogachi (1998), and Foo and Duffull (2011). D-optimal designs are used when the goal of the study is accurate parameter estimation. Box and Lucas (1959) show that D-optimal designs maximize the determinant of the Fisher Information Matrix. In some sense, this can be thought of as minimizing the variance of the parameter estimates.

There have been many attempts at seeking a balance between model discrimination and parameter estimation. Hill et al. (1968) gives motivation for why researches might be interested in a single design for both goals. Instead of having a design for one stage of the experiment where the goal is to find a good model, and a completely separate design for the next stage where model parameters are to be estimated, they suggest that a design to do both at the same time would allow researchers to combine the two stages. This would save both time and precious resources. Most importantly the overall number of subjects for the study would be decreased. Waterhouse et al. (2005) has a practical situation where researchers might actually need to discriminate between models and estimate parameters. They are researching optimal sampling times for subjects with cystic fibrosis. In their study, they need not only to discriminate between two linear and two nonlinear models, but also estimate the model parameters in order to determine the optimal sampling times.

DT-optimal designs, introduced by Atkinson (2008) as a way to balance model discrimination and parameter estimation, are found by simply maximizing a weighted product of T-efficiency and D-efficiency. This represents relative importance of the two objectives. More

details will be given in Chapter 3. A design is considered to be an optimal design when it satisfies a certain optimality criterion, and it is verified by the General Equivalence Theorem in Kiefer (1974). The paper also has background on equivalency theorems in general if more information is desired.

In Chapter 2, more information about dose-response functions is given, as well as background on D-optimal and T-optimal designs. DT-optimal designs and the DT-optimal equivalence theorem are discussed in detail in Chapter 3. Efficiency calculations are shown in Chapter 4 for D-optimal, T-optimal, and DT-optimal designs. Chapter 5 contains concluding remarks as well as potential further research.

CHAPTER 2. BACKGROUND

2.1. Dose-response Studies

Plotting the logarithm of dose on the X-axis and the drug response on the Y-axis can result in a variety of shapes. Three of those shapes are under consideration here. The first is called a "strong-downturn" function. This function resembles a concave parabola. The second is called a "slight-downturn" function. Also resembling a concave parabola, this function is "flatter" than a strong-downturn dose-response function. The third is called a "no-downturn" function and resembles a Sigmoid "S" curve.

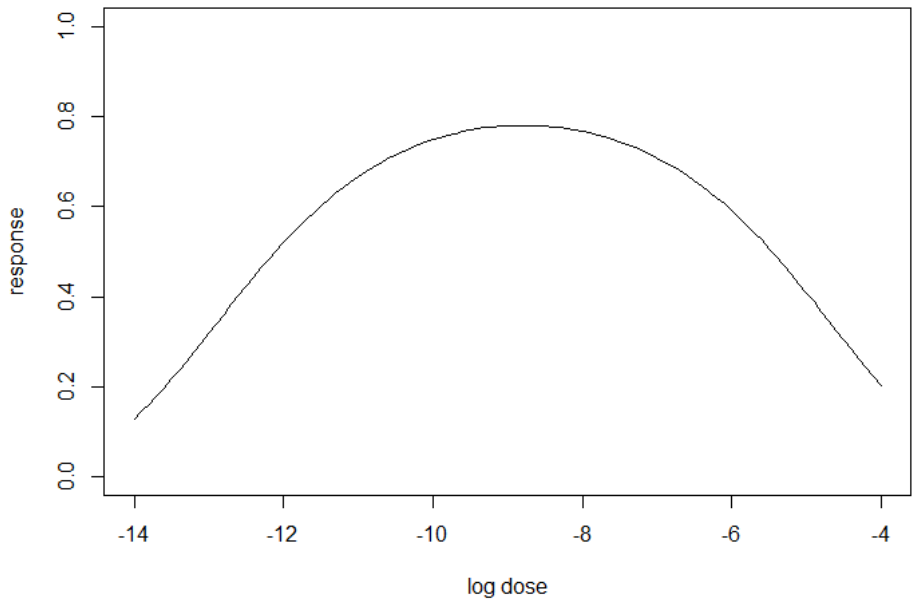

Figure 2.1. Strong-downturn dose response

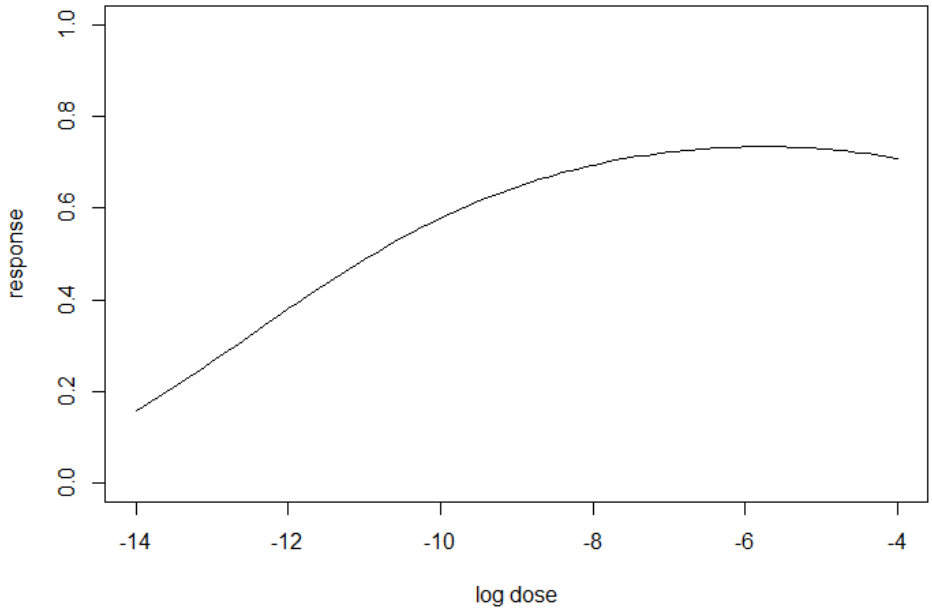

Figure 2.2. Slight-downturn dose response

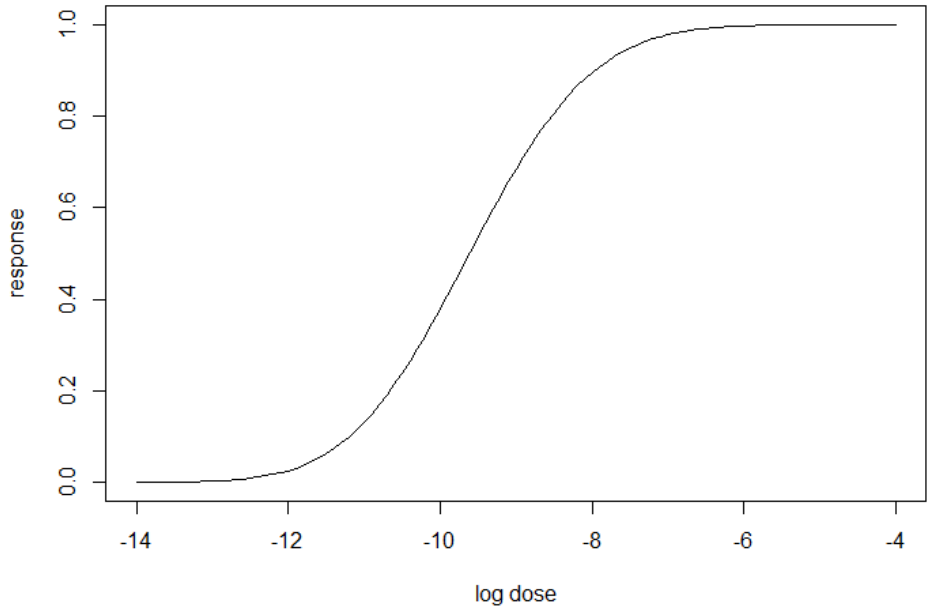

Figure 2.3. No-downturn dose response

2.2. T-optimal Designs

In some studies, researches may have a hard time deciding between competing plausible models. In these cases, designs that are efficient at model discrimination are useful because they can help select the most appropriate model using the least effort and resources. Atkinson and Fedorov (1975) consider designs for discrimination between two rival regression models $\eta_1(x, \Theta_1)$ and $\eta_2(x, \Theta_2)$, whose forms are given in Section 4.1. The model parameters are estimated by least squares. If it is assumed that the first model is true, then

$$y_{ij} = \eta_1(x_i, \Theta_1) + \varepsilon_{ij}, \qquad \varepsilon_{ij} \sim N(0, \sigma^2), \tag{2.1}$$

where y_{ij} is continuous and represents a dose response, $j = 1, 2, 3, \ldots, n_i$, $i = 1, 2, 3, \ldots, k$, n_i is the number of subjects allocated to x_i, and $\Theta_1 = \{\theta_{11}, \theta_{12}, \ldots, \theta_{1p}\}$is the vector of model parameters. The values of x_i are selected for a dose range X. Finally, the sample size, $N = \sum_{i=1}^{k} n_i$.

Let $\xi = \{x_i, w_i\}^k$ denote k design points (dose levels), where x_i is the i^{th} log dose and w_i is the corresponding weight at x_i. Here $x_i \epsilon X$ and the nearest positive inter of $N * w_i$ is the number of subjects assigned to x_i. The lack of fit sum of squares for model $\eta_2(x, \Theta_2)$ is made as large as possible by maximizing

$$\Delta_1(\xi) = \sum_{i=1}^{k} w_i \{\eta_1(x_i, \Theta_1) - \eta_2(x_i, \hat{\Theta}_2)\}^2, \tag{2.2}$$

where $\widehat{\Theta_2}$ are the calculated parameters of η_2 that minimize its distance from η_1 and $\Delta_1(\xi)$ is called the objective function of a design ξ. Note that the values of x_i and w_i are found using the algorithm in Section 3.2. The values of Θ_1 come from previous research and can be seen below in Table 4.1. The design maximizing (2.2) is called T-optimal design, it is denoted as ξ_T^*.

Atkinson and Fedorov (1975) give the Equivalence Theorem for T-optimal designs as the following,

$$\Psi_1(x, \xi_T^*) \leq \Delta_1(\xi_T^*), \tag{2.3}$$

where

$$\Psi_1(x, \xi_T^*) = \{\eta_1(x_i, \Theta_1) - \eta_2(x_i, \hat{\theta}_2)\}^2.$$

Equality in (2.3) holds at the design points of the T-optimal design.

2.3. D-optimal Designs

When the goal of the study is parameter estimation, D-optimal designs are most commonly used. As mention earlier, D-optimal designs maximize the determinant of the Fisher Information Matrix. Under (2.1), the Fisher Information matrix can be written as

$$M_1(\xi) = F_1^T W F_1, \tag{2.4}$$

where $W = diag\{w_i\}$ and

$$F_1(x, \Theta_1) = \left\{\frac{\delta\eta_1(x_1,\Theta_1)}{\delta\Theta_1}, \frac{\delta\eta_1(x_2,\Theta_1)}{\delta\Theta_1}, \dots, \frac{\delta\eta_1(x_k,\Theta_1)}{\delta\Theta_1}\right\}_{p\times k}, \tag{2.5}$$

$$\frac{\delta\eta_1(x_i,\Theta_1)}{\delta\Theta_1} = \left\{\frac{\delta\eta_1(x_i,\Theta_1)}{\delta\theta_{11}}, \frac{\delta\eta_1(x_i,\Theta_1)}{\delta\theta_{12}}, \dots, \frac{\delta\eta_1(x_i,\Theta_1)}{\delta\theta_{1p}}\right\}^T.$$

The D-optimal design ξ_D^* maximizes $|M_1(\xi)|$. Kiefer and Wolfowitz (1960) give the equivalence theorem for D-optimal designs as

$$\Psi_2(x, \xi_D^*) \leq p, \tag{2.7}$$

where p is the number of parameters in the first model and

$$\Psi_2(x, \xi_D^*) = f_1^T(x) M_1^{-1}(\xi_D^*) f_1(x),$$

$$f_1(x, \Theta_1) = \frac{\delta\eta_1(x_1,\Theta_1)}{\delta\Theta_1}.$$

Similarly to T-optimality, equality in (2.7) holds at the design points of the D-optimal design.

CHAPTER 3. DT-OPTIMAL DESIGNS

3.1. DT-optimality

The DT-optimal optimality criterion is a weighted product of the efficiencies of T-optimal and D-optimal designs. The DT-optimal design maximizes the efficiencies for model discrimination (T-optimal design) and parameter estimation (D-optimal design). According to Atkinson (2008), the efficiency of any design ξ relative to the T-optimal design ξ_T^* is called T-efficiency:

$$E^T = \Delta_1(\xi)/\Delta_1(\xi_T^*). \tag{3.1}$$

In other words, it is the ratio of the objective function for the design in question to the optimal design. D-efficiency is

$$E^D = \{|M_1(\xi)|/|M_1(\xi_D^*)|\}^{(1/p)}. \tag{3.2}$$

If we denote the efficiency of a design ξ as $E(\xi)$, we can calculate how many more subjects a researcher would need to have the same accuracy as the optimal design using the formula

$$\pi = 100 \times \left(\frac{1}{E(\xi)} - 1\right)\%. \tag{3.3}$$

For example, $E(\xi) = 0.5$ implies a study would 100% or twice as many subjects to be just as accurate as the optimal design.

To find a DT-optimal design we maximize the weighted product of (3.1) and (3.2).

$$\{E_f^T\}^{(1-\lambda)}\{E_f^D\}^{\lambda} = \{\Delta_1(\xi)/\Delta_1(\xi_T^*)\}^{(1-\lambda)}\{|M_1(\xi)|/|M_1(\xi_D^*)|\}^{\lambda/p}, \quad (0 \le \lambda \le 1). \tag{3.4}$$

If $\lambda = 0$ (3.4) becomes T-efficiency and if $\lambda = 1$ (3.4) becomes D-efficiency. We can simplify (3.4) by first taking the log,

$$(1-\lambda)log\Delta_1(\xi) + (\lambda/p)log|M_1(\xi)| - (1-\lambda)log\Delta_1(\xi_T^*) - (\lambda/p)log|M_1(\xi_D^*)| \tag{3.5}$$

then removing the terms involving the optimal designs ξ_D^* and ξ_T^* because they are constant when we maximize (3.4) over ξ, to find

$$\Phi_1^{DT}(\xi) = (1-\lambda)log\Delta_1(\xi) + (\lambda/p)log|M_1(\xi)|. \quad (3.6)$$

The design maximizing (3.5) is called DT-optimal and is denoted ξ_{DT}^*. Here λ is the relative importance of the D-optimality criterion compared to the T-optimality criterion. When $\lambda = 1$ the design maximizing (3.5) is the D-optimal design, and when $\lambda = 0$ the design maximizing (3.5) is the T-optimal design.

The equivalence theorem for DT-optimality design (Atkinson 2008) states that

$$\Psi^{DT}(x, \xi_{DT}^*) \leq 1, \quad (3.7)$$

where

$$\Psi^{DT}(x, \xi_{DT}^*) = (1-\lambda)\Psi_1(x, \xi_{DT}^*)/\Delta_1(\xi_{DT}^*) + (\lambda/p)\Psi_2(x, \xi_{DT}^*)$$

$$= (1-\lambda)\{\eta_1(x_i, \Theta_1) - \eta_2(x_i, \hat{\Theta}_2)\}^2/\Delta_1(\xi_{DT}^*) + (\lambda/p)f_1^T(x)M_1^{-1}(\xi_{DT}^*)f_1(x).$$

Similarly to both the T-optimal and D-optimal equivalence theorems equality in (3.6) holds at the design points of the DT-optimal design.

3.2. Algorithm to Find Optimal Design

We can find the DT-optimal design numerically using the well-known V-algorithm (Fedorov 1972).

Step 0: Set an initial design ξ_0 with design points $x_{1,0}, x_{2,0}, \dots, x_{k,0}$ and uniform weights $w_{1,0}, w_{2,0}, \dots, w_{k,0} = 1/k$

Step 1: Obtain $\hat{\Theta}_{2,s}$:

$$\hat{\Theta}_{2,s} = \arg\min_{\Theta_{2,s} \in \Theta_S} \sum_{i=1}^k w_{i,s} \left(\eta_1(x_{i,s}, \Theta_{1,s}) - \eta_2(x_{i,s}, \hat{\Theta}_{2,s})\right)^2.$$

Step 2: Find the design point

$$x_s^* = \arg\max_{x_s^* \in X}(1-\lambda)log\sum_{i=1}^{k} w_{i,s}\left(\eta_1\left(x_{i,s},\Theta_{1,s}\right)-\eta_2\left(x_{i,s},\hat{\Theta}_{2,s}\right)\right)^2$$

$$+(\lambda/p)log\left|f_1^T\left(x_{i,s}\right)M_1^{-1}(\xi_s)f_1\left(x_{i,s}\right)\right|.$$

Step 3: Stop the algorithm if $|\Psi^{DT}(x_s^*,\xi_{DT}^*)-1| \le \varepsilon$, $10^{-8} \le \varepsilon \le 10^{-3}$. When the algorithm stops ξ_s is the DT-optimal design.

Step 4: Otherwise, update the design

$$\xi_{s+1} = \left(1-\frac{1}{s+1}\right) * \xi_s + \xi_{new},$$

$$\xi_{new} = \left\{\begin{matrix} x_s^* \\ \frac{1}{s+1} \end{matrix}\right\}.$$

Step 5: Set $s = s+1$ and repeat Steps 1 – 5.

CHAPTER 4. DESIGN RESULTS

4.1. Competing Models

Two probit models are considered. The first has three parameters and includes a quadratic term. The second model only has two parameters and does not include a quadratic term. The two models are

$$\eta_1(x, \Theta_1) = \Phi\left(-(\theta_{11} + \theta_{12}x_i + \theta_{13}{x_i}^2)\right), \tag{4.1}$$

$$\eta_2(x, \Theta_2) = \Phi\left(-(\theta_{21} + \theta_{22}x_i)\right). \tag{4.2}$$

Φ is the cumulative standard normal distribution. The Fisher Information matrix for $\eta_1(x, \Theta_1)$ is

$$\begin{bmatrix} \sum_{i=1}^{k} w_i\, e^{\left(-(\theta_{11}+x_i\theta_{12}+{x_i}^2\theta_{12})^2\right)} & \sum_{i=1}^{k} x_i w_i\, e^{\left(-(\theta_{11}+x_i\theta_{12}+{x_i}^2\theta_{12})^2\right)} & \sum_{i=1}^{k} {x_i}^2 w_i\, e^{\left(-(\theta_{11}+x_i\theta_{12}+{x_i}^2\theta_{12})^2\right)} \\ \sum_{i=1}^{k} x_i w_i\, e^{\left(-(\theta_{11}+x_i\theta_{12}+{x_i}^2\theta_{12})^2\right)} & \sum_{i=1}^{k} {x_i}^2 w_i\, e^{\left(-(\theta_{11}+x_i\theta_{12}+{x_i}^2\theta_{12})^2\right)} & \sum_{i=1}^{k} {x_i}^3 w_i\, e^{\left(-(\theta_{11}+x_i\theta_{12}+{x_i}^2\theta_{12})^2\right)} \\ \sum_{i=1}^{k} {x_i}^2 w_i\, e^{\left(-(\theta_{11}+x_i\theta_{12}+{x_i}^2\theta_{12})^2\right)} & \sum_{i=1}^{k} {x_i}^3 w_i\, e^{\left(-(\theta_{11}+x_i\theta_{12}+{x_i}^2\theta_{12})^2\right)} & \sum_{i=1}^{k} {x_i}^4 w_i\, e^{\left(-(\theta_{11}+x_i\theta_{12}+{x_i}^2\theta_{12})^2\right)} \end{bmatrix} \tag{4.3}$$

4.2. Nominal Parameter Values and Dose Range

In order to illustrate the 3 types of response functions, three sets of nominal parameter values and the log dose range were adopted from Ming (2014). Estimates for the two parameter model provide a minimum distance from $\eta_1(x, \Theta_1)$. They are the values last used to find the minimum distance when the algorithm detailed above stops. Welshons et al. (2003) gives motivation for using the design space $[-14, -4]$ and the initial parameter values for the probit model for the strong-downturn dose response function.

Table 4.1. Parameter sets for η_1 and estimates for η_2

	θ_{11}	θ_{12}	θ_{13}	$\hat{\theta}_{21}$	$\hat{\theta}_{22}$
Strong-downturn	4.630	1.230	0.070	0.03125	-0.0125
Slight-downturn	0.175	0.277	0.024	-1.453125	-0.1515625
No-downturn	-6.690	-0.600	0.010	-5.734375	-0.5984375

The following plots show the two competing models for each dose-response function.

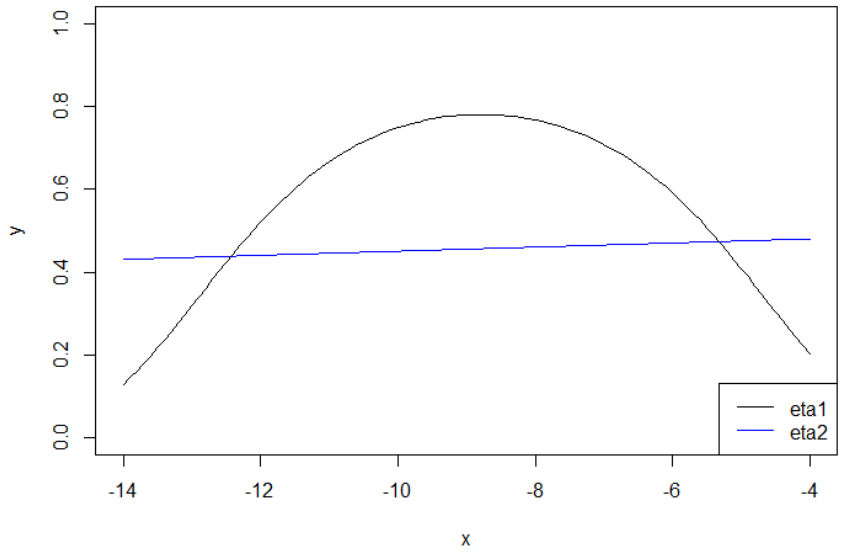

Figure 4.1. The two competing models for the strong-downturn dose response

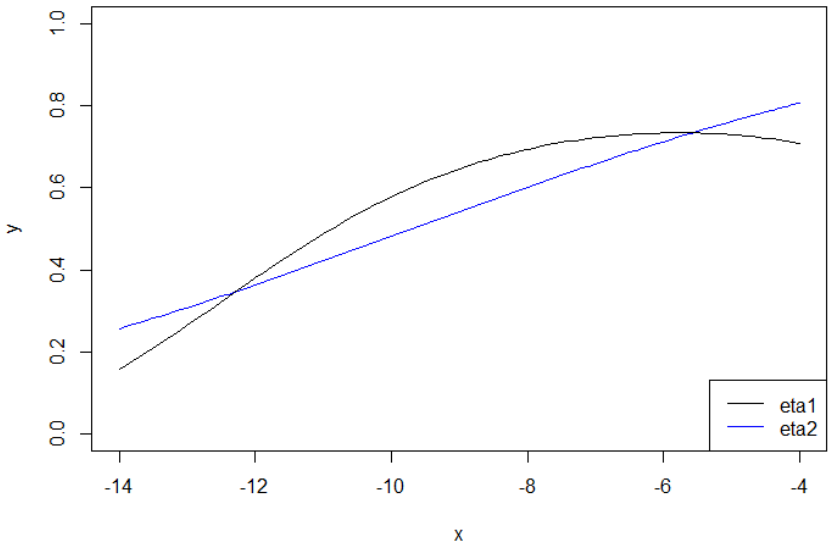

Figure 4.2. The two competing models for the slight-downturn dose response

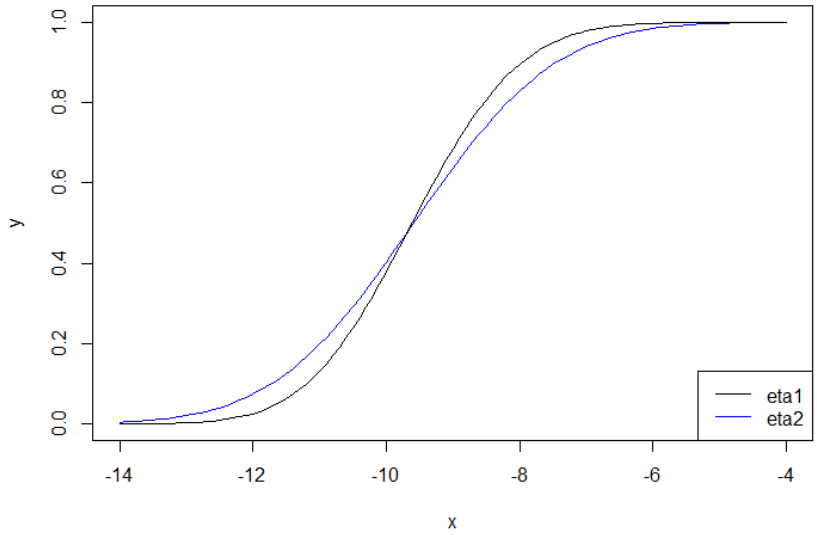

Figure 4.3. The two competing models for the no-downturn dose response

4.3. T-optimal designs

Recall T-optimal designs maximizing the minimum difference between the competing models. The T-optimal designs are:

For the Strong-downturn dose response function

$$\xi_T^* = \begin{Bmatrix} -14 & -9 & -4 \\ 0.25 & 0.50 & 0.25 \end{Bmatrix}, \tag{4.4}$$

For the Slight-downturn dose response function

$$\xi_T^* = \begin{Bmatrix} -14 & -9.2 & -4 \\ 0.27 & 0.43 & 0.30 \end{Bmatrix}, \tag{4.5}$$

For the No-downturn dose response function

$$\xi_T^* = \begin{Bmatrix} -11.1 & -8.2 \\ 0.43 & 0.57 \end{Bmatrix}, \tag{4.6}$$

Consider a drug with a strong-downturn dose response. In this scenario we would assign 25% of the subjects to each log dose boundary and 50% of the subjects to the log dose level -9. This would allow us to most accurately maximize the distance between the two models. Also notice the similarity between (4.4) and (4.5). The design points are essentially identical, and the subjects for the slight-downturn dose response are just a little more evenly distributed with different weights. The T-optimal design for the no-downturn dose response is substantially different. There are only two design points. This will be a problem when trying to estimate model parameters because we need at least as many design points as there are model parameters in order to find estimates. The magnitude of the problem will be quantified when we calculate this design's D-efficiency.

4.4. D-optimal designs

Recall D-optimal designs maximizing the determinant of the Fisher Information matrix. The D-optimal designs are:

For the Strong-downturn dose response function

$$\xi_D^* = \begin{Bmatrix} -13.2 & -10.3 & -7.2 & -4.3 \\ 0.32 & 0.17 & 0.18 & 0.33 \end{Bmatrix}, \tag{4.7}$$

For the Slight-downturn dose response function

$$\xi_D^* = \begin{Bmatrix} -13 & -9.5 & -4 \\ 0.33 & 0.34 & 0.33 \end{Bmatrix}, \tag{4.8}$$

For the No-downturn dose response function

$$\xi_D^* = \begin{Bmatrix} -11.1 & -9.6 & -8 \\ 0.34 & 0.33 & 0.33 \end{Bmatrix}, \tag{4.9}$$

All three dose response functions have unique D-optimal designs. There are still similarities however. In (4.7) there are four design point while in (4.8) there are only three. However notice that the boundaries are similar and that (4.7)'s middle two design points are basically a split in both log dose and weight of the middle design point of (4.8). It is also worth noticing that both (4.8) and (4.9) have essentially uniform weights. This is a typical trait of D-optimal designs when the number of model parameters is equal to the number of design points. To understand how to use a D-optimal design consider (4.9). If our drug exhibits a no-downturn dose response we would allocate and equal amount of subjects to the log dose levels -11.1, -9.6, and -8. This would allow researchers to most accurately estimate the models parameters.

4.5. DT-optimal designs

Recall DT-optimal designs maximizing the weighted product of D-efficiency and T-efficiency. The DT-optimal designs are:

For the Strong-downturn dose response function

$$\xi_{DT}^* = \begin{Bmatrix} -13.9 & -9 & -4 \\ 0.28 & 0.45 & 0.27 \end{Bmatrix}, \tag{4.10}$$

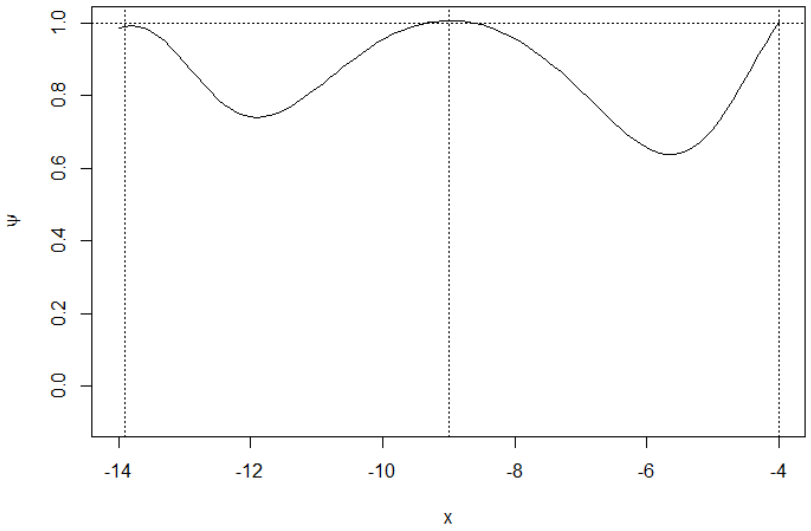

Figure 4.4. Optimality check for parameter set 1

For the Slight-downturn dose response function

$$\xi_{DT}^{*} = \begin{Bmatrix} -14 & -9.3 & -4 \\ 0.30 & 0.40 & 0.30 \end{Bmatrix}, \tag{4.11}$$

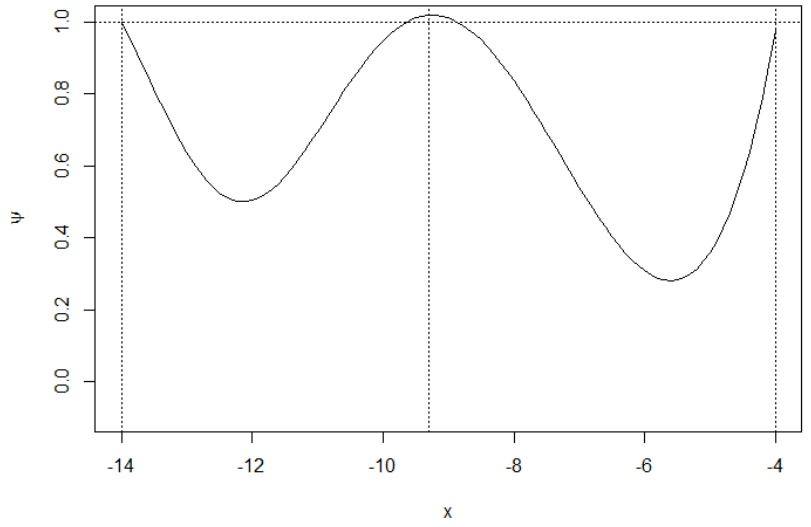

Figure 4.5. Optimality check for parameter set 2

For the No-downturn dose response function

$$\xi_{DT}^{*} = \begin{Bmatrix} -11.1 & -9.5 & -8.1 \\ 0.41 & 0.18 & 0.41 \end{Bmatrix}, \tag{4.12}$$

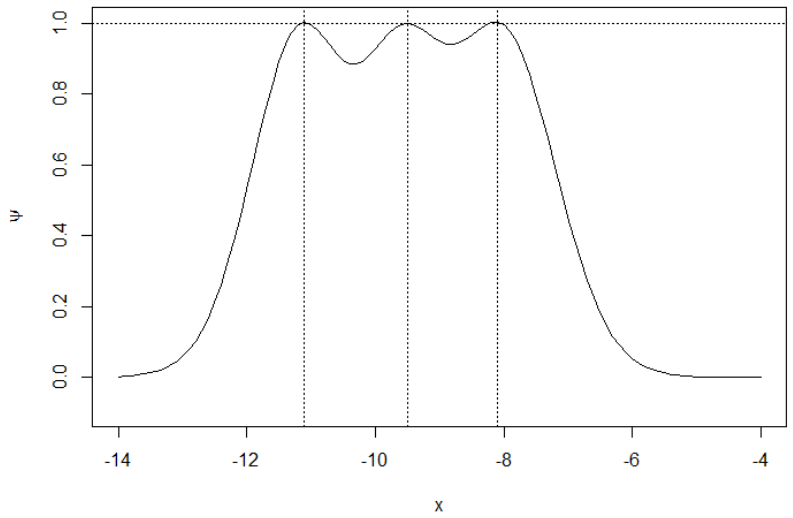

Figure 4.6. Optimality check for parameter set 3

This time consider a drug with a slight-downturn dose response. (4.11) tells us that approximately 30% of the subjects should go to the boundary log dose values, and that 40% of the subjects should be given a log dose value of -9.3. With this we could provide a good balance between model discrimination and parameter estimation. The figure following each designs is a visual way to check if the design is really optimal. According to (3.7), at the design points of the DT-optimal design, $\Psi^{DT}(x, \xi_{DT}^{*}) = 1$. Sticking with the slight downturn dose response example, we can see that this is the case. $\Psi^{DT}(x, \xi_{DT}^{*}) = 1$ at $x = -14, -9.3, and -4$. It is also worth noting that the DT-optimal designs for strong-downturn and slight-downturn dose response functions are very similar to the T-optimal designs. However for a no-downturn dose response function the DT-optimal design is more similar to the D-optimal design.

CHAPTER 5. EFFICIENCY

Efficiency is used to compare a given design to the optimal design. Recall (3.1) and (3.2) for calculating T-efficiency and D-efficiency. Using these we can calculate the efficiency of the DT-optimal design relative to the T-optimal and D-optimal designs for each parameter set.

For the strong-downturn dose-response function:

Table 5.1. Efficiency for parameter set 1

	$E^T(\xi)$	$E^D(\xi)$
ξ_T^*	1.00	0.79
ξ_D^*	0.50	1.00
ξ_{DT}^*	0.95	0.85

Here we can see that the T-optimal design is only about 80% efficient at parameter estimation. This means the about 25% more subjects would be required to reach the same accuracy as the D-optimal designs, according to (3.3). However the D-optimal design preforms even worse at model discrimination. As stated before, an efficiency of 50% means that twice as many subjects would be needed to be as accurate as the T-optimal design. The DT-optimal design preforms much better in both cases. With a T-efficiency of 95% only about 6% more subjects would be needed to reach the same accuracy as the T-optimal design. Also a D-efficiency of 85% means about 18% more subjects would be required to be just as accurate as the D-optimal design at parameter estimation.

Table 5.2. Efficiency for parameter set 2

	$E^T(\xi)$	$E^D(\xi)$
ξ_T^*	1.00	0.97
ξ_D^*	0.85	1.00
ξ_{DT}^*	0.98	0.98

For the slight-downturn dose response we can see that all three optimal designs are fairly efficient. Looking back at (4.5), (4.8), and (4.11), we can see that all three designs are pretty similar, so the high efficiencies aren't surprising. However DT-optimal designs are still slightly better if both parameter estimation and model discrimination are important.

Table 5.3. Efficiency for parameter set 3

	$E^T(\xi)$	$E^D(\xi)$
ξ_T^*	1.00	0.14
ξ_D^*	0.65	1.00
ξ_{DT}^*	0.83	0.92

When we consider the no-downturn dose response function T-optimal designs are very inefficient at parameter estimation. At 14% D-efficiency, about 615% more subjects would be needed to accurately estimate model parameters. The reason this design is so bad at parameter estimation is because there are only two design points, while the model has three parameters. The T-efficiency of the D-optimal design is 65%. About 54% more subjects would be needed to be as accurate at model discrimination as the T-optimal design. The DT-optimal design has a T-efficiency of 83% and a D-efficiency of 92%. 21% and 9% more subjects would be needed to reach the same accuracy as the T-optimal and D-optimal designs, respectively.

CHAPTER 6. CONCLUSIONS

The DT-optimal design for the strong-downturn dose response function was very similar to the T-optimal design. For the slight-downturn dose response function all three designs were similar, only the weights for the D-optimal design were different. The DT-optimal and D-optimal designs were nearly the same for the no-downturn dose response. Overall DT-optimal designs perform well for both model discrimination and parameter estimation for all three dose response functions. For the strong-downturn and no-downturn dose-response functions D-optimality performs poorly when the objective is model discrimination and T-optimality performs poorly for parameter estimation, as expected.

Despite some similarities in the designs, it is best to use the DT-optimal design in all cases if the study has goals to choose a model and estimate parameters. DT-optimal designs will always be at least slightly more efficient at model discrimination than D-optimal designs, and more efficient at parameter estimation than T-optimal designs.

Future research in this area could include studying DT-optimal designs for different values of λ to see the relationship between λ and the efficiencies. Also, because the DT-optimal design relies on the nominal parameter values, studying how DT-optimal designs change when the parameter values are specified incorrectly would be an interesting research question. Finally another extension could be simulating actual data to use with these designs. This would allow a better picture of how well DT-optimal designs preform at both tasks.

BIBLIOGRAPHY

[1] Atkinson, A. C. and Fedorov, V. V. (1975a). *The design of experiments for discriminating between two rival models*. Biometrika 62, 57C70.

[2] Atkinson, A. C. and Fedorov, V. V. (1975b). *Optimal design: Experiments for discriminating between several models*. Biometrika 62, 289C303.

[3] Atkinson, A. C., Bogacka, B. and Bogachi, M. B. (1998). *D- and T-optimum designs for the kinetics of a reversible chemical reaction*. Chemometrics and Intelligent Laboratory Systems 43, 185-198.

[4] Atkinson, A.C., (2008). *DT-optimum designs for model discrimination and parameter estimation*. Journal of Statistical Planning and Inference 138, 56-64.

[5] Box, G.E.P., Lucas, H.L., (1959). *Design of experiments in nonlinear situations*. Biometrika 46, 77–90.

[6] Biswas, A., Chaudhuri, P., 2002. *An efficient design for model discrimination and parameter estimation*. Biometrika 89, 709–718.

[7] Foo, L. K. and Duffull, S. (2011). *Optimal design of pharmacokinetic-pharmacodynamic studies.* Pharmacokinetics in Drug Development: Advances and Applications 3, 175-194.

[8] Fedorov, V. V. (1971). *Asymptotically optimal designs of experiments for discriminating two rival regression models*. Theory of Probability and Its Applications 16, 561-2.

[9] Fedorov V.V. (1972). *Theory of optimal experiments*. Academic Press. New York.

[10] Hill, W.J., Hunter, W.G., Wichern, W.D., (1968). *A joint design criterion for the dual problem of model discrimination and parameter estimation*. Technometrics 10, 145–160.

[11] Hyun, S.W (2013): *Optimal Designs for a Probit Model With a Quadratic Term*, Statistics in Biopharmaceutical Research 5:1, 18-26.

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

[13] Kiefer, J. (1974). *General equivalence theory for optimum designs*. Annals of Statistics 2, 849C879.

[14] Stigler, S. (1971). *Optimal experimental design for polynomial regression*. Journal of the American Statistical Association 66, 311C318.

[15] Ming, Y. (2014). T-*Optimal Designs for Model Discrimination in Probit Models*. North Dakota State University. Master's Thesis.

[16] Studden,W. J. (1982). *Some robust-type D-optimal designs in polynomial regression*. Journal of the American Statistical Association 77, 916C921. MR0686418.

[17] Tommasi, C. and L´opez-Fidalgo, J. (2010). *T-optimum designs for discrimination between two multiresponse dynamic models*. Journal of the Royal Statistical Society: Series B 67, 3-18.

[18] Ucinski, D. and Bogacka, B. (2005). *Bayesian optimum designs for discriminating between models with any distribution*. Computational Statistics and Data Analysis 54, 143C150.

[19] Waterhouse, T.H., Eccleston, J.A., Duffull, S.B., (2004). *On optimal design for discrimination and estimation*. In: Antoch, J. (Ed.), COMPSTAT 2004: Proceedings in Computational Statistics. Physica-Verlag, Heidelberg, pp. 1963–1970.

[20] Waterhouse, T.H., Eccleston, J.A., Duffull, S.B., (2005). *Optimal design for model discrimination and parameter estimation for itraconazole population pharmocokinetics in cystic fibrosis patients*. Journal of Pharmacokinetics and Pharmacodynamics 32, 521-545.

[21] Welshons, W. V., Thayer, K. A., Judy, B. M., Taylor, J. A., Curran, E. M., and Saal, F. S. (2003). *Large Effects From Small Exposures. I. Mechanisms for Endocrine-Disrupting Chemicals with Estrogenic Activity*. Environmental Health Perspectives 111, 994–1006.

[22] Wiens, D. P. (2009). *Robust discrimination designs*, Journal of the Royal Statistical Society: Series B 71, 805-829.

APPENDIX. R CODE

```
#Estimated Model 2 parameters' range:

#theta1 [-10,0]

#theta2 [-0.01, -1]

#True model with Theta1:

#theta1 = 4.63 theta1 = 0.175 theta1 = -6.69

#theta2 = 1.23 theta2 = 0.277 theta2 = -0.60

#theta3 = 0.07 theta3 = 0.024 theta3 = 0.01

#Initial value#

x0 = c(-14, -10, -6, -4)

n0 = length(x0)

w = rep(1/n0, n0)

(D = rbind(x0, w))

p = 1

n = 1

#define M

#number of parameter

k=3

#value of parameter (change to one of sets above; also change values in the

loop)

alpha=-6.69

beta=-.6

gamma=.01
```

```
#define function
f<-function(x)
{matrix(c(exp((-1/2)*(gamma*(x^2)+beta*x+alpha)^2),
          x*exp((-1/2)*(gamma*(x^2)+beta*x+alpha)^2),
          (x^2)*exp((-
1/2)*(gamma*(x^2)+beta*x+alpha)^2)),nrow=3,ncol=1,byrow=F)}

# define information matrix
A1<-rep(0,n0)
A2<-rep(0,n0)
A3<-rep(0,n0)
A4<-rep(0,n0)
A5<-rep(0,n0)

for (i in 1:n0)
{
  A1[i]=w[i]*exp((-1)*(gamma*(x0[i]^2)+beta*x0[i]+alpha)^2)
  A2[i]=x0[i]*A1[i]
  A3[i]=x0[i]^2*A1[i]
  A4[i]=x0[i]^3*A1[i]
  A5[i]=x0[i]^4*A1[i]
}
M0=matrix(c(sum(A1),sum(A2),sum(A3),sum(A2),sum(A3),
            sum(A4),sum(A3),sum(A4),sum(A5)),nrow=3,ncol=3,byrow=F)
```

```
while(p > .0001){
  Rt1 = c(-5, 5)
  Rt2 = c(-1, 1)
  s = c(1, .1)
  while(max(s) > .01){
    theta1 = seq(Rt1[1], Rt1[2], s[1])
    theta2 = seq(Rt2[1], Rt2[2], s[2])
    eta1 = function(x)
      {pnorm(-(-6.69 - .6*x + 0.01*x^2))}
    mod2 = expand.grid(theta1, theta2)
    diff = rep(NA, nrow(mod2))
    for (i in 1:nrow(mod2)){
      diff[i] = sum(w*(sapply(x0, eta1) - pnorm(-(mod2[i,1]
                                                  + mod2[i,2]*x0)))^2)
    }
    (theta1hat = mod2[which.min(diff),1]); (theta2hat = mod2[
      which.min(diff),2])
    s = s/2
    Rt1[1] = theta1hat - s[1]
    Rt1[2] = theta1hat + s[1]
    Rt2[1] = theta2hat - s[2]
    Rt2[2] = theta2hat + s[2]
  }
  theta1hat
  theta2hat
  x = seq(-14, -4, .1)
  a = rep(NA, length(x))
  diff2 = rep(NA, length(x))
```

```
  for (j in 1:length(x)){
    diff2[j] = (.5/(sum(w*(sapply(x0, eta1) - pnorm(-(theta1hat +
theta2hat*x0)))^2)))*
               (sapply(x[j], eta1) - pnorm(-(theta1hat + theta2hat *x[j])))^2
+
               (.5/3)*(t(f(x[j]))%*%solve(M0)%*%f(x[j]))
  }
  (anew = x[which.max(diff2)])
  p = abs(max(diff2)-1)

  x0 = c(x0,anew)
  alpha2 = 1/(n + 1)
  w = c((1 - alpha2) * w, alpha2)
  n = n + 1
  n0 = length(x0)
  A1<-rep(0,n0)
  A2<-rep(0,n0)
  A3<-rep(0,n0)
  A4<-rep(0,n0)
  A5<-rep(0,n0)
```

```
  for (i in 1:n0)
  {
    A1[i]=w[i]*exp((-1)*(gamma*(x0[i]^2)+beta*x0[i]+alpha)^2)
    A2[i]=x0[i]*A1[i]
    A3[i]=x0[i]^2*A1[i]
    A4[i]=x0[i]^3*A1[i]
    A5[i]=x0[i]^4*A1[i]
  }
  M0=matrix(c(sum(A1),sum(A2),sum(A3),sum(A2),sum(A3),
              sum(A4),sum(A3),sum(A4),sum(A5)),nrow=3,ncol=3,byrow=F)
  print(p)
  D = rbind(x0,w)
}
#Summarize the result
DT_optimal = by(D[2,], D[1,], FUN = sum)
DT_optimal
#plot the two models
theta1hat
theta2hat
y = pnorm(-(theta1hat + theta2hat*x))
y1 = eta1(x)
plot(x, y, cex = 0.3, ylim = c(0, 1), col = "blue",
     type="l",pch=1)
lines(x, y1,type="l",pch=1)
cont.txt=c(expression(eta1),
           expression(eta2))
legend("bottomright", legend=cont.txt,
       col=c(1,4), lwd=1, lty=c(1,1))
```

```
#Verify DT-optimal
X = D[1,]
W = D[2,]
x = seq(-14, -4, .1)
ds = rep(0,length(x))
for (i in 1:length(x))
{ds[i] = (.5/(sum(w*(sapply(x0, eta1) - pnorm(-(theta1hat +
theta2hat*x0)))^2)))*
         (sapply(x[i], eta1) - pnorm(-(theta1hat + theta2hat *x[i])))^2 +
         (.5/3)*(t(f(x[i]))%*%solve(M0)%*%f(x[i]))}

cont.txt2 = (expression(psi))
plot(x, ds, cex = 0.3, type = "l", pch=1, ylab = cont.txt2,
     ylim = c(-0.09400517,1))
abline(h = 1,pch =1, lty = 3)
abline(v = c(-11.1, -9.5, -8.1), pch =1, lty = 3)


#Calculating Efficiency
#Obj function:
sum(w*(sapply(x0, eta1) - pnorm(-(theta1hat + theta2hat*x0)))^2)

#DetM0
det(M0)
```

```
#########
#THETA-1#
#########
# T-optimal
# x0 = c(-14, -9, -4)
# w = c(0.25, 0.50, 0.25)
# Obj = 0.09395754
# DetM0 = 151.6461

# DT-optimal
# x0 = c(-13.9, -9, -4)
# w = c(0.28, 0.45, 0.27)
# Obj = 0.08957937
# DetM0 = 189.0502
# theta1hat = 0.03125
# theta2hat = -0.0125

# D-optimal
# x0 = c(-13.2, -10.3, -7.2, -4.3)
# w = c(0.32, 0.17, 18, 0.33)
# Obj = 0.04710415
# DetM0 = 2.210567e-07

(Eff.DT.T =  0.08957937/0.09395754) # 0.9534027
(Eff.DT.D = (189.0502/308.7837)^(1/3)) # 0.8491301
(Eff.T.D = (151.6461/308.7837)^(1/3)) # 0.7889672
(Eff.D.T = 0.04710415/0.09395754) # 0.5013344
```

```
#########
#THETA-2#
#########
# T-optimal
# x0 = c(-14, -9.2, -4)
# w = c(0.27, 0.43, 0.30)
# Obj = 0.01029732
# DetM0 = 524.7601

# DT-optimal
# x0 = c(-14, -9.3, -4)
# w = c(0.30, 0.40, 0.30)
# Obj = 0.01007072
# DetM0 = 543.74
# theta1hat = -1.453125
# theta2hat = -0.1515625

# D-optimal
# x0 = c(-13.7, -9.5, -4)
# w = c(0.33, 0.33, 0.33)
# Obj = 0.008726613
# DetM0 = 577.8466

(Eff.DT.T =  0.01007072/0.01029732) # 0.9779943
(Eff.DT.D = (543.74/577.8466)^(1/3)) # 0.9799252
(Eff.T.D = (524.7601/577.8466)^(1/3)) # 0.968388
(Eff.D.T = 0.008726613/0.01029732) # 0.8474645
```

```
#########
#THETA-3#
#########
# T-optimal
# x0 = c(-11.1, -8.2)
# w = c(0.43, 0.57)
# Obj = 0.004559819
# DetM0 = 0.0002780678

# DT-optimal
# x0 = c(-11.1, -9.5, -8.1)
# w = c(0.41, 0.18, 0.41)
# Obj = 0.003777845
# DetM0 = 0.07830435
# theta1hat = -5.734375
# theta2hat = -0.5984375

# D-optimal
# x0 = c(-11.1, -9.6, -8)
# w = c(0.34, 0.33, 0.33)
# Obj = 0.002985904
# DetM0 = 0.1005924

(Eff.DT.T =  0.003777845/0.004559819) # 0.8285077
(Eff.DT.D = (0.07830435/0.1005924)^(1/3)) # 0.9198992
(Eff.T.D = (0.0002780678/0.1005924)^(1/3)) # 0.1403444
(Eff.D.T = 0.002985904/0.004559819) # 0.6548295
```