D_S-OPTIMAL DESIGN FOR MODEL DISCRIMINATION

IN A PROBIT MODEL

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

By

Ruifeng Liu

In Partial Fulfillment
for the Degree of
MASTER OF SCIENCE

Major Department:
Statistics
Master of Applied Statistics

May 2013

Fargo, North Dakota

North Dakota State University
Graduate School

Title

D_S-OPTIMAL DESIGN FOR MODEL DISCRIMINATION IN A PROBIT MODEL

By

Ruifeng Liu

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. Rhonda Magel

Dr. Gang Shen

Dr. Kevin K.Sedivec

Approved:

5/8/2013 — Date

Dr. Rhonda Magel — Department Chair

ABSTRACT

In toxicology studies, dose response functions with a downturn at higher doses are often observed. For such response functions, researchers often want to see if the downturn of the response is significant. A probit model with a quadratic term is adopted to demonstrate the dose response with a downturn. Under the probit model, we obtain optimal designs to study the significance of the downturn and their efficiencies are compared. Our approach identifies the upper bound of the number of optimal design points and searches for the optimal design numerically based on the upper bound.

ACKNOWLEDGMENTS

I remind back my journey of pursuing Master Degree of Applied Statistics for last three years in Statistics Department of North Dakota State University and I realized that I own lots of debt to my department, my major advisor, my advisory committee members, and my family and friends. Thanks went to Statistics Department for this valuable further education opportunity and logistic help along my studies.

Dr. Seung Won Hyun, my major advisor, was much appreciated for his guidance, courage, and academic input for my study. His attitude and energy to teach and to do the research will foster my career goals and I will cherish him as my major advisor for my whole life.

I want to thank Dr.Rhonda Magel for her acceptance of application and following study plan suggestions. I appreciate your willingness being my committee member.

I must thank Dr. Kevin Sedivec for his scholarship opportunity and real on-campus job experience, without his financial help, my study will be much harder for sure.

Dr. Gang Shen, thank you for your willingness being my committee member. Your credible suggestions, criticisms, and challenges to my study and paper were appreciated.

Finally, I would like to thank my parents, sisters, brothers, parents-in-law for their love and support. Paul and Anne Nyren, my dear American family members, thank you so much for your warm welcome to your family. For a foreign student, your love and courage sustained my interest to study and my energy to finish my degree. Thank you all, without your inputs, it is impossible for me to finish my studies and my paper. I especially thank my loved husband, Dr.Guojie Wang, Without his endless encouragement and support, I was unable to finish my research paper. I thank our adorable son, Michael Wang, who brings us fun and happiness.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

Dose-response functions with a downturn are observed in many toxicology studies (see Margolin et al., 1981; Welshons et al., 2003; Bhatti et al., 2010). When the dose-response function has a downturn at higher dose levels, researcher often want to study whether or not the downturn of the response is significant. In this paper, optimal designs to study the downturn at higher dosage levels are obtained.

Optimal design is a class of experimental design. It provides precise parameter estimates by specifying design points to be used and identifying the distribution of samples over these selected design points. Also, it is an efficient way to estimate appropriate parameters while lowering the cost. Optimal design uses different optimality criterion based on the goal of experiment. The goal in this paper is to study the downturn of dose-response effectively. Ting (2006) states that adding a quadratic term of the dose can describe the downturn well. The downturn of dose-response can be studied by discriminating between two rival nonlinear models: one contains the quadratic term and the other one does not. Many studies on the optimal design have been done on discriminating among the competing models (Waterhouse et al., 2008; Atkinson and Fedorov, 1975; Atkinson et al., 2008; Dette et al., 2010; O'Brein, 2005). Dette (1994) found the optimal designs for model discrimination in the polynomial regression model. Also, optimal discrimination designs for trigonometric and Fourier regression models were found (see Dette and Melas, 2005; Dette and Roeder, 1997). The two rival models considered here are nested model. In this case, Ds-optimality can be used to find the optimal design for discriminating between the two models. Ds-optimal design provides accurate estimation of the quadratic term that distinguishes between the two models.

In this paper, a probit model with a quadratic term is adopted to demonstrate the dose response with a downturn at higher dose levels. A probit model is often used in toxicology study and the probit model with a quadratic term provides a good fit to the dose response functions with a downturn at high doses (see Hyun, 2013). The probit model can be adopted when the response takes any values between 0 and 1. For the example of such responses, DNA (mcg/well) for log concentration of estradiol in Welshons et al. (2003) have responses between 0 and 1. Under the probit model, Ds-optimal design for discriminating the two models is obtained. We also want to see what other designs work well for the model discrimination. The uniform design and D-optimal design are considered. The uniform design is a traditional design that can be used without any previous knowledge. D-optimal design is one of the most widely used optimal designs. In order to find the optimal designs, we first find the upper bounds of optimal designs using theorems from Hyun et al. (2013). Then Ds- and D-optimal designs are searched for using a numerical algorithm called V-algorithm (Fedorov, 1972). To see the performances of designs or the model discrimination, their efficiencies are obtained and compared.

Chapter 2 presents background theory about this study. Chapter 3 introduces the dose response functions that have a downturn at higher doses. In Chapter 4 the uniform design is presented, and Ds- and D-optimal designs are obtained using the V-algorithm. Chapter 5 presents the efficiencies of all the obtained designs and compares their efficiencies to evaluate the performances of the designs. Finally, a conclusion is presented based on the results along with a plan for future studies.

CHAPTER 2. BACKGROUND

Many researchers are focused on the toxicity studies which are related to environmental concerns and health issues. In toxicology, the dose-response can be characterized as bell-shaped or S-shaped distribution curves. To find the distribution curves, we need to estimate model parameters as precise as possible. Minimizing the variances of model parameter estimates leads us to estimating the parameters accurately and ensuring unbiased valid results. Here, optimal design is used to minimize the variances.

Optimal design is widely applied in many areas of statistics. It is a very flexible, and powerful experimental design. Optimal design specifies the dose levels we should use and how to assign the subjects on these levels in the most efficient manner. It provides unbiased parameter estimates with valid results while lowering cost. There are several criteria that are used to identify optimal designs. Optimality criterion is denoted by Ψ which is a convex function of the Fisher information matrix in general. Because the Fisher information matrix truly depends on unknown values of the model parameters, the parameter values must be specified to obtain optimal designs. So the optimal designs provide the best design that minimizes or maximizes the optimality criterion for the given values of the parameters. Here, we present two optimality criterion used to search for optimal designs.

2.1. Optimality Criteria

We considered a situation with the regression model in which observations are given by

$$y_{ij} = \mu(\Theta, x_i) + \varepsilon_{ij} \quad (i = 1, 2, \ldots, k; \quad j = 1, 2, \ldots, n_i)$$

—$\mu(\Theta, x_i)$: mean response functions of Θ and x_i,

—x_i : the ith log dose,

—Θ : the vector of the model parameters,

—n_i : number of subjects allocated to x_i,

$$\sum_{i=1}^{k} n_i = N$$

—ε_{ij} are independently normally distributed with mean of zero and unknown constant variance of σ^2.

A design is represented by the measure ξ over x_i,

$$\xi = \begin{pmatrix} x_1 & x_2 & \dots & x_k \\ \omega_1 & \omega_2 & \dots & \omega_k \end{pmatrix},$$

where the first line gives the value of the design points and ω_i represents the design weight corresponding to design point x_i, where $\omega_i = \frac{n_i}{N}$. By Taylor series expansion, the Fisher information matrix for Θ :

$$M(\xi; \Theta) = \sum_{i=1}^{k} \omega_i f(x) f^T(x),$$

where $f(x) = (\frac{\partial\mu(\theta,x)}{\partial\theta_1}, \frac{\partial\mu(\theta,x)}{\partial\theta_2}, \dots, \frac{\partial\mu(\theta,x)}{\partial\theta_k})^T$, $f^T(x)$ is a transpose of $f(x)$.

To ensure minimum-variance, optimal design maximizes its criterion $\Psi\{M(\xi, \Theta)\}$. Now the problem becomes to obtain the design ξ that maximize $\Psi\{M(\xi, \Theta)\}$.

2.2. D-optimality

D-optimal design works well when the interest is in estimating parameters in the model. It maximizes the determinant of the Fisher information matrix, which means minimizing the joint confident region of estimating parameters. D-optimal design criterion is

$$\Psi = |M(\xi; \Theta)|.$$

2.3. Ds-optimality

Ds-optimal design is applied when the goal of research is estimating a subset of the parameters precisely. For a nonlinear response, $E(y_{ij}) = \mu(\Theta, x_i)$, the parameter vector Θ is of dimension $p \times 1$, it can be partitioned into two sections $\begin{pmatrix} \Theta_1 \\ \Theta_2 \end{pmatrix}$. Θ_1 is of dimension $p_1 \times 1$ and Θ_2 is of dimension $(p - p_1) \times 1$. Then, the Fisher information matrix can be partitioned as

$$M(\xi; \Theta) = \begin{pmatrix} M_{11}(\xi; \Theta) & M_{12}(\xi; \Theta) \\ M_{21}(\xi; \Theta) & M_{22}(\xi; \Theta) \end{pmatrix},$$

where

$$M_{st}(\xi; \Theta) = \sum_{i=1}^{k} \omega_i f_s(x_i)^T f_t(x_i), \quad (s, t = 1, 2)$$

here

$$f_1(x) = \frac{\partial \mu(\Theta, x)}{\partial \Theta_1}, \qquad f_2(x) = \frac{\partial \mu(\Theta, x)}{\partial \Theta_2}.$$

Suppose we are interested in estimating Θ_2. The covariance matrix for Θ_2 is

$$[M_{22}(\xi; \Theta) - M_{21}(\xi; \Theta) M_{11}^{-1}(\xi; \Theta) M_{12}(\xi; \Theta)]^{-1},$$

where $M_{11}^{-1}(\xi; \Theta)$ is the inverse of $M_{11}(\xi; \Theta)$.

Ds-optimal design for estimating Θ_2 maximizes the determinant of the information matrix of Θ_2 (Atkinson and Donev, 1992),

$$\Psi = |M_{22}(\xi; \Theta) - M_{21}(\xi; \Theta) M_{11}^{-1}(\xi; \Theta) M_{12}(\xi; \Theta)| = \frac{M(\xi, \Theta)}{M_{11}(\xi; \Theta)}.$$

2.4. The General Equivalence Theorem

The General Equivalence Theorem (Kiefer, 1958; cf. Pukelsheim, 2006) provides methods for identifying optimal designs and verifying it. For a nonlinear model when researcher only consider locally optimal design, equivalence theorems are formulated in respect to a compact and convex set of matrices. The General Equivalence Theorem can be viewed as an application of the result that derivatives are zero at the minimum of an objective function over region. Here, the objective function depends on the design ξ through the information matrix $M(\xi;\Theta)$. The following is General Equivalence Theorem stated in Atkinson et al. (2007). Let the measure $\bar{\xi}$ put unit mass at the point x and let the measure $\xi^{'}$ be given by

$$\xi' = (1-\alpha)\xi + \alpha\bar{\xi}.$$

Then,

$$M(\xi';\Theta) = (1-\alpha)M(\xi;\Theta) + \alpha M(\bar{\xi};\Theta)$$

Accordingly, the derivative of Ψ in the direction $\bar{\xi}$ is

$$\phi(x,\xi) = \lim_{\alpha\to 0^+} \frac{1}{\alpha}[\Psi\{(1-\alpha)M(\xi;\Theta) + \alpha M(\bar{\xi};\Theta)\} - \Psi\{M(\xi;\Theta)\}].$$

The necessary and sufficient conditions for ξ^* to be optimal design are fulfillment of following inequalities:

1. the design ξ^* minimizes $\Psi\{M(\xi;\Theta)\}$;
2. ξ^* maximizes the minimum over design space of $\phi(x,\xi^*)$;
3. The minimum over design space $\phi(x,\xi^*) = 0$ at the points which support our design ξ^* .

White (1973) extended the General Equivalence Theorem to nonlinear model. Thus, we will use this method to verify D and Ds-optimal design which we will discuss in Chapter 4.

2.5. The V-algorithm

V-algorithm (Fedorov, 1972) is a numerical algorithm to search optimal designs based on the General Equivalence Theorem. It is an efficient and popular method to obtain optimal designs. The details of this algorithm will be given later in section 4.

CHAPTER 3. MODEL

In this chapter, we will introduce the model that demonstrates dose-response functions with a downturn. Then, we find the Fisher information matrix which is used to search D- and Ds-optimal designs.

We define $\Theta = (\theta_1, \theta_2, \theta_3)$, suppose a response y_{ij} is continuous, it follows

$$y_{ij} = \mu(\Theta, x_i) + \varepsilon_{ij}; \quad \varepsilon_{ij} \sim N(0, \sigma^2), \tag{1}$$

where x_i is the ith log dose, $\mu(\Theta, x_i) = \Phi(-(\theta_1 + \theta_2 x_i + \theta_3 x_i^2))$, Φ is the cumulative distribution function of standard normal distribution; and σ^2 is unknown. To obtain optimal design, we need construct the Fisher information matrix for Θ.

By Taylor expansion, an approximate Fisher information matrix for Θ can be written as

$$M(\xi;\Theta) = \frac{1}{\sigma^2}\sum_{i=1}^{k} \omega_i f(x_i) f^T(x_i),$$

where $f(x)$ can be expressed as

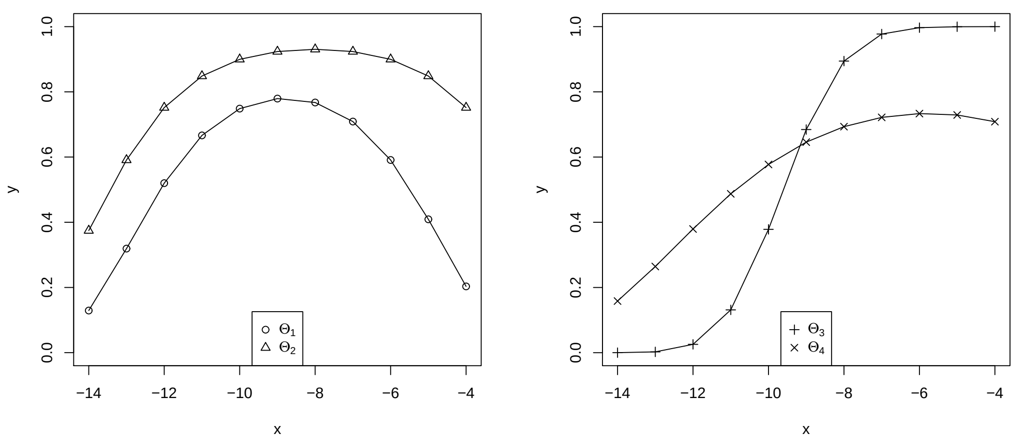
$$f(x) = \exp(-\frac{(\theta_1 + \theta_2 x + \theta_3 x^2)^2}{2})(1 \qquad x \qquad x^2)^T.$$

The goal of research is to find Ds- and D-optimal designs for given model parameters. As mentioned earlier, the values of model parameters need to be specified in order to obtain optimal designs. Here we consider four different sets of values of parameters to study Ds- and D-optimal designs (see Table 1). They provide four different shapes of dose-response function. We classify these four sets of the parameters into two categories. Each category has two sets of the parameters. One category (Θ_1 and Θ_2) shows dose-response curves with a strong downturn at high dose levels for the probit model (see Figure 1 (a)), and the other one (Θ_3 and Θ_4)

has a weak downturn at high dose levels (see Figure 1 (b)).

Table 1: Four sets of parameters values

Θ	Parameter values
Θ_1	{4.63, 1,23, 0.07}
Θ_2	{1.72, 0.80, 0.05}
Θ_3	{0.175, 0.277, 0.024}
Θ_4	{-6.69, -0.60, 0.01}



(a) A strong downturn

(b) A weak downturn

Figure 1: Four different shapes of probit model with a quadratic term: Θ

CHAPTER 4. OPTIMAL DESIGNS

In this chapter, we will discuss Ds- and D-optimal designs under the probit model. Here, we consider four sets of the values of the parameters as mentioned in the previous section. First, the upper bounds of the optimal designs are obtained using next theorem.

4.1. The Upper Bounds on Design Points

Theorem. Under the probit model, regardless of the values of the model parameters, the upper bound of design points that maximize Ds-optimality criterion is 4.

Proof. The proof of follows directly from Hyun et al. (2013). The paper identify the upper bound of optimal design points based on the number of taking derivatives to the objective function from General Equivalence Theorem to reach a quadratic form. Under Ds-optimality criterion, the differentiable function $F(x)$ comes from General Equivalence Theorem is:

$$f(x)[M(\xi^*;\Theta)]^{-1}f(x)^T - f_1(x)[M_{11}(\xi^*;\Theta)]^{-1}f_1(x)^T,$$

Where

$$f(x) = \exp(-\frac{(\theta_1 + \theta_2 x + \theta_3 x^2)^2}{2})(1 \qquad x \qquad x^2)^T.$$

$$f_1(x) = \exp(-\frac{(\theta_1 + \theta_2 x + \theta_3 x^2)^2}{2})(1 \qquad x)^T.$$

Here, $[M(\xi^*;\Theta)]^{-1}$ and $[M_{11}(\xi^*;\Theta)]^{-1}$ are constants since ξ^* is Ds-optimal design. So the $F(x)$ becomes:

$$F(x) = f(x)\begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix} f(x)^T - f_1(x)\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} f_1(x)^T$$

$$= \exp\{-(\theta_1 + \theta_2 x + \theta_3 x^2)^2\}\{(c_{33}x^4 + 2c_{23}x^3 + (2c_{13} + c_{22})x^2 + 2c_{12}x + c_{11}) - (c_{22}x^2 + 2c_{12}x + c_{11})\}$$

$$= \exp\{-(\theta_1 + \theta_2 x + \theta_3 x^2)^2\}\{c_{33}x^4 + 2c_{23}x^3 + 2c_{13}x^2\} \tag{2}$$

The derivative of $F^{'}(x) = G_1(x) = P_1(x)F_1(x)$, here $P_1(x) = \exp\{-(\theta_1 + \theta_2 x + \theta_3 x^2)^2\}$ is a positive, we ignore this positive factor. Continue to take derivative of $F_1(x)$, the sixth derivative becomes the equation in quadratic form with respect to x:

$$\frac{G_6(x)}{P_6(x)} = -210\theta_3^2 c_{33}x^2 - (90\theta_2\theta_3 c_{33} + 120\theta_3^2 c_{23})x - 10\theta_1\theta_3 c_{33} - 5\theta_2^2 c_{33} - 6\theta_2\theta_3 c_{13} - 20\theta_3^2 c_{13}$$

where c_{33} is positive since it is a diagonal element of the covariance matrix. So, the sign of coefficient for x^2 is negative. Based on Theorem 2 in Hyun et al. (2013), we have no more than $\frac{6}{2} + 1 = 4$ upper bounds.

In the case of D-optimality, Hyun et al. (2013) show that the upper bound of D-optimal design point under the probit model is 4.

4.2. Uniform Design

Welshons et al. (2003) considered a uniform design for studying dose-response function with a downturn. In general, the uniform design is used as a traditional design when there is no information available for the study. The uniform design has 11 equally spaced design points scattered in the design space [-14,-4], with 9% of the subjects assigned to each point. The uniform design is:

$$\xi_u = \begin{pmatrix} -14 & -13 & -12 & -11 & -10 & -9 & -8 & -7 & -6 & -5 & -4 \\ \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} & \frac{1}{11} \end{pmatrix}.$$

4.3. Ds-optimal Design

Ds-optimal criteria is applied when a researcher want to estimate a particular subset of the parameters of a given model. When response functions have a downturn at high doses, the downturn can be described by adding the quadratic term to the model. The significance of the downturn can be studied by discriminating between nested models. In order to study the nested models effectively, we need to minimize the variance of estimating the coefficient of the quadratic term. So, Ds-optimal design is appropriate in this case, because we are only interested in estimating θ_3 precisely.

The Fisher information matrix for Θ can be partitioned as:

$$M(\xi;\Theta) = \begin{pmatrix} M_{11}(\xi;\Theta) & M_{12}(\xi;\Theta) \\ M_{21}(\xi;\Theta) & M_{22}(\xi;\Theta) \end{pmatrix},$$

where

$$M_{st}(\xi;\Theta) = \sum_{i=1}^{k} \omega_i f_s(x_i) f_t(x_i)^T, \quad (s,t=1,2)$$

here

$$f_1(x) = \exp(-\frac{(\theta_1+\theta_2 x+\theta_3 x^2)^2}{2})(1 \qquad x)^T,$$

$$f_2(x) = \exp(-\frac{(\theta_1+\theta_2 x+\theta_3 x^2)^2}{2})x^2.$$

To find Ds-optimal design, we use V-algorithm. First, set the values of parameters and the initial design ξ_0 based on the upper bound of Ds-optimal design from Theorem. For the initial design, 4 design points in the design space $[-14,-4]$ with equal weights are used:

$$\xi_0 = \begin{pmatrix} -14 & -10 & -8 & -4 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}.$$

Initial information matrices $M(\xi_0;\Theta)$ and $M_{11}(\xi_0;\Theta)$ are evaluated using the initial design ξ_0. Based on the General Equivalence Theorem, a next design point x_{n+1} is obtained by:

$$\begin{aligned}\bar{d}_{s_n} &= \max_{x\in[-14,-4]}(f(x)[M(\xi_n;\Theta)]^{-1}f(x)^T - f_1(x)[M_{11}(\xi_n;\Theta)]^{-1}f_1(x)^T) \\ &= f(x_{n+1})[M(\xi_n;\Theta)]^{-1}f(x_{n+1})^T - f_1(x_{n+1})[M_{11}(\xi_n;\Theta)]^{-1}f_1(x_{n+1})^T\end{aligned}$$

For each step, $M(\xi_{n+1};\Theta)$, $M_{11}(\xi_{n+1};\Theta)$ are updated by following equations;

$$M(\xi_{n+1};\Theta) = (1-\alpha_{n+1})M(\xi_n;\Theta) + \alpha_{n+1}f(x_{n+1})f(x_{n+1})^T,$$

$$M_{11}(\xi_{n+1};\Theta) = (1-\alpha_{n+1})M_{11}(\xi_n;\Theta) + \alpha_{n+1}f_1(x_{n+1})f_1(x_{n+1})^T.$$

where α_{n+1} can be set as $\frac{1}{n+1}$. The algorithm is stopped when ds_n is close to $p-p_1$, here p and p_1 are corresponding to the dimension of the $M(\xi;\Theta)$ and $M_{11}(\xi;\Theta)$. In our case, $p-p_1=1$.

$$ds_n = f(x)[M(\xi^*;\Theta)]^{-1}f(x)^T - f_1(x)[M_{11}(\xi^*;\Theta)]^{-1}f_1(x)^T \leq 1 \tag{3}$$

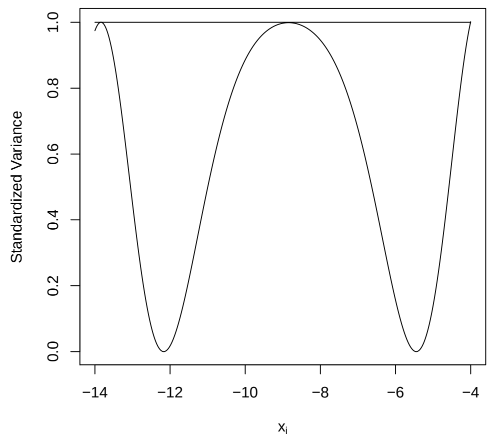
When ξ^* is Ds-optimal design, (3) is satisfied and the equality holds if x is one of Ds-optimal design points.

V-algorithm is used to obtain Ds-optimal designs under the probit model with the four sets of the values of the parameters (see Table 2). For example, Ds-optimal design under the probit model with $\Theta = \Theta_1$ allocates 28.5% of subjects at $x_1 = -13.84$, 46.7% at $x_2 = -8.84$ and 24.8% at $x_3 = -4$. This design minimizes $Var(\widehat{\theta_3})$ under the probit model with the given parameter values. Figure 2 shows that the plot of standardized variances over design space hits the maximum 1 when design points are Ds-optimal design points. It verifies the Ds-optimal design using General

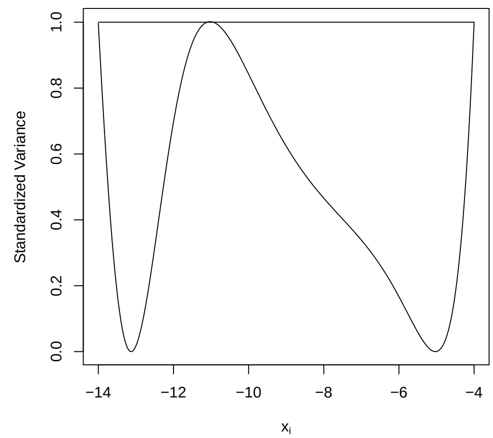
Equivalence Theorem. As you can see in Table 2, Ds-optimal designs vary with the different values of the parameters.

Table 2: Ds-optimal designs under the different values of the parameters

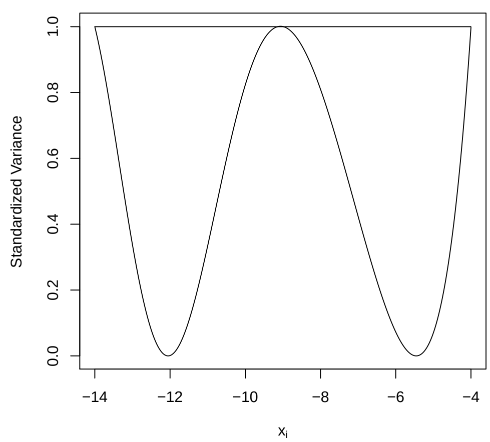
Θ	Ds-optimal design
$\Theta_1 = \{4.63, 1.23, 0.07\}$	$\begin{pmatrix} -13.84 & -8.84 & -4 \\ 0.285 & 0.467 & 0.248 \end{pmatrix}$
$\Theta_2 = \{1.72, 0.80, 0.05\}$	$\begin{pmatrix} -14 & -11.02 & -4 \\ 0.264 & 0.249 & 0.134 \end{pmatrix}$
$\Theta_3 = \{0.175, 0.277, 0.024\}$	$\begin{pmatrix} -14 & -9.06 & -4 \\ 0.337 & 0.431 & 0.232 \end{pmatrix}$
$\Theta_4 = \{-6.69, -0.60, 0.01\}$	$\begin{pmatrix} -11.54 & -9.57 & -7.49 \\ 0.381 & 0.217 & 0.402 \end{pmatrix}$



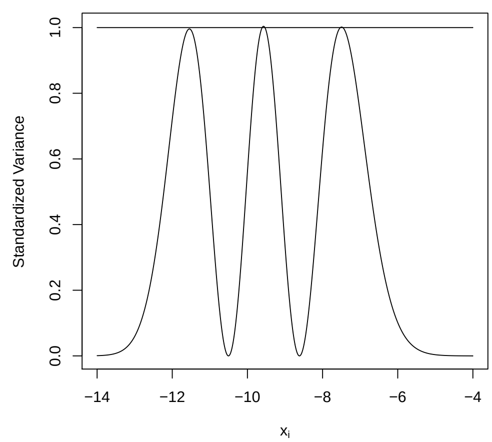
(a) $\Theta = \Theta_1$



(b) $\Theta = \Theta_2$



(c) $\Theta = \Theta_3$



(d) $\Theta = \Theta_4$

Figure 2: Verifying Ds-optimal design using G.E.T

4.4. D-optimal Design

D-optimal design is used when the goal is to estimate parameters in the model. D-optimal design maximizes the determinant of the Fisher information matrix $M(\xi;\Theta)$.

$$\Psi = |M(\xi;\Theta)| \tag{4}$$

To obtain D-optimal design, V-algorithm based on the General Equivalence Theorem is used. First, we set the values of the parameters and initial design ξ_0 based on the upper bound of D-optimal design from Hyun et al.(2013):

$$\xi_0 = \begin{pmatrix} -14 & -10 & -8 & -4 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}.$$

The initial information matrix $M(\xi_0; \Theta)$ is evaluated using the initial design. Then the next design point x_{n+1} among the candidate points is obtained by the equation below:

$$\begin{aligned} \bar{d}_n &= \max_{x \in [-14,-4]} (f(x)[M(\xi_n; \Theta)]^{-1} f(x)^T) \\ &= f(x_{n+1})[M(\xi_n; \Theta)]^{-1} f(x_{n+1})^T \end{aligned}$$

The Fisher information matrix which is correspond to the new design point x_{n+1} is replaced by:

$$M(\xi_{n+1}; \theta) = (1 - \alpha_{n+1}) M(\xi_n, \theta) + \alpha_{n+1} f(x_{n+1}) f(x_{n+1})^T,$$

where

$$\alpha_{n+1} = \frac{\bar{d}_n - 3}{3(\bar{d}_n - 1)}.$$

V-algorithm continue updating the design points x_{n+1} and α_{n+1} until the difference of function:

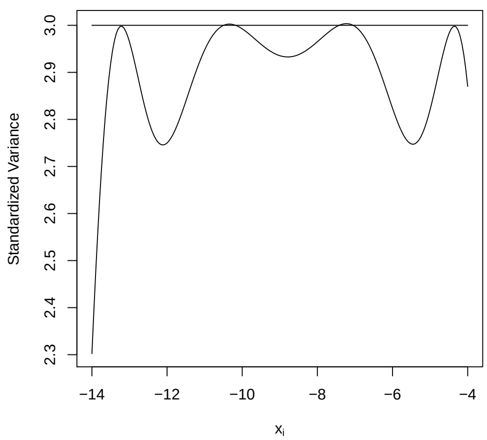
$$f(x_{n+1})[M(\xi_{n+1}; \theta)]^{-1} f(x_{n+1})^T - 3 \tag{5}$$

is close to 0. When the difference (5) is close to 0, the design ξ_{n+1} becomes D-optimal design that maximize the D-optimality criterion (4). The optimal design under D-optimal criteria is comprised of the design points when the design points and weights satisfy the condition above.

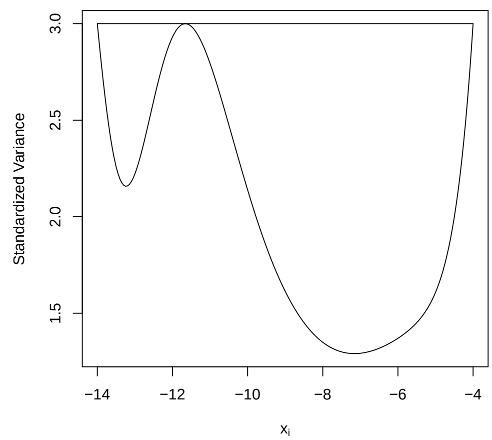
We obtain the D-optimal designs under the probit model with the four sets of the parameters in Table 3. For example, the D-optimal design under the values of the parameters Θ_1 is to assign 32.3% of the subjects to the design points -13.21 and -4.36, and 17.7% of subjects to each of the design point -10.34 and -7.24. This design maximizes $|M(\xi;\Theta)|$ under the probit model with the given values of the parameters. We also verify the D-optimal designs using General Equivalence Theorem, plot a standardized variance of predicted response over design space [-14, -4] (Figure 3). The plot hits the maximum 3 when the points are D-optimal design points.

Table 3: D-optimal design under the different values of the parameters

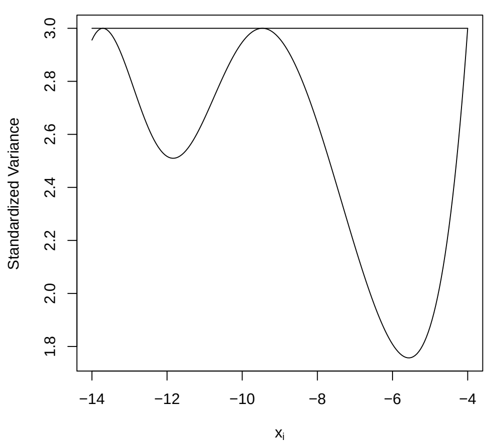
Θ	D-optimal design
$\Theta_1 = \{4.63, 1.23, 0.07\}$	$\begin{pmatrix} -13.22 & -10.34 & -7.23 & -4.35 \\ 0.323 & 0.177 & 0.177 & 0.323 \end{pmatrix}$
$\Theta_2 = \{1.72, 0.80, 0.05\}$	$\begin{pmatrix} -14 & -11.66 & -4 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$
$\Theta_3 = \{0.175, 0.277, 0.024\}$	$\begin{pmatrix} -13.71 & -9.47 & -4 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$
$\Theta_4 = \{-6.69, -0.60, 0.01\}$	$\begin{pmatrix} -11.09 & -9.57 & -7.99 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$



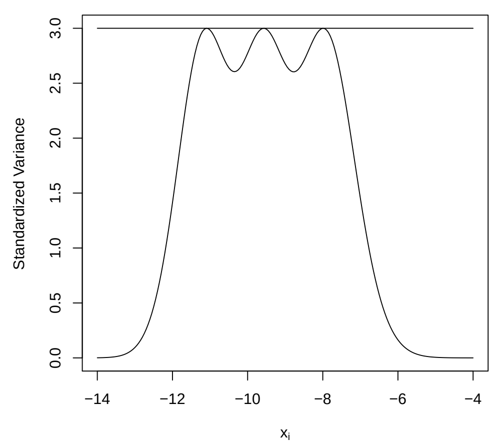
(a) $\Theta = \Theta_1$



(b) $\Theta = \Theta_2$



(c) $\Theta = \Theta_3$



(d) $\Theta = \Theta_4$

Figure 3: Verifying D-optimal design using G.E.T

CHAPTER 5. EFFICIENCY

In the previous section, we discussed the uniform design, the Ds-optimal design and the D-optimal design under the probit model. In this section, we will discuss the efficiencies of the optimal designs. Design efficiency is a measure which compares any design to the optimal design. For example, the efficiency of design ξ is ℓ, the design ξ would need $100(\frac{1}{\ell}-1)\%$ more subjects than the optimal design to achieve as accurate of estimates. Since our goal in this paper is to estimate θ_3 accurately, Ds-efficiency is considered here.

As mentioned earlier, we are interested in estimating θ_3 precisely. Thus, we check the ratio of variance of estimating θ_3 based on design ξ to the minimum variance of estimating θ_3. Let ξ^*_{Ds} and ξ^*_D denote the Ds-optimal design and D-optimal design under the probit model respectively, ξ_u is the uniform design, $A = (0 \quad 0 \quad 1)$, the variance of estimating parameter θ_3 can be approximated by $A^T M^{-1}(\xi;\Theta)A$. Ds-efficiency is the ratio of the variance of $\widehat{\theta}_3$ under the Ds-optimal design ξ^*_{Ds} to the variance of $\widehat{\theta}_3$ under any design. It is calculated by the equation below:

$$Eff_{Ds}(\xi) = \frac{A^T M^{-1}(\xi^*_{Ds};\Theta)A}{A^T M^{-1}(\xi;\Theta)A}.$$

where ξ^*_{Ds} is the Ds-optimal design under the model with different parameters.

Uniform design is the traditional design and D-optimal design works well for estimating model parameters. So we want to know how they works for estimating θ_3. Based on the Ds-efficiency (see Table 4), uniform design performs poorly for the all sets of the parameters. The uniform design requires at least 70% more subjects to provide the same accuracy as the Ds-optimal design does for estimating θ_3. For example, the uniform design under the probit model with true values of the parameters $\Theta = \Theta_2$ needs 101% more subjects to achieve as accurate estimates as Ds-optimal design dose. D-optimal design performs better than uniform design for estimating θ_3. However, it still provides lower efficiency than Ds-optimal design.

Table 4: Ds-efficiencies under the different values of the parameters

Designs	ξ^*_{Ds}		ξ^*_D		ξ_u	
$\Theta = \Theta_1$	1	(0%)	0.673	(48.6%)	0.570	(75.4%)
$\Theta = \Theta_2$	1	(0%)	0.722	(38.5%)	0.540	(85.2%)
$\Theta = \Theta_3$	1	(0%)	0.860	(16.3%)	0.497	(101%)
$\Theta = \Theta_4$	1	(0%)	0.746	(34.0%)	0.330	(203%)

CHAPTER 6. CONCLUSIONS

Optimal design is applied in many research areas. It specifies design points and design weights to minimize the variance of estimating interesting features in the most efficient manner. In toxicology, dose-response functions with a downturn are often observed. For such response functions, researchers often want to study the significance of the downturn of dose-response. In order to study the downturn effectively, model discrimination between nested model is used. The quadratic term of the dose distinguishes the nested models. In this paper, a probit model is adopted to study dose-response functions. Under the probit model, optimal designs for the model discrimination are studied.

Ds-optimal works well for the model discrimination because Ds-optimal design provides minimum variance for estimating the coefficient of the quadratic term. We also checked how the traditional uniform design and D-optimal design works for the model discrimination. As shown in the efficiency, the traditional uniform design performs poorly for the model discrimination. D-optimal design works better than the uniform design but dose not provide as good efficiency as Ds-optimal design.

In the future, we want to study T-optimal design for the model discrimination between the nested models. Waterhouse et al.(2008) mentioned that T-optimal design works better than Ds-optimal design for studying effectively a subset of parameters. Studying the significance of the downturn can be expressed as model discrimination between two models. One describes the response with a downturn and the other one describe the response without the downturn. So optimal designs for model discrimination between other possible models can be considered. As we discussed earlier, optimal designs truly depend on the values of parameters. Therefore, optimal designs which are less sensitive to the value of unknown parameters can be included for my future work.

REFERENCES

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

[2] Atkinson, A.C., Donev, A.N., 1992. *Optimum Experimental Designs*. Oxford University Press.

[3] Atkinson, A. C., Donev, A. N., Tobias,R. D., 2007. *Optimum Experimental Designs with SAS*. Oxford University Press.

[4] Atkinson, A. C., Fedorv, V. V., 1975. *The design of experiments for discriminating between two rival models*. Biomatrika 62, 57-70.

[5] Bhatti, P., Veiga, L. H., Ronckers, C. M., et al., 2010. *Risk of Second Primary Thyroid Cancer after Radiotherapy for a Childhood Cancer in a Large Cohort Study: an Update from the Childhood Cancer Survivor Study*. Radiation Research 174(6a), 741-752.

[6] Brez, F., Hothorn, L. A., 2003. *Statistical Analysis of Monotone or Non-monotone Dose-Response Data from In Vitro Toxicological Assays*. Alternatives To Laboratory Animals 31, 81-96.

[7] Dette, H., 1994. *Discrimination Designs for Polynomial Regression on a Compact Interval*. The Annals of Statistic 22, 890-904.

[8] Dette,H., Melas, V. B., Wong, W. K., 2005. *Optimal Design for Goodness-of-Fit of the Michaelis-Menten Enzyme Kinetic Function*. Journal of the American Statistical Association 472, 1370-1381.

[9] Dette, H., Pepelyshev, A., Shpilev, Piter., Weng kee wong., 2010. *Optimal design for discriminating between dose-response models in toxicology studies*. Bernoulli 16, 1164-1176.

[10] Dette, H., Roeder, I., 1997. *Optimal discrimination designs for multi-factor experiments*. The Annals of Statistic 25, 1161-1175.

[11] Fedorov, V. V., 1972. *Theory of Optimal Experiments*. Academic press, New York.

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

[13] Hyun, S. W., Flournoy, N., Yang, M., 2011. *Optimal designs for response functions with a downturn*. Jounal of Statistical Planning and Inference 141, 559-575.

[14] Hyun, S. W., Yang, M., Flournoy, N., 2013. *A procedure for finding an improved upper bound on the number of optimal design points*. Computational statistics and Data Analysis 58, 276-282.

[15] Kiefer, J., 1958. *On the Nonrandomized Optimality and Randomized Nonoptimality of Symmetrical Designs*. Annals of Mathematical Statistics 29, 675-699.

[16] Margolin, B. H., Kaplan, N., Zeiger, E., 1981. *Statistical analysis of the Ames Salmonella/microsome test*. Proceedings of the National Academy of Sciences of the United States of America 78, 3779-3781.

[17] O'Brien, T. E., 2005. *Designing for Parameter Subsets in Gaussian Nonlinear Regression Models*. Journal of Data Science 3, 179-197.

[18] Pukelsheim, F., 1993. *Optimal Design of Experiment*. In: Society for Industrial and Applied Mathematics. SIAM, Philadelphia, PA.

[19] Ting, N., 2006. *Dose Finding in Drug Development*. New York.

[20] Waterhouse, T. H., Eccleston, J. A., Woods, D. C., Lewis, S. M., 2008. *Design selection criteria for discrimination/estimation for nested models and a binomial response*. Journal of statistical planning and inference 138, 132-144.

[21] Welshons, W. V., Thayer, K. A., Judy, B. M., Taylor, J. A., Curran, E. M., 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] White, L. V., 1973. *An Extension of the General Equivalence Theorem to Nonlinear Models*. Biometrika 60, 345-348.

APPENDIX A. R-CODE FOR D_S-OPTIMAL DESIGN UNDER MODEL WITH $\Theta = \Theta_1$

Note: For other sets of parameters, only the values of parameters in the code are changed.

```
#number of parameter
k=3
#value of parameter
alpha=4.63
beta=1.23
gamma=0.07
#Initial value
x0=c(-14,-10,-8,-4)
n0=length(x0)
w=rep(1/n0,n0)
D=rbind(x0,w)
#Initial Information matrix
A1<-rep(0,n0)
A2<-rep(0,n0)
A3<-rep(0,n0)
A6<-rep(0,n0)
A9<-rep(0,n0)
for (i in 1:n0)
{
A1[i]=w[i]*exp(-(alpha+(beta*x0[i])+(gamma*x0[i]^2))^2)
A2[i]=x0[i]*A1[i]
A3[i]=x0[i]^2*A1[i]
```

```
A6[i]=x0[i]^3*A1[i]
A9[i]=x0[i]^4*A1[i]
}
M0=matrix(c(sum(A1),sum(A2),sum(A3),sum(A2),sum(A3),sum(A6),
sum(A3),sum(A6),sum(A9)),nrow=3,ncol=3,byrow=F)
IM0=solve(M0)
M1=matrix(c(sum(A1),sum(A2),sum(A2),sum(A3)),nrow=2,ncol=2,byrow=F)
IM1=solve(M1)
#Find ds
f0<-function(x)
{matrix(c(exp(-0.5*(alpha+beta*x+gamma*x^2)^2),x*exp(-0.5*(alpha+
beta*x+gamma*x^2)^2),x^2*exp(-0.5*(alpha+beta*x+gamma*x^2)^2)),
nrow=3,ncol=1,byrow=F)}
f1<-function(x)
{matrix(c(exp(-0.5*(alpha+beta*x+gamma*x^2)^2),x*exp(-0.5*(alpha+
beta*x+gamma*x^2)^2)),nrow=2,ncol=1,byrow=F)}
p=1
while(p>.001){
x1=seq(-14,-4,.01)
n1=length(x1)
ds=rep(0,n1)
for (j in 1:n1)
{ds[j]=t(f0(x1[j]))%*%solve(M0)%*%f0(x1[j])
-t(f1(x1[j]))%*%solve(M1)%*%f1(x1[j])}
for (j in 1:n1)
{if(max(ds)==ds[j])x1[j]=x1[j] else x1[j]=NA}
```

```
newX=na.omit(x1)
newds=max(ds)
#Find alpha(n+1)
an=1/(n1+1)
p<-newds-1
#Get M(n+1)
newM0=(1-an)*M0+an*f0(newX)%*%t(f0(newX))
newM1=(1-an)*M1+an*f1(newX)%*%t(f1(newX))
M0<-newM0
M1<-newM1
newW=(1-an)*D[2,]
W=c(newW,an)
X=c(D[1,],newX)
newD=rbind(X,W)
D=newD
dsoptimal<-by(D[2,], D[1,],FUN=sum)}
dsoptimal
#Verify Ds-optimal design
#number of parameter
k=3
#value of parameter
alpha=4.63
beta=1.23
gamma=0.07
#Ds-optimal design
x=c(-13.84,-8.84,-4)
```

```
n=length(x)
w=c(0.285,0.467,0.248)
Ds=rbind(x,w)
#information matrix for Ds-optimal design
A1<-rep(0,n)
A2<-rep(0,n)
A3<-rep(0,n)
A6<-rep(0,n)
A9<-rep(0,n)
for (i in 1:n)
{
A1[i]=w[i]*exp(-(alpha+(beta*x[i])+(gamma*x[i]^2))^2)
A2[i]=x[i]*A1[i]
A3[i]=x[i]^2*A1[i]
A6[i]=x[i]^3*A1[i]
A9[i]=x[i]^4*A1[i]
}
M=matrix(c(sum(A1),sum(A2),sum(A3),sum(A2),sum(A3),sum(A6),sum(A3),
sum(A6),sum(A9)),nrow=3,ncol=3,byrow=F)
IM=solve(M)
M2=matrix(c(sum(A1),sum(A2),sum(A2),sum(A3)),nrow=2,ncol=2,byrow=F)
IM2=solve(M2)
#Find ds
f0<-function(x)
{matrix(c(exp(-0.5*(alpha+beta*x+gamma*x^2)^2),x*exp(-0.5*(alpha+
beta*x+gamma*x^2)^2),x^2*exp(-0.5*(alpha+beta*x+gamma*x^2)^2)),
```

```
nrow=3,ncol=1,byrow=F)}
f1<-function(x)
{matrix(c(exp(-0.5*(alpha+beta*x+gamma*x^2)^2),x*exp(-0.5*(alpha+
beta*x+gamma*x^2)^2)),nrow=2,ncol=1,byrow=F)}
p=1
x1=seq(-14,-4,.01)
n1=length(x1)
ds=rep(0,n1)
for (j in 1:n1)
{ds[j]=t(f0(x1[j]))%*%solve(M)%*%f0(x1[j])
-t(f1(x1[j]))%*%solve(M2)%*%f1(x1[j])}
plot(x1,ds, type="l", xlab=quote(x [i]), ylab="Standardized Variance" )
ds=rep(1, n1)
lines(x1, ds, type="l")
#D-optimal design
x1=c(-13.22,-10.34,-7.23,-4.35)
n=length(x1)
w1=c(0.323,0.177,0.177,0.323)
D=rbind(x1,w1)
#information matrix for D-optimal design
A1<-rep(0,n)
A2<-rep(0,n)
A3<-rep(0,n)
A6<-rep(0,n)
A9<-rep(0,n)
for (i in 1:n)
```

```
{
A1[i]=w1[i]*exp(-(alpha+(beta*x1[i])+(gamma*x1[i]^2))^2)
A2[i]=x1[i]*A1[i]
A3[i]=x1[i]^2*A1[i]
A6[i]=x1[i]^3*A1[i]
A9[i]=x1[i]^4*A1[i]
}
M=matrix(c(sum(A1),sum(A2),sum(A3),sum(A2),sum(A3),
sum(A6),sum(A3),sum(A6),sum(A9)),nrow=3,ncol=3,byrow=F)
IMd=solve(M)
#Uniform design
x2=c(-14.00,-13.00,-12.00,-11.00,-10.00,-9.00,-8.00,
-7.00,-6.00,-5.00,-4.00)
n=length(x2)
w2=rep(1/11,11)
U=rbind(x2,w2)
#information matrix for Ds-optimal design
A1<-rep(0,n)
A2<-rep(0,n)
A3<-rep(0,n)
A6<-rep(0,n)
A9<-rep(0,n)
for (i in 1:n)
{
A1[i]=w2[i]*exp(-(alpha+(beta*x2[i])+(gamma*x2[i]^2))^2)
A2[i]=x2[i]*A1[i]
```

```
A3[i]=x2[i]^2*A1[i]
A6[i]=x2[i]^3*A1[i]
A9[i]=x2[i]^4*A1[i]
}
Mu=matrix(c(sum(A1),sum(A2),sum(A3),sum(A2),sum(A3),
sum(A6),sum(A3),sum(A6),sum(A9)),nrow=3,ncol=3,byrow=F)
IMu=solve(Mu)
A=c(0,0,1)
var_ds=t(A)%*%IM%*%A
var_d=t(A)%*%IMd%*%A
var_u=t(A)%*%IMu%*%A
effds_d=var_ds/var_d
effds_d
effds_u=var_ds/var_u
effds_u
effd_u=var_d/var_u
effd_u
```

APPENDIX B. R-CODE FOR D-OPTIMAL DESIGN UNDER MODEL WITH $\Theta = \Theta_1$

Note: For other sets of parameters, only the values of parameters in the code are changed.

```
#number of parameter
k=3
#value of parameter
alpha=4.63
beta=1.23
gamma=0.07
#Initial value
x0=c(-14,-10,-8,-4)
n0=length(x0)
w=rep(1/n0,n0)
D=rbind(x0,w)
#Initial Information matrix
A1<-rep(0,n0)
A2<-rep(0,n0)
A3<-rep(0,n0)
A6<-rep(0,n0)
A9<-rep(0,n0)
for (i in 1:n0)
{
A1[i]=w[i]*exp(-(alpha+(beta*x0[i])+(gamma*x0[i]^2))^2)
A2[i]=x0[i]*A1[i]
A3[i]=x0[i]^2*A1[i]
```

```
A6[i]=x0[i]^3*A1[i]
A9[i]=x0[i]^4*A1[i]
}
M0=matrix(c(sum(A1),sum(A2),sum(A3),sum(A2),sum(A3),sum(A6),sum(A3),
sum(A6),sum(A9)),nrow=3,ncol=3,byrow=F)
IM0=solve(M0)
#Find d
f<-function(x)
{matrix(c(exp(-0.5*(alpha+beta*x+gamma*x^2)^2),x*exp(-0.5*(alpha+beta*x
+gamma*x^2)^2),x^2*exp(-0.5*(alpha+beta*x+gamma*x^2)^2)),nrow=3,ncol=1,
byrow=F)}
p=1
while(p>.001){
x1=seq(-14,-4,.01)
n1=length(x1)
dn=rep(0,n1)
for (j in 1:n1)
{dn[j]=t(f(x1[j]))%*%solve(M0)%*%f(x1[j])}
for (j in 1:n1)
{if(max(dn)==dn[j])x1[j]=x1[j] else x1[j]=NA}
newX=na.omit(x1)
newdn=max(dn)
#Find alpha(n+1)
an=(newdn-k)/(k*(newdn-1))
p<-newdn-k
#Get M(n+1)
```

```
newM=(1-an)*M0+an*f(newX)%*%t(f(newX))
M0<-newM
newW=(1-an)*D[2,]
W=c(newW,an)
X=c(D[1,],newX)
newD=rbind(X,W)
D=newD
doptimal<-by(D[2,], D[1,],FUN=sum)}
doptimal
#Verify D-optimal design
#number of parameter
k=3
#value of parameter
alpha=4.63
beta=1.23
gamma=0.07
#D-optimal design
x=c(-13.22,-10.34,-7.23,-4.35)
n=length(x)
w=c(0.323,0.177,0.177,0.323)
D=rbind(x,w)
#information matrix for D-optimal design
A1<-rep(0,n)
A2<-rep(0,n)
A3<-rep(0,n)
A6<-rep(0,n)
```

```
A9<-rep(0,n)
for (i in 1:n)
{
A1[i]=w[i]*exp(-(alpha+(beta*x[i])+(gamma*x[i]^2))^2)
A2[i]=x[i]*A1[i]
A3[i]=x[i]^2*A1[i]
A6[i]=x[i]^3*A1[i]
A9[i]=x[i]^4*A1[i]
}
M=matrix(c(sum(A1),sum(A2),sum(A3),sum(A2),sum(A3),sum(A6),
sum(A3),sum(A6),sum(A9)),nrow=3,ncol=3,byrow=F)
IM=solve(M)
#Find dn
f<-function(x)
{matrix(c(exp(-0.5*(alpha+beta*x+gamma*x^2)^2),x*exp(-0.5*
(alpha+beta*x+gamma*x^2)^2),x^2*exp(-0.5*(alpha+beta*x+
gamma*x^2)^2)),nrow=3,ncol=1,byrow=F)}
p=1
x1=seq(-14,-4,.05)
n1=length(x1)
dn=rep(0,n1)
for (j in 1:n1)
{dn[j]=t(f(x1[j]))%*%solve(M)%*%f(x1[j])}
plot(x1,dn,type="l",ylab="Standardized Variance", xlab=quote(x [i]))
dn=rep(3, n1)
lines(x1, dn, type="l")
```