OPTIMAL DESIGNS FOR DOSE-FINDING IN CONTINGENT RESPONSE MODELS

A Dissertation
presented to
the Faculty of the Graduate School
University of Missouri - Columbia

In Partial Fulfillment
of the Requirements for the Degree

Doctor of Philosophy

---

by
HUWAIDA RABIE

Dr. Nancy Flournoy, Dissertation Supervisor

DECEMBER 2004

The undersigned, appointed by the Dean of the Graduate School, have examined the dissertation entitled

OPTIMAL DESIGNS FOR DOSE-FINDING IN
CONTINGENT RESPONSE MODELS

presented by Huwaida Rabie

a candidate for the degree of Doctor of Philosophy

and hereby certify that in their opinion it is worthy of acceptance.

Nancy Flournoy

Shanna Swan

Youssef Saab

Zhuoqiong He

Farroll T.Wright

Paul Speckman

DEDICATION

To my family especially my mother who encouraged me to do the best always.

ACKNOWLEDGMENTS

I would like to express my gratitude and thanks to my advisor Dr. Nancy Flournoy for all the help, support, and the time she spent supervising my dissertation. My thanks are extended to my committee members for their suggestions, especially Dr. Farroll Wright and Paul Speckman for the useful comments they provided when reviewing my dissertation. My thanks to Dr. Leonard Hearne for the help in understanding the Fortran program that used in finding the optimal designs.

I would like to thank my parents Salim and Rahma for the love, passion, and encouragement during my study. Also my sincere appreciation to my sister Wafa and my brothers George and Khalid for taking care of my parents and providing full support to me.

Foremost, I thank God who makes everything possible and takes care of me and my family.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

OPTIMAL DESIGNS FOR DOSE-FINDING IN
CONTINGENT RESPONSE MODELS

Huwaida Rabie

Dr. Nancy Flournoy, Dissertation Supervisor

ABSTRACT

We study D and c optimal designs for dose–finding with opposing failure functions. In particular, we study the contingent response models of Li, Durham and Flournoy (1995). In the contingent response model, there are two opposing types of failure. We call one failure type *toxicity* and the other *disease failure*, short for *failure due to disease*. No disease failure is *efficacy*. No toxicity and no disease failure is a *success* or *cure*. We assume disease failures are contingent on toxicity in that they are only observed in the absence of toxicity. We also assume the probability of toxicity increases with the dose, and the probability of disease failure given no toxicity decreases with dose.

We find canonical c- and D-optimal designs and show that other designs in the location–scale family can be obtained from a canonical design. For c-optimality, interest is in finding designs for estimating the dose that maximizes the cure probability, which we call the *optimal dose*. We use the positive–negative extreme value contingent response model to provide a specific illustration of the D and C optimal designs. We examine the efficiency of relevant up–and–down procedures in the literature for estimating the optimal dose based on the maximum likelihood estimation.

We show that these procedures are inefficient for estimating the optimal dose.

Chapter 1

INTRODUCTION

An experimental design problem concerns the allocation of treatment(s) to experimental units (subjects) for the purpose of investigating the nature of dependency of a response variable(s) on these treatment(s). In a model-based statistical procedure, the distribution of the response variable(s) is specified in terms of unknown parameters and treatments. The goals of the experimental design could be to efficiently estimate the unknown parameters or some function of them, and or to efficiently test hypotheses about the parameters.

The theory of optimal experimental designs has been studied widely, but research has largely focused on regression settings, especially on linear models with normal errors. Many authors have contributed to this research including Elfving [10], Wynn [40], Fedorov [14], Atkinson and Donev [2], Pazman [33] and Silvey [36]. With linear models, Fishers information is independent of model parameters, and hence finding the optimal designs gives the researcher an explicit procedure for allocating treatments to subjects. In many applications, such as in a bio-assay or a phase II clinical trial, the response variable(s) is modelled by a nonlinear function for which Fisher's information depends on the parameters of interest, which are unknown. Optimal designs for nonlinear models have been studied for a univariate response by, for example, White [39], Silvey [36], Sitter [38], Atkinson and Donev [2], and He [42]. Optimal designs for linear models are well studied, but the same cannot be said for nonlinear models. Multivariate responses have received little attention and more research is needed in this field. Zocci and Atkinson [43] and Fan and Chaloner [11] have studied some multivariate responses. In particular, they studied optimal designs for the logistic continuation–ratio model.

In this dissertation, we study optimal designs for nonlinear response models. In particular, we study D and c optimal designs for the contingent response models of Li, Durham and Flournoy [29]. In the contingent response model, there are two types of failure. We call one failure type *toxicity* and the other *disease failure*. No disease failure is *efficacy*. No toxicity and no disease failure is a *success* or *cure*. We assume disease failures are *contingent* on toxicity in that they can only be observed in the absence of toxicity. We also assume the probability of toxicity increases with dose, and the probability of disease failure given no toxicity decreases with dose. See Figures 1.1 and 1.2.

Examples of data well fit by a contingent response model arise in many areas of study. When a new drug is to be tested two concerns arise: the safety and the

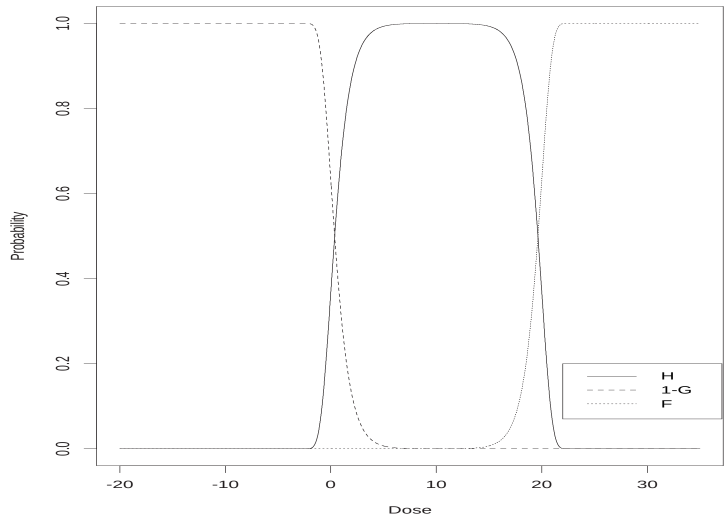

Figure 1.1: Positive-negative extreme value model for $r = 1, \mu = -20$.

efficacy of a drug, that is, how toxic is the drug and how well the drug produces the intended result. In many phase II clinical trials, a toxicity failure is fatal and so severe as to stop the trial for these subjects. Then efficacy results are obtained only in the absence of toxicity failures.

Finding an effective dose with little toxicity is important, and hence a good experimental design is important. For example, the sooner a drug's efficacy at a tolerate dose is established, the sooner a decision can be made whether or not to continue development of the drug; an efficient determination that a drug is ineffective or too toxic prevents wasted resources in the larger and more expensive phase III study; when an effective dose is found efficiently approval time is shortened.

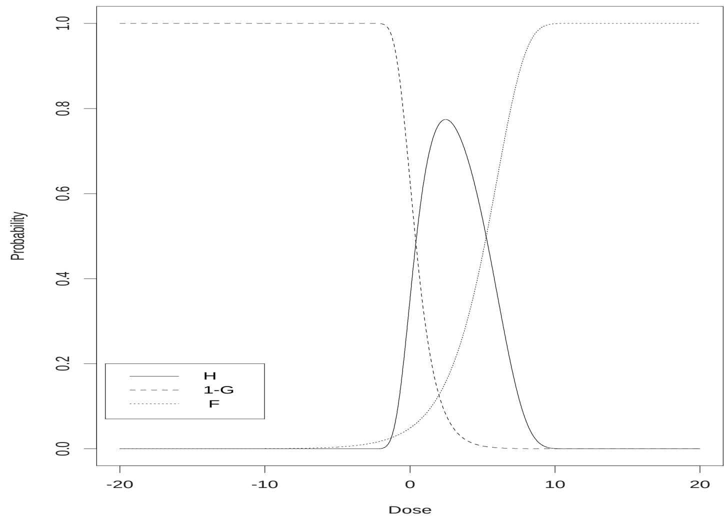

Figure 1.2: Positive-negative extreme value model for $r = 0.5, \mu = -3$.

The contingent response model is also useful in some stress testing situations. For example, Hayes, Edie and Durham [25] describe testing the compressive strength of fibers. A fiber may fail after it is stressed under tension to a predetermined level. Only if the fiber does not break under the initial tension is a recoil test initiated. If the initial stress level is sufficiently high (but not high enough to lead to a failure), the fiber may then fail due to compressive stresses generated as the stored strain energy is recovered. The goal is to find the stress level that maximizes the probability of a recoil success without tensile failure.

Fan and Chaloner [12] found D and c-optimal designs for the continuation ratio model which we show to be a special case of the contingent response model in the next chapter. Zocchi and Atkinson [43] study the influence of gamma radiation of the emergence of house flies. "In this study, seven sets of 500 pupae were exposed to one of several doses of radiation. Observations from each set of pupae after a period of time included the number of flies that died before the opening (unopened pupae), the number of flies that died before complete emergence, and the number of flies that completely emerged from the set of 500 pupae. They found the D-optimal designs for this application.

In Chapter 2, we give some background theory on optimal designs. In Chapter 3, we define the contingent response model and give examples. In Chapter 4, we present general theorems for D-optimal designs for the contingent response model and find the optimal designs for the positive–negative extreme value model. In Chapter 5, we find the c-optimal designs for the positive–negative extreme value model and present some relative theorems. In Chapter 6, we define the limiting optimal designs and find the limiting D-optimal designs for the positive–negative extreme value model. In Chapter 7, we introduce an up–and–down designs for implementing the optimal designs and we give some comparisons with other up–and–down designs that are often used.

Chapter 2

BACKGROUND THEORY ON OPTIMAL DESIGNS

When the distribution of an observable response(s) depends on control variable(s), decisions must be made concerning what levels of the control variable(s) to use and how to allocate treatments to experimental units. The decisions typically depend on how many experimental units are available and the range of the control(s) variable(s). Different criteria can be adopted based on the goals of the experiments.

In nonlinear models, Fisher's Information depends on the parameters of the underlying distribution of the response(s). This problem led Chernoff [5] to propose locally optimal designs which are produced by evaluating the optimal designs at guessed parameter values. Estimating parameters with data from previous experiments is an attractive alternative. The optimal design also can be approximated by sequentially updating the maximum likelihood estimates of the parameters, and reevaluating the optimal design after every subject, or group, is treated. Empirical Bayesian methods also can be used. Grovagnoli [20] proposed a nonparametric sequential approximation to optimal designs using up and down designs.

2.1 The Design Problem

Let $\boldsymbol{\theta}$ be the vector of the parameters in the distribution of the response(s) variable, and let $I(x,\boldsymbol{\theta})$ be the per observation Fisher's information matrix at control variable value x. If a design puts n_i independent observations at $x_i, i = 1, 2, ..., K$ for fixed $N = \Sigma_i n_i$, then Fisher's information matrix is $\Sigma_i n_i \mathbf{I}(x_i,\boldsymbol{\theta})$. The problem is to find $\{x_i\}$, $\{n_i\}$, and K to maximize an optimality criteria (ϕ). This is the exact design problem. Because of the discrete domain, calculus optimization techniques cannot be implemented. Kiefer [27] suggested an *approximate theory* in which $n_i/N = \xi_i$ with $0 \leq \xi_i \leq 1$. Fisher's information becomes N $\mathbf{M}$ $(\boldsymbol{\xi}, \boldsymbol{\theta})$, where $\mathbf{M}(\boldsymbol{\xi}, \boldsymbol{\theta}) = \Sigma_i \xi_i I(x_i, \boldsymbol{\theta})$ is the *average Fisher's information matrix* with the measure $\boldsymbol{\xi} = \begin{pmatrix} x_1 & \ldots & x_K \\ \xi_1 & \ldots & \xi_K \end{pmatrix}$.
The problem now becomes finding $\boldsymbol{\xi}$ and K to maximize the optimality criteria. Exact designs are found by integer approximation of continuous designs. In another words, for any ϕ we want to find $\boldsymbol{\xi}^*$ such that $\phi\{\mathbf{M}(\boldsymbol{\xi}, \boldsymbol{\theta})\} = max_{\xi \in H} \phi\{\mathbf{M}(\boldsymbol{\xi}, \boldsymbol{\theta}\}$, where H is the set of all possible designs.

The following are some important properties of Fisher's information matrix:

1) The Fisher's information matrix is symmetric and positive definite.

2) If the number of support points of the measure $\boldsymbol{\xi}$ is less than the number of unknown parameters, then Fisher's information is singular.

3) The set of Fisher's information matrices over all possible designs $\boldsymbol{\xi} \in H$ is the convex hull of the family of information matrices for a given $\boldsymbol{\theta}$.

4) The number of support points for any design is at most $p(p+1)/2$, where p is the number of parameters in the model, by the Caratheodory's theorem. For more details, see (Silvey [37], p.72).

2.2 Optimality Criteria

In this section we describe some commonly used optimality criteria (*cf.* Silvey[37] and Fedorov [14]). We give more details for the criteria used in our work, D and c optimality.

2.2.1 $D-$optimality

The most intensively studied criterion is D-optimality. It is well known that the asymptotic variance of the maximum likelihood estimates (MLE) can be approximated by the inverse of Fisher's information. The D-optimality criterion maximizes the determinant of Fisher's information, but it is usually expressed in terms of logarithm determinant of Fisher's information to assure the concavity of the criterion. This is instrumental for verifying the optimality of a candidate design. Define

$$\phi_D = max_{\xi}[log(|\mathbf{M}(\boldsymbol{\xi}, \boldsymbol{\theta})|)] = min[log(|(\mathbf{M})^{-1}(\boldsymbol{\xi}, \boldsymbol{\theta})|)].$$

If we are interested in estimating all the model's parameters, D-optimality criteria is used.

An advantage of this criterion is that the optimal designs do not depend on the scale of the variables. In other words, non-degenerate linear transformations in the space of the parameters estimates leave D-optimal designs unchanged, unlike A and E optimality criterion described below. Some more properties of the D-optimality criteria ϕ_D are given below:

1) ϕ_D is an increasing function of Fisher's information matrices. That is, if $\boldsymbol{\xi_1}, \boldsymbol{\xi_2}$ are two design measures then $\phi_D(\mathbf{M}(\boldsymbol{\xi_1}, \boldsymbol{\theta}) + \mathbf{M}(\boldsymbol{\xi_2}, \boldsymbol{\theta})) > \phi_D(\mathbf{M}(\boldsymbol{\xi_1}, \boldsymbol{\theta}))$.

2) ϕ_D is concave on the set of Fisher's information matrices defined on all $\xi \in H$.

3) The D-optimal design need not to be unique. If $\boldsymbol{\xi_1}$ and $\boldsymbol{\xi_2}$ are D-optimal designs, then the design $\boldsymbol{\xi}^* = \epsilon\ \boldsymbol{\xi_1}^* + (1-\epsilon)\ \boldsymbol{\xi_2}^*$, $0 \leq \epsilon \leq 1$, is D-optimal.

In linear models when errors are assumed normal, the confidence ellipsoid for θ is proportional to $\det \mathbf{M}^{-1/2}(x)$. The result is to make this ellipsoid as small as possible, i.e. maximizing $\det \mathbf{M}(x)$. In non-linear models, the confidence ellipsoid for θ is not proportional to $\det \mathbf{M}^{-1/2}(x, \boldsymbol{\theta})$. However, we still use $\det \mathbf{M}^{-1/2}(x, \boldsymbol{\theta})$ as a measure of variation.

2.2.2 D_A-optimality

This criterion is used when we are interested in s linear combinations of the parameters, that is, elements of $\mathbf{A}^T\boldsymbol{\theta}$, where $\mathbf{A}$ is a $p \times s$ matrix with $s < p$. The covariance matrix for these linear combinations is $\mathbf{A}^T(\mathbf{M}^{-1}(\boldsymbol{\xi}, \boldsymbol{\theta}))\mathbf{A}$. Hence the criteria minimizes $log|\mathbf{A}^T\mathbf{M}^{-1}(\boldsymbol{\xi}, \boldsymbol{\theta}))\mathbf{A}|$.

2.2.3 $A-$optimality

A optimality is used when we are interested in estimating the average of the parameters. A-optimality minimizes the sum of the asymptotic variances of the parameter estimates. That is, A-optimality minimizes the trace of $\mathbf{M}^{-1}(\boldsymbol{\xi}, \boldsymbol{\theta})$.

2.2.4 $E-$optimality

The E-optimality criterion minimizes the variance of the least well-estimated contrast $a^T\theta$, with $a^Ta = 1$. That is, minimize the maximum eigenvalue of $\mathbf{M}^{-1}(\boldsymbol{\xi}, \boldsymbol{\theta})$.

2.2.5 $c-$optimality

The c-optimality criterion, a special case of D_A-optimality, is used when we are interested in estimating a specific linear combination of $\boldsymbol{\theta}$, say $c^T\boldsymbol{\theta}$ where c is a $p \times 1$ vector, with minimum variance, that is proportional to $c^T\mathbf{M}^{-1}(\boldsymbol{\xi}, \boldsymbol{\theta})c$. The criterion is

$$\phi_c = \begin{cases} max(-c^T\mathbf{M}^{-1}(\boldsymbol{\xi}, \boldsymbol{\theta})c) & \text{if } \mathbf{M}(\boldsymbol{\xi}, \boldsymbol{\theta}) \text{ is nonsingular} \\ \\ -\infty & \text{else.} \end{cases}$$

A disadvantage of the c-optimality criterion is that the optimal designs are often singular. Hence ϕ_c is not continuous and the maximum might not exist. Silvey[37] modified the definition to

$$\phi_c = \begin{cases} max(-c^T \mathbf{M}^{-1}(\boldsymbol{\xi}, \boldsymbol{\theta})c) & \text{if } \mathbf{M}(\boldsymbol{\xi}, \boldsymbol{\theta}) \text{ is nonsingular} \\ \\ max(-c^T \ \mathbf{M}^{-}(\boldsymbol{\xi}, \boldsymbol{\theta})c) & \text{if } c^T\boldsymbol{\theta} \text{ is estimable when } \ \mathbf{M}(\boldsymbol{\xi}, \boldsymbol{\theta}) \text{ is singular} \\ \\ -\infty & \text{else,} \end{cases}$$

where $\mathbf{M}^{-}(\boldsymbol{\xi}, \boldsymbol{\theta})$ is a generalized inverse of $\mathbf{M}(\boldsymbol{\xi}, \boldsymbol{\theta})$. With Silvey's modification, the criterion becomes concave and continuous so the maximum exists and any local maximum is a global maximum.

2.2.6 Bayesian Optimality Criteria

Bayesian optimal criteria average the optimality criteria described above with respect to a prior distribution on $\boldsymbol{\theta}$.

2.3 Verifying the Optimality Criteria

The General Equivalence Theorem was developed by Kiefer [27] and used to verify the optimality criteria. We define ϕ and state the General Equivalence Theorem as they are stated in Silvey [37]. Before we restate Silvey's [37] General Equivalence Theorem 6.1.2, we define the directional derivative, which is used in this theorem.

Definition 2.3.1 *The directional derivative of ϕ at a matrix $\boldsymbol{M}_1$ in the direction of* *a matrix* $\boldsymbol{M}_2$ *is* $F_\phi(\boldsymbol{M}_1, \boldsymbol{M}_2) = \lim_{\epsilon\to 0^+} \frac{1}{\epsilon}[\phi\{(1-\epsilon)\boldsymbol{M}_1 + \epsilon\boldsymbol{M}_2\} - \phi\{\boldsymbol{M}_1\}]$.

Theorem 2.3.1 (The General Equivalence Theorem) *If* ϕ *is concave and differentiable at* $\boldsymbol{M}(\boldsymbol{\xi}, \boldsymbol{\theta})$, *then* $\boldsymbol{\xi}$ *is optimal if and only if* $F_\phi\{\boldsymbol{M}(\boldsymbol{\xi}, \boldsymbol{\theta}), \boldsymbol{I}(x, \boldsymbol{\theta})\} \leq 0$ *for all* x *in the domain of* x. *The supremum of* F_ϕ *over all possible values of* x *is 0 and it is attained at the optimal design points.*

When a candidate optimal design produces a singular $M(\boldsymbol{\xi}, \boldsymbol{\theta})$ another version of the general equivalence theorem was proved by Silvey [37]. Let r be the rank of the singular Fisher's information matrix and p is the dimension of $\boldsymbol{\theta}$. Note that $r < p$.

Theorem 2.3.2 (General Equivalence Thm. for Singular Optimal Designs) *If* ϕ *is concave and differentiable at* $\boldsymbol{M}(\boldsymbol{\xi}, \boldsymbol{\theta})$, *then a sufficient condition that* ϕ *is optimal at* $\boldsymbol{M}_r(\boldsymbol{\xi}^*, \boldsymbol{\theta})$ *is that there exist a matrix* $H_{p\times(p-r)}$ *of rank* $p - r$ *such that* $F_\phi\{[\boldsymbol{H}\boldsymbol{H}^T + \boldsymbol{M}_r(\boldsymbol{\xi}^*, \boldsymbol{\theta}), \boldsymbol{I}(x, \boldsymbol{\theta})\} \leq 0$ *over all possible of* x *in the domain of* x. *The supremum of* F_ϕ *over all possible of* x *is 0 and it is attained at the design points.*

Chapter 3

THE CONTINGENT RESPONSE MODEL

We now describe the contingent response model. Let

$$Y_{1j} = \begin{cases} 1 & \text{if the } j^{th} \text{ subject has a toxic response} \\ 0 & \text{else} \end{cases}$$

$$Y_{2j} = \begin{cases} 1 & \text{if the } j^{th} \text{ subject has disease failure} \\ 0 & \text{else} \end{cases}$$

for $j = 1, \ldots, N$. Only three outcomes are possible, namely, $\{Y_{1j} = 0, Y_{2j} = 0,\}, \{Y_{1j} = 0, Y_{2j} = 1\}$, and $\{Y_{1j} = 1\}$. We consider a location-scale family of parametric models: $P\{Y_{1j} = 1 \mid x\} = F(\alpha_1 + \beta_1 x) = F_x$; and $P\{Y_{2j} = 0 \mid Y_{1j} = 0, x\} = G(\alpha_2 + \beta_2 x) = G_x$, with $\bar{F}_x = 1 - F_x$ and $\bar{G}_x = 1 - G_x$; x is log dose. The probability of success is

$$H_x = P\{Y_{1j} = 0, Y_{2j} = 0 \mid x\} = \bar{F}_x G_x; \tag{3.1}$$

the probability of toxicity is F_x; and the probability of disease failure is $P(Y_{1j} = 0, Y_{2j} = 1) = \bar{F}_x \bar{G}_x$.

The optimal dose is the maximum of H_x. Assuming the derivatives of F and G exists, the optimal dose may be found by setting the derivative of (3.1) equal to zero, i.e.,

$$H'(\cdot) = \bar{F}(\cdot)G'(\cdot) - G(\cdot)F'(\cdot) = 0, \tag{3.2}$$

where F' and G' are the derivatives of F and G. See Figures 1 and 2.

When (3.2) does not have a closed form solution, numerical methods will yield a practical solution. Conditions for $H(\cdot)$ to have a maximum are given by Li, Durham and Flournoy [29]. These conditions are satisfied if $F(x)$ and $G(x)$ are probability continuous distribution functions.

The *continuation ratio model* is a special case of the contingent response model in which F_x and G_x are logistic. To see this, we follow Agresti ([1], p. 319) and define $\pi_1(x) = P(toxicity|x) = F_x$; $\pi_2(x) = P(cure|x) = \bar{F}_x G_x$; and $\pi_3(x) = P(disease failure|x) = \bar{F}_x \bar{G}_x$. Continuation ratio logits are defined as $L_j =$ logit $\rho_j(x)$, where $\rho_j(x) = \pi_j(x)/\sum_{i>j} \pi_i(x), i = 1, 2, 3; j = 1, 2$. Modeling $L_j = \alpha_j + \beta_j x$ is equivalent to assuming $F_x = \exp(\alpha_1 + \beta_1 x)/(1 + \exp(\alpha_1 + \beta_1 x))$ and $G_x = \exp(\alpha_2 + \beta_2 x)/(1 + \exp(\alpha_2 + \beta_2 x)), \beta_1, \beta_2 > 0$. Designs for continuation ratio model have been studied by Zocchi and Atkinson [43]. They found optimal design for a particular application of the continuation ratio model. Fan and Chaloner ([11], [12]) studied optimal designs for the continuation ratio model more generally. So this thesis is a generalization of some of their work. They give the solution to (3.2) when $\beta_1 = \beta_2$. For $\beta_1 \neq \beta_2$, (3.2) does not have an explicit solution. In this case Fan and Chaloner ([12]) used the implicit function theory introduced by Atkinson and Haines [3] to find the c-optimal designs. Melas [31] used the implicit function theory in finding the E-optimal designs for polynomial regression on a segment.

Li, Durham and Flournoy [29] describe two contingent response models for which the optimal dose does have an explicit solution:

1) **The positive-negative extreme value contingent response model (PNEV).**

Let $\bar{F}_x = \exp(-\exp(\alpha_1 + \beta_1 x))$ and $G_x = \exp(-\exp(-(\alpha_2 + \beta_2 x))), \beta_1,$ $\beta_2 > 0$. Then the optimal dose for $\beta_1 \neq \beta_2$ is $\nu = [log\{\beta_2/\beta_1\} - \alpha_1 - \alpha_2]/(\beta_1 + \beta_2)$. For $\beta_1 = \beta_2 = \beta$, the optimal dose is $\nu = -(\alpha_1 + \alpha_2)/2\beta$.

2) **The logistic-exponential contingent response model (LE).**

Let $F_x = \exp(\alpha_1 + \beta_1 x)/1 + \exp(\alpha_1 + \beta_1 x)$ and $G_x = \exp(\alpha_2 + \beta_2 x), \beta_1, \beta_2 > 0$, $x \in (-\infty, \frac{-\alpha_2}{\beta_2})$. For $\beta_1 > \beta_2$, the optimal dose is $\nu = [log\{\beta_2/(\beta_1 - \beta_2)\} - \alpha_1)]/\beta_1$. No solution to (3.2) exists for $\beta_1 \leq \beta_2$, but in that case the optimal dose is $x = -\alpha_2/\beta_2$.

3.1 Fisher's Information

Assume we have n_i independent observations taken at x_i, $i = 1, \ldots, K$. Let r_i be the number of toxic responses, m_i the number of cures and $(n_i - m_i - r_i)$ the number of disease failures at dose level x_i. Define $v(x) = F_x'^2/(F_x\bar{F}_x)$ and $w(x) = \bar{F}_x G_x'^2/(G_x\bar{G}_x)$. The likelihood function is proportional to

$$\prod_{i=1}^{K} F_{x_i}^{r_i}(\bar{F}_{x_i}G_{x_i})^{m_i}(\bar{F}_{x_i}\bar{G}_{x_i})^{n_i-r_i-m_i}.$$

Lemma 3.1.1 *If $\beta_1 \neq \beta_2$, then $\mathbf{\Theta} = (\alpha_1, \beta_1, \alpha_2, \beta_2)$ and $\mathbf{I_{4\times 4}}(x, \mathbf{\Theta}) = diag(\boldsymbol{A}, \boldsymbol{B})$, where*

$$\mathbf{A_{2\times 2}} = v(x, \mathbf{\Theta}) \begin{pmatrix} 1 & x \\ x & x^2 \end{pmatrix}; \mathbf{B_{2\times 2}} = w(x, \mathbf{\Theta}) \begin{pmatrix} 1 & x \\ x & x^2 \end{pmatrix}.$$

If $\beta_1 = \beta_2 = \beta$, then $\mathbf{\Theta} = (\alpha_1, \beta, \alpha_2)$ and

$$\boldsymbol{I}_{3\times 3}(x, \mathbf{\Theta}) = v(x, \mathbf{\Theta}) \begin{pmatrix} 1 & x & 0 \\ x & x^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} + w(x, \mathbf{\Theta}) \begin{pmatrix} 0 & 0 & 0 \\ 0 & x^2 & x \\ 0 & x & 1 \end{pmatrix}. \qquad (3.3)$$

Proof: The likelihood function for a single design point is proportional to

$$F_x^r(\bar{F}_x G_x)^m(\bar{F}_x\bar{G}_x)^{n-r-m} = F_x^r(G_x)^m(\bar{F}_x)^{n-r}(\bar{G}_x)^{n-m-r}.$$

So the loglikelihood function at x is proportional to

$$L = \{rlog(F_x) + mlog(G_x) + (n-r)log(\bar{F}_x) + (n-m-r)log(\bar{G}_x)\}.$$

For $\beta_1 \neq \beta_2$, define $\boldsymbol{\theta} = (\boldsymbol{\theta}_j) = ((\alpha_1, \beta_1, \alpha_2, \beta_2), j = 1, \ldots, 4,)$, $z_1 = \alpha_1 + \beta_1 x$ and $z_2 = \alpha_2 + \beta_2 x$. Then

$$\partial L/\partial\theta_j = \begin{cases} \frac{\partial z_1}{\partial \theta_j}\{r\frac{F'_{z_1}}{F_{z_1}} - (n-r)\frac{F'_{z_1}}{\bar{F}_{z_1}}\} & j = 1, 2 \\ \\ \frac{\partial z_2}{\partial \theta_j}\{m\frac{G'_{z_2}}{G_{z_2}} - (n-m-r)\frac{G'_{z_2}}{\bar{G}_{z_2}}\} & \text{j = 3, 4,} \end{cases} \tag{3.4}$$

where $\dfrac{\partial z_1}{\partial \theta_1} = 1$, $\dfrac{\partial z_1}{\partial \theta_2} = x$, $\dfrac{\partial z_2}{\partial \theta_3} = 1$ and $\dfrac{\partial z_2}{\partial \theta_4} = x$. For $j = l = 1, 2$,

$$\frac{\partial L^2}{\partial\theta_j \partial\theta_l} = \frac{\partial z_1}{\partial \theta_j}\frac{\partial z_1}{\partial \theta_l}\{r\frac{F_{z_1}F''_{z_1} - F'^2_{z_1}}{F^2_{z_1}} - (n-r)\frac{\bar{F}_{z_1}F''_{z_1} + F'^2_{z_1}}{\bar{F}^2_{z_1}}\}. \tag{3.5}$$

Now the a_{ij} element of the matrix $\mathbf{A}_{2\times 2}$ is

$$\begin{aligned} -E\frac{\partial L^2}{\partial\theta_j \partial\theta_l} &= \frac{\partial z_1}{\partial \theta_j}\frac{\partial z_1}{\partial \theta_l}\{-nF_{z_1}\frac{F_{z_1}F''_{z_1} - F'^2_{z_1}}{F^2_{z_1}} + n\bar{F}_{z_1}\frac{\bar{F}_{z_1}F''_{z_1} + F'^2_{z_1}}{\bar{F}^2_{z_1}}\} \\ &= (\frac{\partial z_1}{\partial \theta_j}\frac{\partial z_1}{\partial \theta_l})\{\frac{-n\bar{F}_{z_1}(F_{z_1}F''_{z_1} - F'^2_{z_1}) + nF_{z_1}(\bar{F}_{z_1}F''_{z_1} + F'^2_{z_1})}{F_{z_1}\bar{F}_{z_1}}\} \\ &= (\frac{\partial z_1}{\partial \theta_j}\frac{\partial z_1}{\partial \theta_l})\frac{nF'^2_{z_1}}{F_{z_1}\bar{F}_{z_1}}, \end{aligned} \tag{3.6}$$

and the upper block sub-matrix of Fisher's information matrix can be written as

$$\mathbf{A} = \begin{pmatrix} 1 & x \\ \\ x & x^2 \end{pmatrix} \frac{n\bar{F}'^2_{z_1}}{F_{z_1}\bar{F}_{z_1}}.$$

For $j, l = 3, 4$,

$$\frac{\partial L^2}{\partial\theta_j \partial\theta_l} = \frac{\partial z_2}{\partial \theta_j}\frac{\partial z_2}{\partial \theta_l}\{m\frac{G_{z_2}G''_{z_2} - G'^2_{z_2}}{G^2_{z_2}} - (n-m-r)\frac{\bar{G}_{z_2}G''_{z_2} + G'^2_{z_2}}{\bar{G}^2_{z_2}}\}. \tag{3.7}$$

The b_{ij} element of the matrix $\mathbf{B}_{2\times 2}$ is given by

$$\begin{aligned} -E\frac{\partial L^2}{\partial\theta_j \partial\theta_l} &= \frac{\partial z_2}{\partial \theta_j}\frac{\partial z_2}{\partial \theta_j}\{-n\bar{F}_{z_1}G_{z_2}\frac{G_{z_2}G''_{z_2} - G'^2_{z_2}}{G^2_{z_2}} + n\bar{F}_{z_1}\bar{G}_{z_2}\frac{\bar{G}_{z_2}G''_{z_2} + Gz'^2}{\bar{G}^2_{z_2}}\} \\ &= \frac{\partial z_2}{\partial \theta_j}\frac{\partial z_2}{\partial \theta_l}\{\frac{-n\bar{F}_{z1}\bar{G}_{z_2}(G_{z_2}G''_{z_2} - G'^2_{z_2}) + n\bar{F}_{z_1}G_{z_2}(G''_{z_2} + Gz'^2)}{G_{z_2}\bar{G}_{z_2}}\} \\ &= \frac{\partial z_2}{\partial \theta_j}\frac{\partial z_2}{\partial \theta_l}\frac{n\bar{F}_{z_1}{G_{z_2}}'^2}{G_{z_2}\bar{G}_{z_2}}, \end{aligned} \tag{3.8}$$

and the lower block sub-matrix of Fisher's information matrix can be written as

$$\mathbf{B} = \begin{pmatrix} 1 & x \\ x & x^2 \end{pmatrix} \frac{n\bar{F}_{z_1}\bar{G}'^2_{z_2}}{G_{z_2}\bar{G}_{z_2}}.$$

Note that $\partial L^2/(\partial\theta_l\partial\theta_j) = 0$ for $l = 1, 2,\ j = 3, 4$ and $l = 3, 4,\ j = 1, 2$. So the Fishers information matrix for a single subject at x is

$$I(x, \Theta) = \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{pmatrix}.$$

For $\beta_1 = \beta_2$, define $\boldsymbol{\theta = (\theta_j)}= (\alpha_1, \beta, \alpha_2)$, $z_1 = \alpha_1 + \beta x$ and $z_2 = \alpha_2 + \beta x$. For

$$\frac{\partial L}{\partial \beta} = \frac{\partial z_1}{\partial \theta_j}\{r\frac{F'_{z_1}}{F_{z_1}} - (n-r)\frac{F'_{z_1}}{F_{z_1}}\} + \frac{\partial z_2}{\partial \theta_j}\{m\frac{G'_{z_2}}{G_{z_2}} - (n-m-r)\frac{G'_{z_2}}{G_{z_2}}\},$$

$$-E\frac{\partial L^2}{\partial \beta^2} = \frac{\partial z_1}{\partial \theta_j}\frac{\partial z_1}{\partial \theta_l}\frac{nF_{z_1}{}'^2}{F_{z_1}\bar{F}_{z_1}} + \frac{\partial z_2}{\partial \theta_j}\frac{\partial z_2}{\partial \theta_l}\frac{n\bar{F}_{z_1}G_{z_2}{}'^2}{G_{z_2}\bar{G}_{z_2}},$$

with $j = l = 2$. For $j = l = 1, 2$, both $\partial L/\partial\theta_j$ and $-E(\partial L^2/(\partial\theta_j\partial\theta_l))$ are given by (3.4) and (3.6), respectively. For $j = l = 2, 3$, $\partial L/\partial\theta_j$ and $-E(\partial L^2/(\partial\theta_j\partial\theta_l))$ are given by (3.4) and (3.8), respectively. Note that $\partial L^2/(\partial\theta_j\partial\theta_l) = 0, (j, l) = (1, 3)$ and $(j, l) = (3, 1)$. So the Fishers's information matrix for a single subject is

$$\mathbf{I}_{3\times 3}(x, \Theta) = \frac{F'^2(x, \Theta_1)}{F(x, \Theta_1)\bar{F}(x, \Theta_1)}\begin{pmatrix} 1 & x & 0 \\ x & x^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \frac{\bar{F}(x, \Theta_1)G'^2(x, \Theta_2)}{G(x, \Theta_2)\bar{G}(x, \Theta_2)}\begin{pmatrix} 0 & 0 & 0 \\ 0 & x^2 & x \\ 0 & x & 1 \end{pmatrix}.$$

□

Tables 3.1 and 3.2 give the components of Fishers information for PNEV and LE for unequal slopes and equal slopes, respectively.

Component	Positive-Negative Extreme Value	Logistic-Exponential
A_{11}	$\exp(2\alpha_1 + 2\beta_1 x_i)\bar{F}_i/F_i$	$\bar{F}_i F_i$
A_{12}	$x_i \exp(2\alpha_1 + 2\beta_1 x_i)\bar{F}_i/F_i$	$x_i \bar{F}_i F_i$
A_{22}	$x_i^2 \exp(2\alpha_1 + 2\beta_1 x_i)\bar{F}_i/F_i$	$x_i^2 \bar{F}_i F_i$
B_{11}	$\exp(-2\alpha_2 - 2\beta_2 x_i)\bar{F}_i G_i/\bar{G}_i$	$\bar{F}_i G_i/\bar{G}_i$
B_{12}	$x_i \exp(-2\alpha_2 - 2\beta_2 x_i)\bar{F}_i G_i/\bar{G}_i$	$x_i \bar{F}_i G_i/\bar{G}_i$
B_{22}	$x_i^2 \exp(-2\alpha_2 - 2\beta_2 x_i)\bar{F}_i G_i/\bar{G}_i$	$x_i^2 \bar{F}_i G_i/\bar{G}_i$

Table 3.1: Fisher's information for a single subject at x_i(unequal slopes)

Component	Positive-Negative extreme Value	Logistic-Exponential
I_{11}	$\exp(2\alpha_1 + 2\beta x_i))\bar{F}_i/F_i$	$\bar{F}_i F_i$
I_{12}	$x_i \exp(2\alpha_1 + 2\beta x_i)\bar{F}_i/F_i$	$x_i \bar{F}_i F_i$
I_{22}	$x_i^2 \exp(2\alpha_1 + 2\beta x_i)\bar{F}_i/F_i +$ $x_i^2 \exp(-2\alpha_2 - 2\beta x_i)\bar{F}_i G_i/\bar{G}_i$	$x_i^2[\bar{F}_i F_i + \bar{F}_i G_i/\bar{G}_i]$
I_{23}	$x_i \exp(-2\alpha_2 - 2\beta x_i)\bar{F}_i G_i/\bar{G}_i$	$x_i \bar{F}_i G_i/\bar{G}_i$
I_{33}	$\exp(-2\alpha_2 - 2\beta x_i))\bar{F}_i G_i/\bar{G}_i$	$\bar{F}_i G_i/\bar{G}_i$

Table 3.2: Fisher's information for a single subject at x_i (equal slopes)

Chapter 4

LOCALLY D-OPTIMAL DESIGNS

The D-optimality criterion, used when we are interested in estimating the model parameters, is the maximum of the determinant of Fisher's Information. Since there is no closed form solution for the optimal designs we study, the NPSOL [19] algorithm was used to find designs that maximize the optimality criterion (see appendix C); then the General Equivalence theorem was used to verify global optimality (see Silvey [37]).

4.1 Unequal slopes $\beta_1 \neq \beta_2$

Consider the contingent response model with distribution functions $F_x = F(\alpha_1+\beta_1 x)$ and $G_x = G(\alpha_2+\beta_2 x)$. Reparameterize $\mathbf{\Theta} = (\alpha_1, \beta_1, \alpha_2, \beta_2)$ to $\boldsymbol{\theta}=(\alpha_2, \beta_2, \mu, r)$, where $\mu = \alpha_1 - r\alpha_2$ and $r = \beta_1/\beta_2$. Following the idea of Ford, Torsney, and Wu [18], we say we have the *canonical* (μ, r) *model* when $\boldsymbol{\theta}= (0, 1, \mu, r)$. Let $\boldsymbol{\xi}$ be the D-optimal design for the canonical (μ, r) model. Theorem 4.1.1 tells us that the optimal designs for all other models with parameters $(\alpha_2, \beta_2, \mu, r)$ can be generated by a linear transformation of the canonical optimal designs.

Theorem 4.1.1 *If* $\boldsymbol{\xi}_0^* = \begin{pmatrix} x_1^* & \dots & x_K^* \\ \xi_1^* & \dots & \xi_K^* \end{pmatrix}$ *is locally D-optimal for* $\boldsymbol{\theta_0}= (0, 1, \mu, r)$, *then* $\boldsymbol{\xi}^* = \begin{pmatrix} \frac{x_1^*-\alpha_2}{\beta_2} & \dots & \dots, \frac{x_K^*-\alpha_2}{\beta_2}) \\ \xi_1^* & \dots & \xi_K^* \end{pmatrix}$ *is locally D-optimal for* $\boldsymbol{\theta_0}= (\alpha_2, \beta_2, \mu, r)$.

Proof: Recall that $v(x) = F_x'^2/(F_x\bar{F}_x)$, $w(x) = \bar{F}_x G_x'^2/(G_x\bar{G}_x)$, $v_i = v(x_i, \boldsymbol{\theta_0}) = v(x, \boldsymbol{\theta}) \mid_{x=(x_i-\alpha_2)/\beta_2}$ and $w_i = w(x_i, \boldsymbol{\theta_0}) = w(x, \boldsymbol{\theta})|_{x=(x_i-\alpha_2)/\beta_2}$. It can be seen from Lemma 3.1.1 that $\det \mathbf{I}(x, \boldsymbol{\theta})) = \det(\mathbf{A}(x, \boldsymbol{\theta})))\times \det(\mathbf{B}(x, \boldsymbol{\theta})))$. This implies that $\det\ \mathbf{M}(\boldsymbol{\xi}, \boldsymbol{\theta}) =$

$\det \sum_i^K \xi_i \mathbf{I}(x_i, \boldsymbol{\theta}) = \det \sum_i^K \xi_i \mathbf{A}(x_i, \boldsymbol{\theta})\times \det \sum_i^K \xi_i \mathbf{B}(x_i, \boldsymbol{\theta})$.

$$\begin{aligned}
\det \mathbf{A}(\boldsymbol{\theta}, \xi) &= \det\left(\sum_i^k \xi_i v_i \begin{pmatrix} 1 & \frac{x_i-\alpha_2}{\beta_2} \\ \frac{x_i-\alpha_2}{\beta_2} & (\frac{x_i-\alpha_2}{\beta_2})^2 \end{pmatrix}\right) \\
&= \frac{1}{\beta_2^2} \det \begin{pmatrix} \sum_i^K \xi_i v_i & \sum_i^K \xi_i v_i (x_i - \alpha_2) \\ \sum_i^K \xi_i v_i (x_i - \alpha_2) & \sum_i^K \xi_i v_i (x_i - \alpha_2)^2 \end{pmatrix} \\
&= \frac{1}{\beta_2^2}\{(\sum_i^K \xi_i v_i)\{\sum_i^K \xi_i v_i (x_i^2 - 2\alpha_2 x_i + \alpha_2^2)\} - (\sum_i^K \xi_i v_i (x_i - \alpha_2))^2\} \\
&= \frac{1}{\beta_2^2}\{(\sum_i^K \xi_i v_i) \sum_i^K \xi_i v_i x_i^2 - 2\alpha_2 (\sum_i^K \xi_i v_i) \sum_i^K \xi_i v_i x_i +
\end{aligned}$$

$$
\begin{aligned}
&\alpha_2^2(\sum_i^K \xi_i v_i)^2 - \sum_i^K \sum_j^K \xi_i v_i \xi_j v_j (x_i - \alpha_2)(x_j - \alpha_2)\} \\
= & \frac{1}{\beta_2^2}\{(\sum_i^K \xi_i v_i)\sum_i^K \xi_i v_i x_i^2 - 2\alpha_2(\sum_i^K \xi_i v_i)\sum_i^K \xi_i v_i x_i \\
&+\alpha_2^2(\sum_i^K \xi_i v_i)^2 + \sum_i^K \sum_j^K (-\xi_i v_i \xi_j v_j x_i x_j \\
&+\alpha_2 \xi_i v_i \xi_j v_j x_j + \alpha_2 \xi_i v_i \xi_j v_j x_i - \alpha_2^2 \xi_i v_i \xi_j v_j)\} \\
= & \frac{1}{\beta_2^2}\left((\sum_i^K \xi_i v_i)\sum_i^K \xi_i v_i x_i^2 - \sum_i^K \sum_j^K \xi_i v_i \xi_j v_j x_i x_j\right) \\
= & \frac{1}{\beta_2^2}\left((\sum_i^K \xi_i v_i)\sum_i^K \xi_i v_i x_i^2 - (\sum_i^K \xi_i v_i x_i)^2\right) \\
= & \frac{1}{\beta_2^2}\det\begin{pmatrix} \sum_i^K \xi_i v_i & \sum_i^K \xi_i v_i x_i \\ \sum_i^K \xi_i v_i x_i & \sum_i^K \xi_i v_i x_i^2 \end{pmatrix} \\
= & \frac{1}{\beta_2^2} det \sum_i^K \xi_i v_i \begin{pmatrix} 1 & x_i \\ x_i & x_i^2 \end{pmatrix} = \frac{1}{\beta_2^2}\det \mathbf{A}(\boldsymbol{\xi_0}, \boldsymbol{\theta_0}).
\end{aligned}
$$

In a similar way it can be shown that $\det \mathbf{B}\ (\boldsymbol{\xi}, \boldsymbol{\theta}) = 1/\beta_2^2\ \det \mathbf{B}(\boldsymbol{\xi_0}, \boldsymbol{\theta_0})$. So

$$
\begin{aligned}
\det \mathbf{M}(\boldsymbol{\xi}, \boldsymbol{\theta}) &= \det \mathbf{A}(\boldsymbol{\xi}, \boldsymbol{\theta}) \det \mathbf{B}(\boldsymbol{\xi}, \boldsymbol{\theta}) \\
&= \frac{1}{\beta_2^2}\ \det \mathbf{A}(\boldsymbol{\xi_0}, \boldsymbol{\theta_0})\frac{1}{\beta_2^2}\det \mathbf{B}(\boldsymbol{\xi_0}, \boldsymbol{\theta_0}) \\
&= \frac{1}{\beta_2^4}\mathbf{M}(\boldsymbol{\xi_0}, \boldsymbol{\theta_0}).
\end{aligned}
\tag{4.1}
$$

Now consider the one-to-one onto transformation $\tau(\boldsymbol{\xi_0}) = \boldsymbol{\xi}$ defined on the designs: $\boldsymbol{\xi}_0^* = \begin{pmatrix} x_1^* & \dots & x_K^* \\ \xi_1^* & \dots & \xi_K^* \end{pmatrix}$; $\boldsymbol{\xi}^* = \begin{pmatrix} \frac{x_1^* - \alpha_2}{\beta_2} & \dots & \dots, \frac{x_K^* - \alpha_2}{\beta_2}) \\ \xi_1^* & \dots & \xi_K^* \end{pmatrix}$. Then τ^{-1} is well defined.

Now

$$
\begin{aligned}
\boldsymbol{\xi}^* &= argmax_{\xi}\ \det(\mathbf{M}(\boldsymbol{\xi},\boldsymbol{\theta})) && definition\ of\ \boldsymbol{\xi}^* \\
&= argmax_{\xi}\ \beta_2^4 \det(\mathbf{M}(\boldsymbol{\xi},\boldsymbol{\theta}) && \beta_2^4\ is\ constant \\
&= argmax_{\xi}\ \det(\mathbf{M}(\boldsymbol{\xi_0},\boldsymbol{\theta_0})) && by\ (4.1) \\
&= argmax_{\tau(\xi_0)}\ \det(\mathbf{M}(\boldsymbol{\xi_0},\boldsymbol{\theta_0})) && \tau(\boldsymbol{\xi_0}) = \boldsymbol{\xi} \\
&= argmax_{\tau^{-1}\tau(\xi_0)}\ \det(\mathbf{M}(\boldsymbol{\xi_0},\boldsymbol{\theta_0})) && \tau\, is 1-1\, onto
\end{aligned}
$$

$$\Rightarrow$$

$$
\begin{aligned}
\boldsymbol{\xi_0}^* &= argmax_{\xi_0}\ \det(\mathbf{M}(\boldsymbol{\xi_0},\boldsymbol{\theta_0})) && definition\ of\ \boldsymbol{\xi_0}^* \\
& && and\ \tau^{-1}(\boldsymbol{\xi}) = \boldsymbol{\xi_0}.
\end{aligned}
$$

Maximizing $\det \mathbf{M}(\boldsymbol{\xi},\boldsymbol{\theta})$ is the same as maximizing $log \det \mathbf{M}(\boldsymbol{\xi},\boldsymbol{\theta})$ over all possible designs $\boldsymbol{\xi}$. Therefore, if $\boldsymbol{\xi_0}^*$ is locally D-optimal for $\boldsymbol{\theta_0}$, then $\boldsymbol{\xi}$ is locally D-optimal for $\boldsymbol{\theta}$. □

Locally D-optimal designs for several canonical (μ, r) positive-negative extreme value models are given in Tables 4.1–4.5. For each model, we give the optimal dose, the optimal design points and their weights, and the probabilities of toxicity, disease failure, and success. All optimal designs were verified by the General Equivalence Theorem given by Theorem 2.3.1. Figures 4.1 - 4.11. show that these designs' directional derivatives are nonpositive and attain their maxima at the design points (see Appendix B.1 for directional derivative code.) It can be seen that the optimal designs consist of two, three, four points for small, moderate and large negative values of μ, respectively. For positive values of μ, the optimal designs consist of two and three design points depending on the value of r. For the many combinations of r, μ values we studied, the number of optimal design points for the positive-negative extreme value model is the same as the number of optimal designs points for the analogous logistic-logistic model found by Fan and Chaloner [12].

The D-optimal design for a single negative extreme value model $F(\alpha_1 + \beta_1 x)$, with $\alpha_1 = 0, \beta_1 = 1$, was found by Ford, Torsney, and Wu [18] to be $\boldsymbol{\xi}^*_{NE} = \begin{pmatrix} -1.338 & 0.980 \\ 0.5 & 0.5 \end{pmatrix}$. This is the optimal design for $\bar{F}_{\alpha_1+\beta_1 x}$ also. The optimal design for the single positive extreme value model $G(\alpha_2 + \beta_2 x)$ with $\alpha_2 = 0, \beta_2 = 1$ is found using NPSOL [19] to be $\boldsymbol{\xi}^*_{PE} = \begin{pmatrix} 1.3377 & -0.9796 \\ 0.5000 & 0.5000 \end{pmatrix}$ which reflects the points of $\boldsymbol{\xi}^*_{NE}$ around the origin. This is the same optimal design for $\bar{G}_x$. For the positive-negative model, one sees from Table 4.1 and Figure 1.1 that F_x and $\bar{G}_x$ become quite separate when $-\mu$ gets larger, the optimal designs tend toward having four equally weighted points which are the optimal design points for the separate $F_{\mu+rx}$ and $\bar{G}_x$ combined: $\boldsymbol{\xi}^*_{PNE} = \begin{pmatrix} -1.338 & -0.980 & 0.980 & 1.338 \\ 0.25 & 0.25 & 0.25 & 0.25 \end{pmatrix}$. This is analogous to findings by Fan and Chaloner [12] for logistic-logistic model. As $\mu \to -\infty$, we show in Section 6.1 that the limiting optimal design for the canonical (μ, r) positive-negative extreme value model does indeed have four equally weighted points of the form $[x_1, x_2, (-x_1-\mu)/r, (-x_2-\mu)/r]$, with (x_1, x_2) equal to the optimal design points (x^*_1, x^*_2) of $\boldsymbol{\xi}^*_{PE}$.

r	μ	$g(\boldsymbol{\Theta})$	x	Weights	$F(x)$	$\bar{F}(x)\bar{G}(x)$	$H(x)$
0.5	0	0.46	-1.2752	0.4720	0.41	0.57	0.02
			0.5985	0.3382	0.74	0.11	0.15
			1.9480	0.1898	0.93	0.01	0.06
	-1	1.13	-1.0982	0.4005	0.19	0.77	0.04
			0.8243	0.3635	0.43	0.20	0.37
			3.9569	0.2360	0.93	0.00	0.07
	-3	2.46	-0.9329	0.3312	0.03	.89	0.08
			1.4913	0.4200	0.10	0.18	0.72
			7.6891	0.2488	0.90	0.00	0.10
	-5	3.80	-0.9101	0.2931	0.00	0.91	0.08
			1.6895	0.3729	0.02	0.17	0.82
			7.6511	0.0948	0.27	0.00	0.73
			11.6989	0.2392	0.90	0.00	0.10
	-10	7.13	-0.9778	0.2509	0.00	0.93	0.07
			1.3465	0.2524	0.00	0.23	0.77
			17.3336	0.2473	0.23	0.00	0.77
			21.955	0.2493	0.93	0.00	0.07
	-15	10.46	-0.9796	0.2500	0.00	0.93	0.07
			1.3379	0.2500	0.00	0.23	0.77
			27.3247	0.2500	0.23	0.00	0.77
			31.9592	0.2500	0.93	0.00	0.07

Table 4.1: D-optimal designs for the positive-negative extreme value model $r = 0.5$.

r	μ	$g(\mathbf{\Theta})$	x	Weights	$F(x)$	$\bar{F}(x)\bar{G}(x)$	$H(x)$
1	3	1.50	-4.2496	0.2350	0.25	0.75	0.00
			-2.0621	0.4262	0.92	0.08	0.00
			-1.4308	0.3388	0.99	0.01	0.00
	1	-0.50	-1.566	0.5000	0.43	0.56	0.01
			-0.2085	0.5000	0.89	0.08	0.03
	0	0.00	0.4755	0.5000	0.80	0.09	0.11
			-1.2808	0.5000	0.24	0.74	0.02
	-1	0.50	-1.1222	0.41	0.11	0.85	0.04
			0.4647	0.3447	0.44	0.26	0.30
			1.8528	0.2466	0.90	0.01	0.08
	-3	1.50	-0.9414	0.3092	0.02	0.91	0.08
			1.2863	0.4393	0.17	0.20	0.63
			3.8610	0.2515	0.91	0.00	0.09
	-5	2.50	-0.8454	0.2717	0.00	0.90	0.10
			2.2797	0.477	0.06	0.09	0.85
			5.8125	0.2513	0.90	0.00	0.11
	-10	5.00	-0.973	0.25	0.00	0.93	0.07
			1.362	0.2515	0.00	0.23	0.77
			8.6396	0.2793	0.22	0.00	0.77
			10.9725	0.2491	0.93	0.00	0.07
	-15	7.50	-.9795	0.2500	0.00	0.93	0.07
			1.3381	0.2500	0.00	0.23	0.77
			13.6618	0.2500	0.23	0.00	0.77
			15.9795	0.2500	0.93	0.00	0.07
	-20	10.00	-0.9796	0.2500	0.00	0.93	0.07
			1.3378	0.2500	0.00	0.23	0.77
			18.6623	0.2500	0.23	0.00	0.77
			20.9796	0.2500	0.93	0.00	0.07

Table 4.2: D-optimal designs for the positive-negative extreme value model $r = 1$.

r	μ	$g(\mathbf{\Theta})$	x	Weights	$F(x)$	$\bar{F}(x)\bar{G}(x)$	$H(x)$
2	3	-1.23	-1.8656	0.5000	0.38	0.62	0.00
			-1.0637	0.5000	0.91	0.07	0.01
	1	-0.56	-1.3831	0.5000	0.16	0.83	0.02
			-0.2548	0.5000	0.81	0.14	0.05
	0	-0.23	-1.2158	0.4391	0.08	0.89	0.03
			-0.1719	0.2660	0.51	0.34	0.15
			0.3417	0.2945	0.86	0.07	0.07
	-1	0.10	-1.1287	0.3566	0.04	0.92	0.04
			0.0812	0.3725	0.35	0.39	0.26
			0.9055	0.2708	0.90	0.04	0.07
	-3	0.77	-1.0136	0.2810	0.01	0.91	0.06
			0.7675	0.4573	0.21	0.30	0.50
			1.9332	0.2618	0.91	0.00	0.08
	-5	1.435618	-0.9256	0.2584	0.00	0.92	0.08
			1.5614	0.4831	0.14	0.16	0.70
			2.9311	0.2585	0.91	0.01	0.09
	-7	2.10	-0.8520	0.2521	0.00	0.90	0.10
			2.4150	0.4913	0.11	0.08	0.82
			3.9240	0.2566	0.90	0.00	0.10
	-10	3.10	-0.8987	0.2418	0.00	0.91	0.09
			1.3106	0.1511	0.00	0.24	0.76
			4.0744	0.3544	0.15	0.01	0.84
			5.4483	0.2526	0.91	0.00	0.09
	-15	4.77	-0.9709	0.2476	0.00	0.93	0.07
			1.3222	0.2385	0.00	0.23	0.77
			6.7937	0.2635	0.22	0.00	0.78
			7.98	0.2505	0.93	0.00	0.07
	-20	6.44	-0.9784	0.2496	0.00	0.93	0.07
			1.3352	0.2485	0.00	0.23	0.77
			9.3257	0.2519	2.30	0.00	0.77
			10.4889	0.2501	0.93	0.00	0.07

Table 4.3: D-optimal designs for the positive-negative extreme value model for $r = 2$.

r	μ	$g(\mathbf{\Theta})$	x	Weights	$F(x)$	$\bar{F}(x)\bar{G}(x)$	$H(x)$
3	3	-1.03	-1.5878	0.5000	0.16	0.84	0.01
			-0.8112	0.5000	0.83	0.15	0.02
	1	-0.53	-1.2783	0.4145	0.06	0.92	0.03
			-0.5062	0.26675	0.45	0.45	0.11
			-0.1083	0.3188	0.86	0.09	0.05
	0	-0.28	-1.2187	0.3374	0.03	0.94	0.03
			-0.2938	0.3749	0.34	0.49	0.17
			0.2638	0.2878	0.89	0.06	0.05
	-1	-0.03	-1.1701	0.2956	0.01	0.95	0.04
			-0.0434	0.4275	0.28	0.47	0.26
			0.6136	0.2769	0.90	0.04	0.06
	-3	0.48	-1.081	0.2621	0.00	0.95	0.05
			0.5085	0.4702	0.21	0.36	0.44
			1.2903	0.2676	0.91	0.02	0.07
	-5	0.98	-0.9980	0.2530	0.00	0.93	0.07
			1.0965	0.4838	0.17	0.24	0.60
			1.9570	0.2632	0.91	0.01	0.08
	-10	2.23	-0.8393	0.25	0.00	0.90	0.10
			2.6464	0.4912	0.12	0.06	0.82
			3.618	0.2588	0.91	0.00	0.10
	-15	3.48	-0.9263	0.2409	0.00	0.92	0.08
			1.2348	0.1719	0.00	0.25	0.75
			4.4443	0.3336	017	0.01	0.82
			5.3069	0.2536	0.921	0.00	0.08
	-20	4.72	-0.9663	0.2464	0.00	0.93	0.07
			1.3088	0.2310	0.00	0.24	0.76
			6.1891	0.2716	0.21	0.00	0.79
			6.9873	0.2510	0.93	0.00	0.07

Table 4.4: D-optimal designs for the positive-negative extreme value model for r=3.

r	μ	$g(\mathbf{\Theta})$	x	Weights	$F(x)$	$\bar{F}(x)\bar{G}(x)$	$H(x)$
4	3	-0.88	-1.3721	0.5000	0.08	0.90	0.02
			-0.6477	0.5000	0.78	0.19	0.04
	1	-0.48	-1.2976	0.308	0.02	0.96	0.03
			-0.4986	0.3971	0.31	0.56	0.13
			-0.0501	0.2949	0.89	0.07	0.04
	0	-0.28	-1.2558	0.2796	0.01	0.96	0.03
			-0.2914	0.4365	0.27	0.54	0.19
			0.2099	0.2839	0.90	0.06	0.04
	-1	-0.08	-1.2133	0.2649	0.00	0.96	0.04
			-0.07508	0.4578	0.24	0.50	0.26
			0.4649	0.2773	0.91	0.04	0.05
	-3	0.32	-1.1278	0.2534	0.00	0.95	0.05
			0.3732	0.4768	0.20	0.40	0.40
			0.9683	0.2698	0.91	0.03	0.06
	-5	0.72	-1.0481	0.2505	0.00	0.94	0.06
			0.8347	0.4838	0.17	0.29	0.54
			1.4682	0.2657	0.91	0.02	0.07
	-7	1.12	-0.9786	0.2500	0.00	0.93	0.07
			1.3053	0.4869	0.16	0.20	0.64
			1.9671	0.2632	0.91	0.01	0.08
	-10	1.72	-0.8149	0.2499	0.00	0.90	0.11
			2.02326	0.489	0.14	0.11	0.76
			2.7155	0.2611	0.91	0.01	0.09
	-20	3.72	-0.9337	0.2413	0.00	0.92	0.08
			1.2397	0.1814	0.00	0.25	0.75
			4.6004	0.3235	0.18	0.01	0.81
			5.2322	0.2538	0.92	0.00	0.08
	-30	5.72	-0.9732	0.2481	0.00	0.93	0.07
			1.3236	0.2411	0.00	0.23	0.77
			7.1552	0.2603	0.00	0.58	0.43
			7.7428	0.2506	0.93	0.00	0.07

Table 4.5: D-optimal designs for the positive-negative extreme value model for $r = 4$.

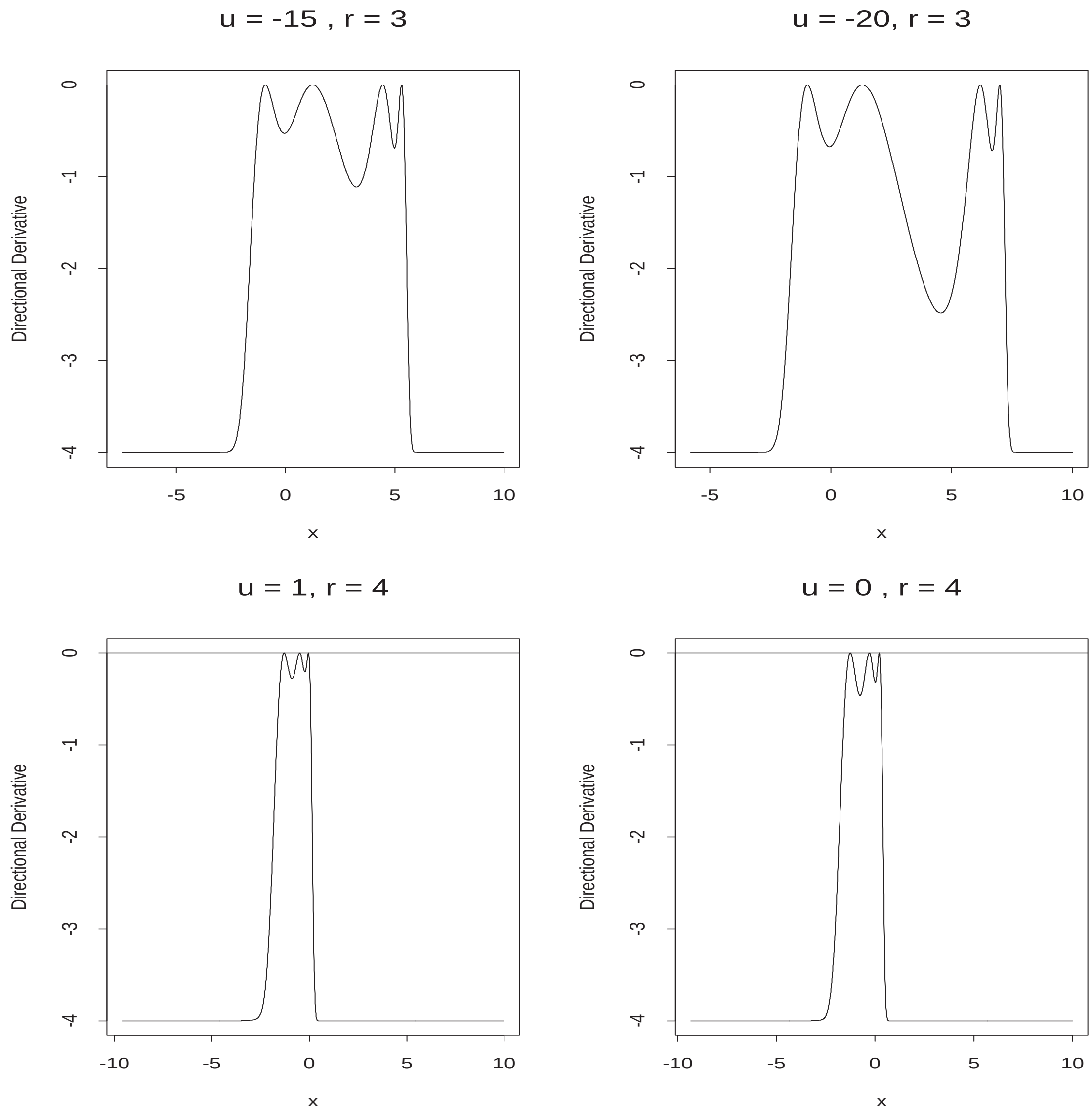

Figure 4.1: Directional derivatives of D-optimal designs for the positive-negative extreme value model with different values of μ and $r = 3, 4$.

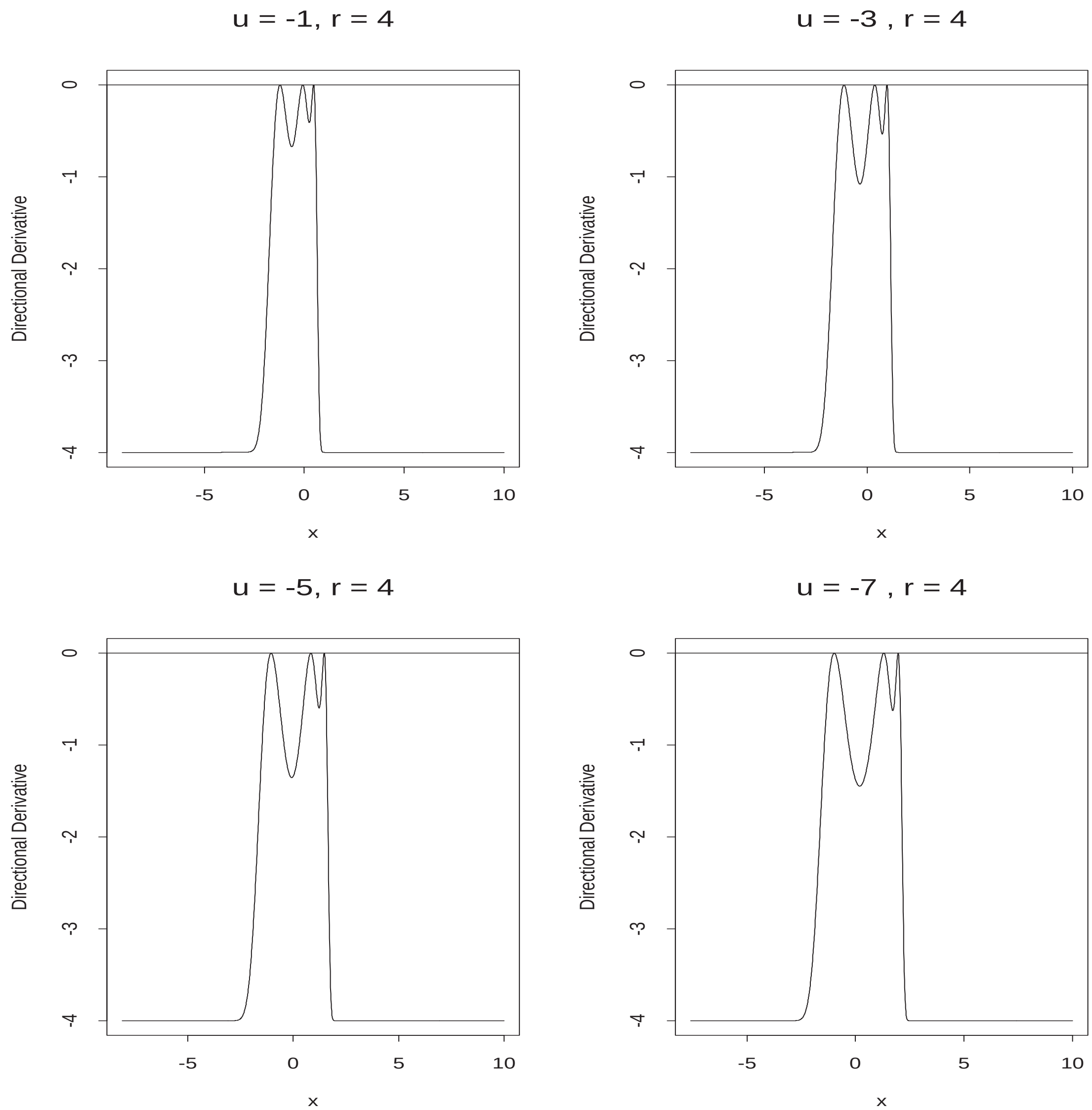

Figure 4.2: Directional derivatives of D-optimal designs for the positive-negative extreme value model for different values of μ and $r = 4$.

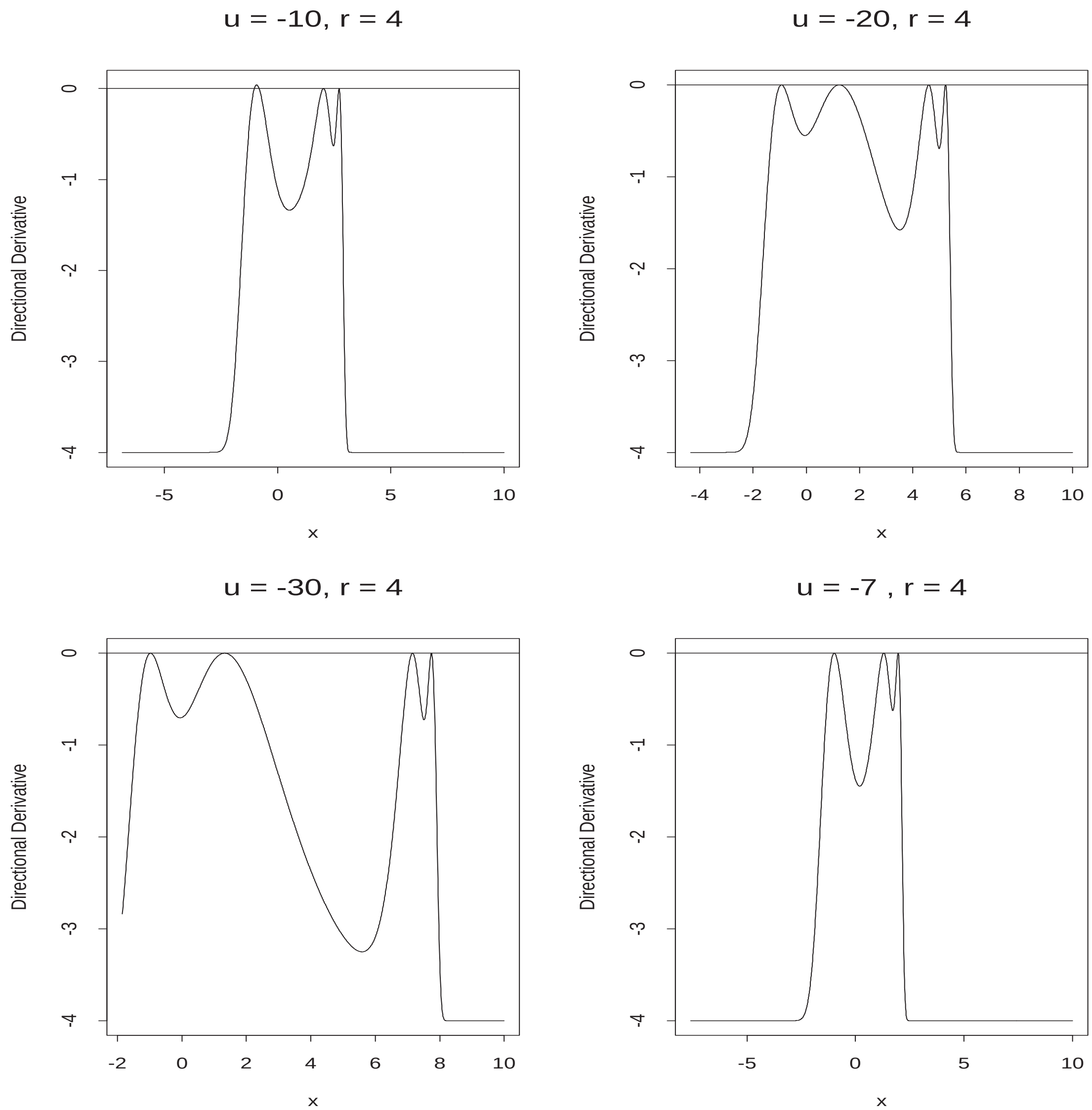

Figure 4.3: Continued directional derivatives of D-optimal designs for the positive-negative extreme value model for different values of μ and $r = 4$.

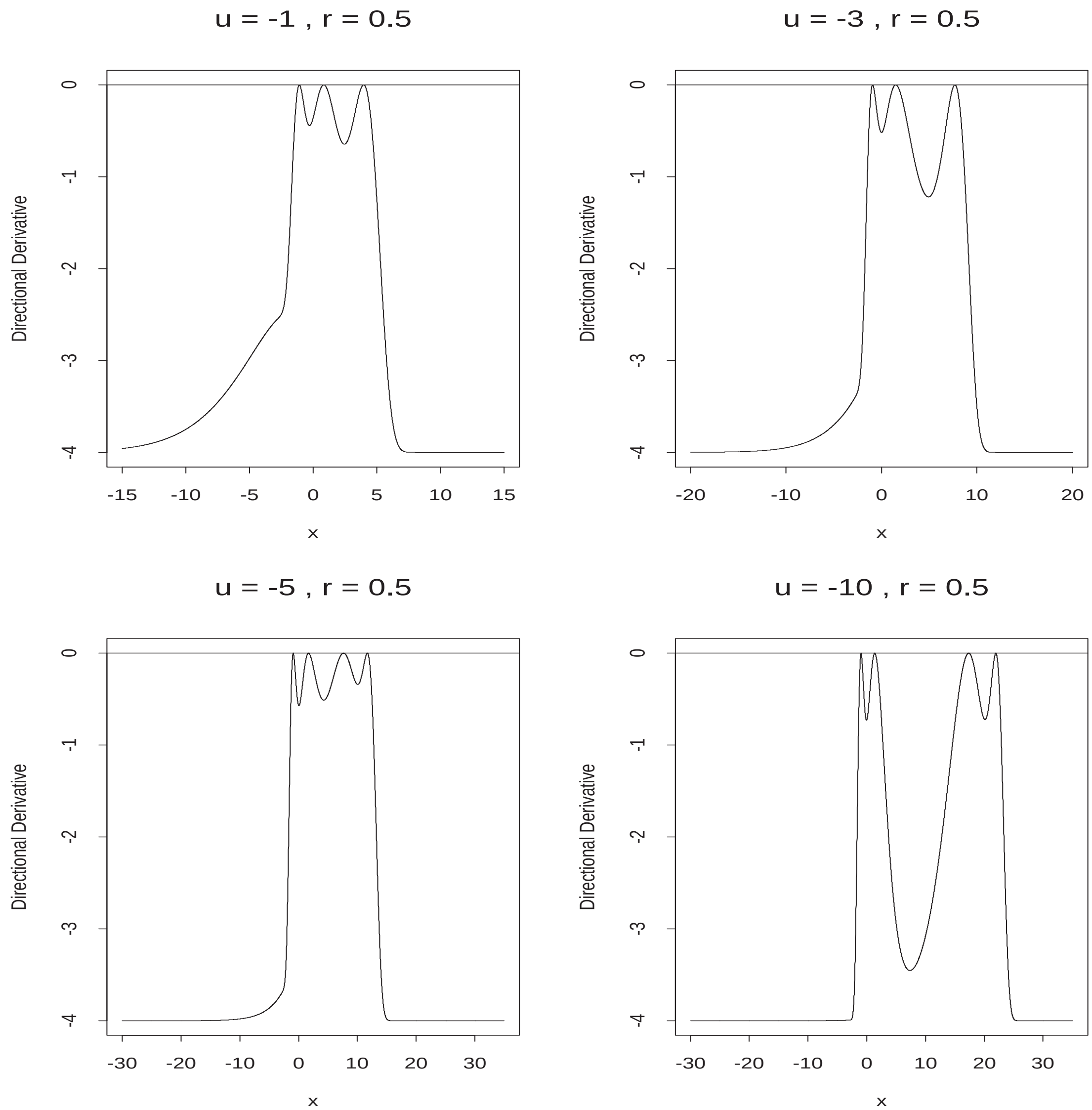

Figure 4.4: Directional derivatives of D-optimal designs for the positive-negative extreme value model for different values of μ and $r = 0.5$.

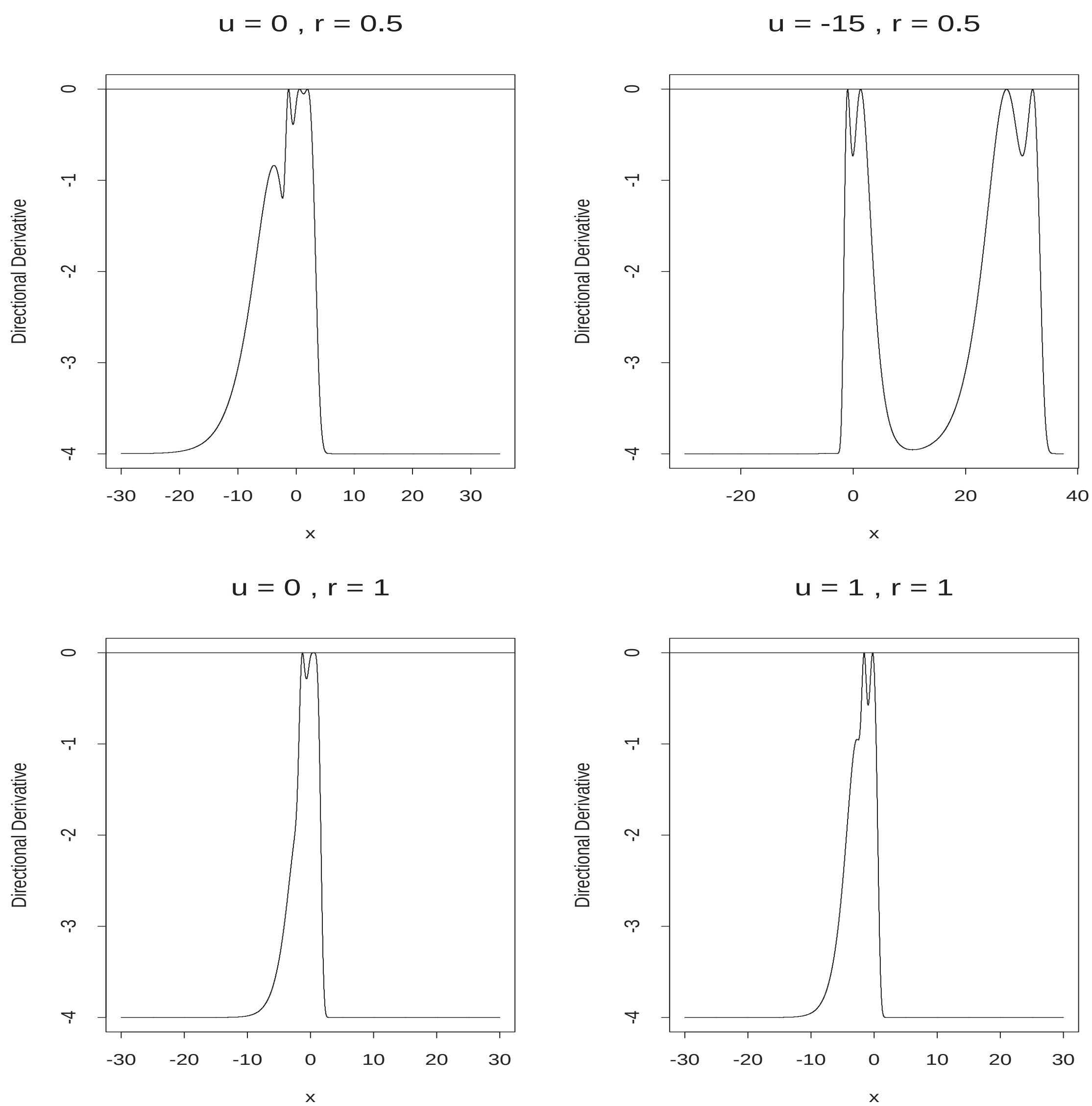

Figure 4.5: Directional derivatives of D-optimal designs for the positive-negative extreme value model with different values of μ and $r = 0.5, 1$.

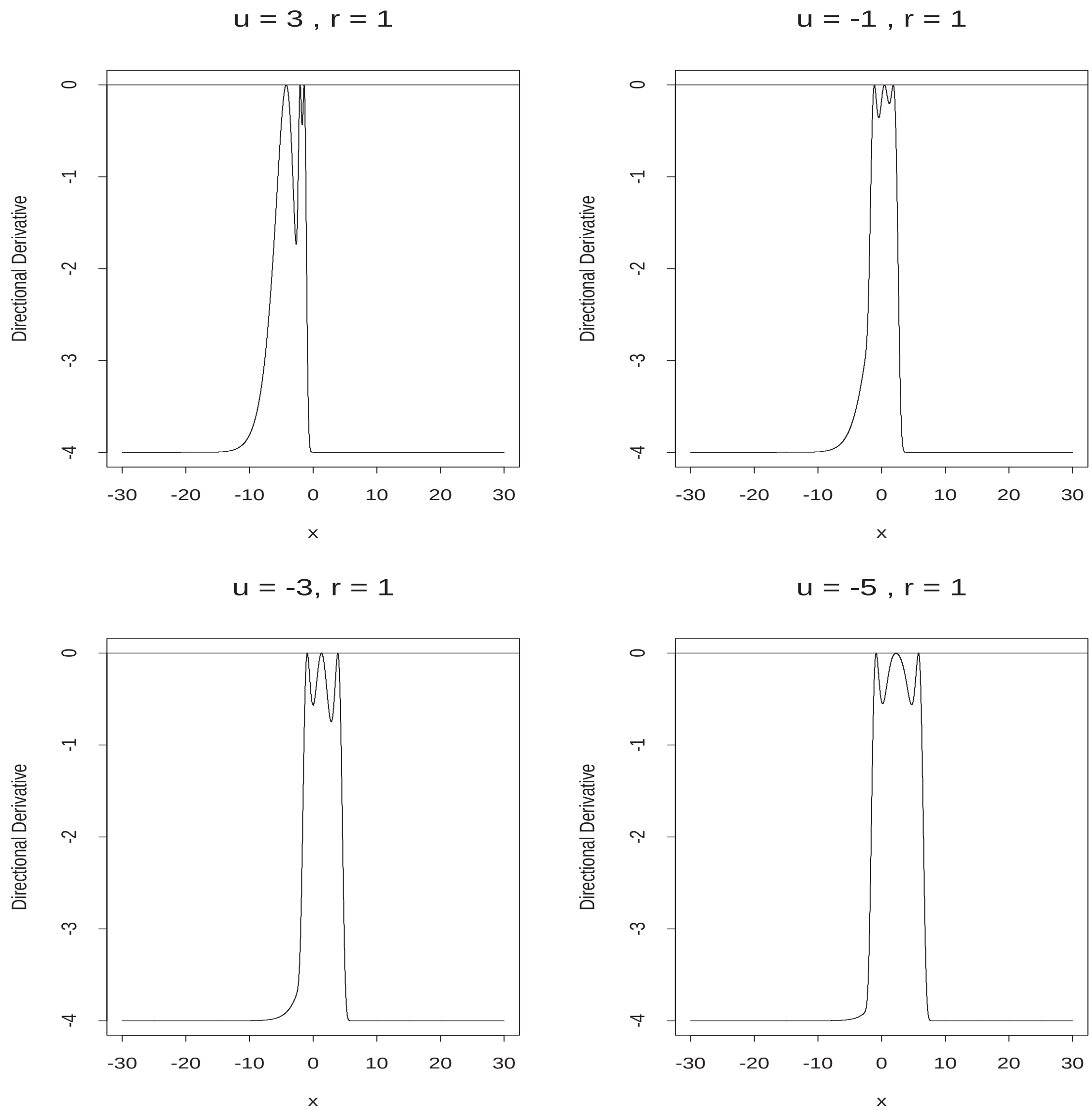

Figure 4.6: Directional Derivatives of D-optimal designs for the positive-negative extreme value model for different values of μ and $r = 1$.

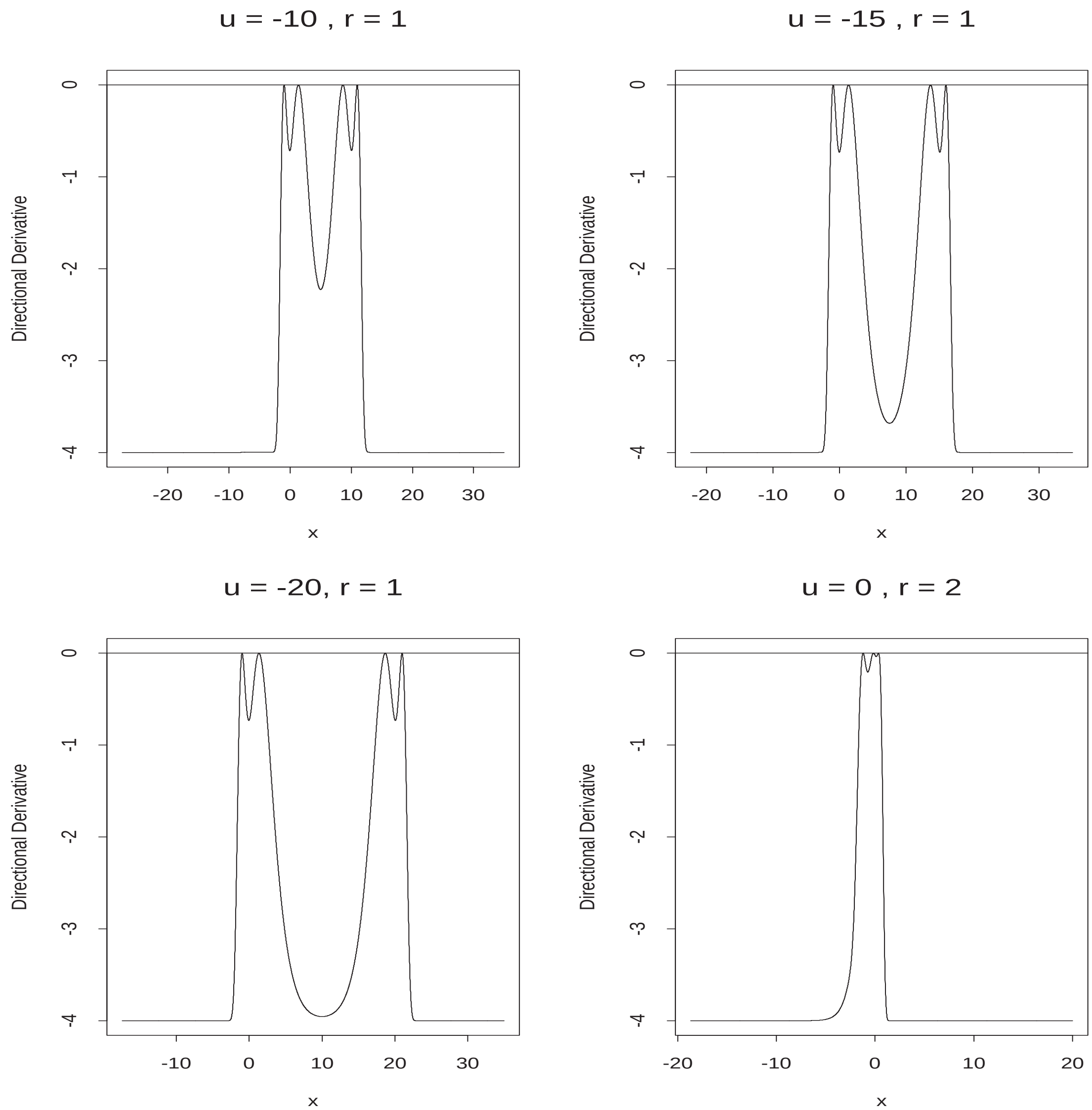

Figure 4.7: Directional derivatives of D-optimal designs for the positive-negative extreme value model for different values of μ and $r = 1, 2$.

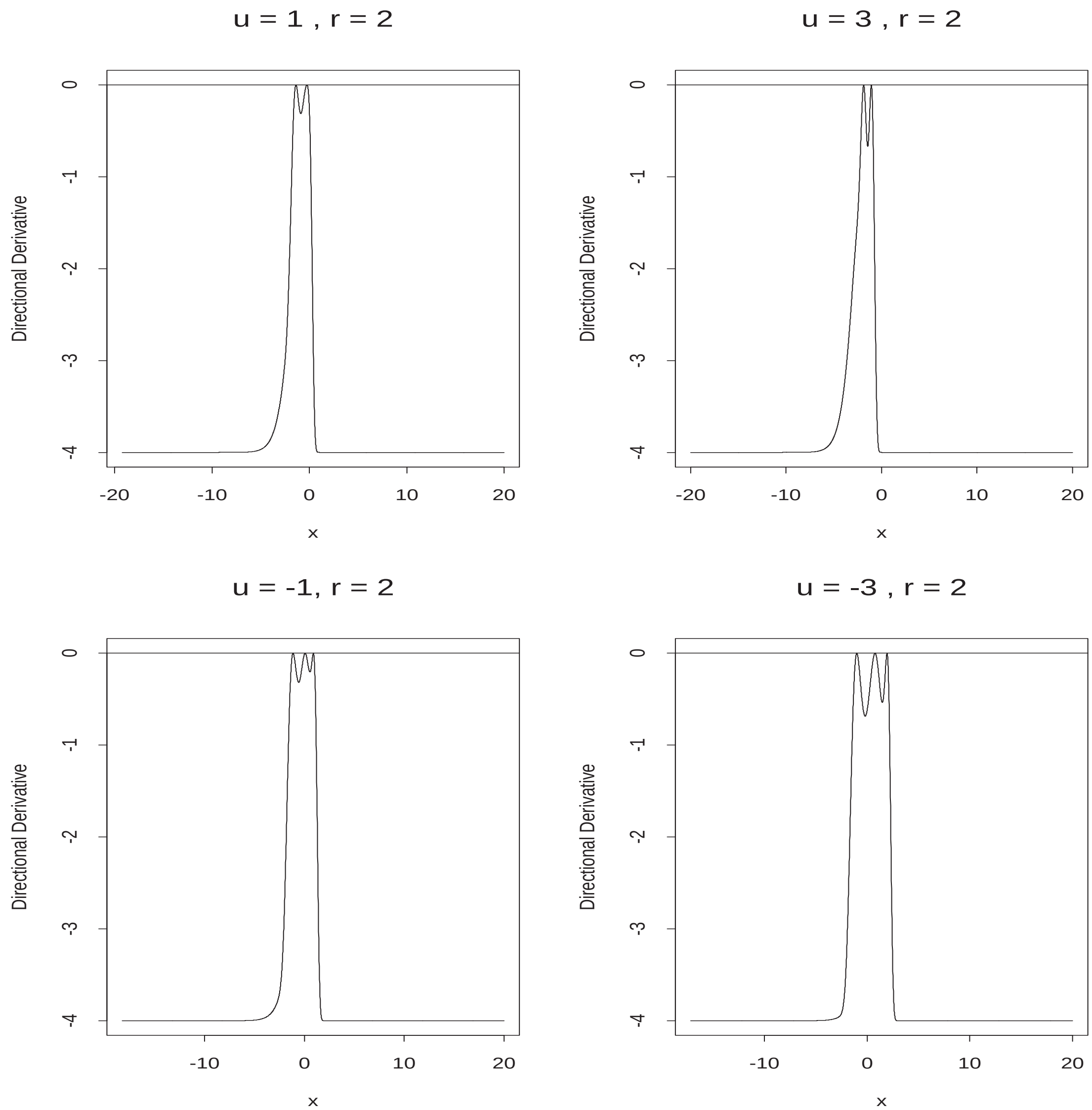

Figure 4.8: Directional derivatives of D-optimal designs for the positive-negative extreme value model for different values of μ and $r = 2$.

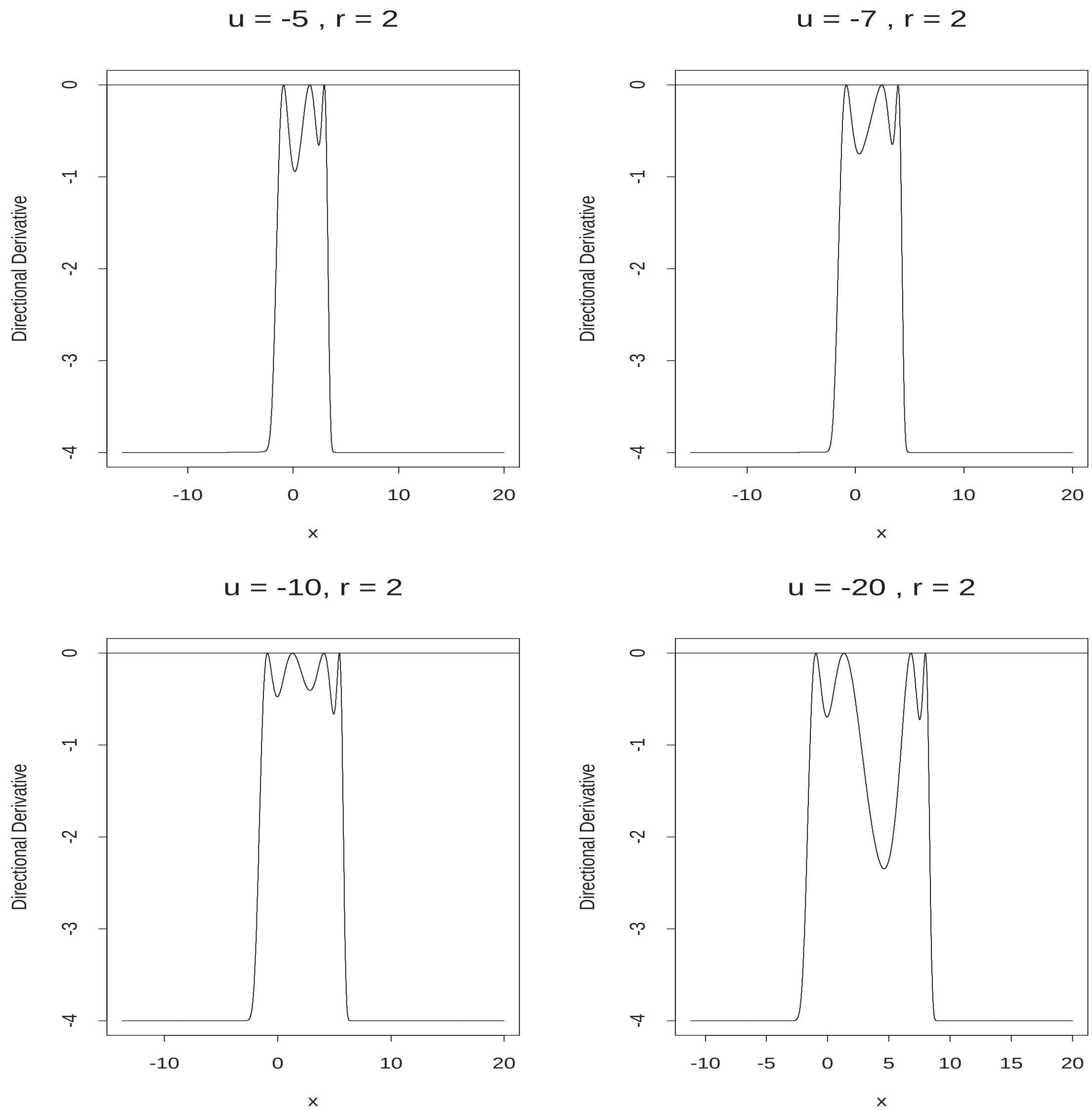

Figure 4.9: Continued directional Derivatives of D-optimal designs for the positive-negative extreme value model for different values of μ and $r = 2$.

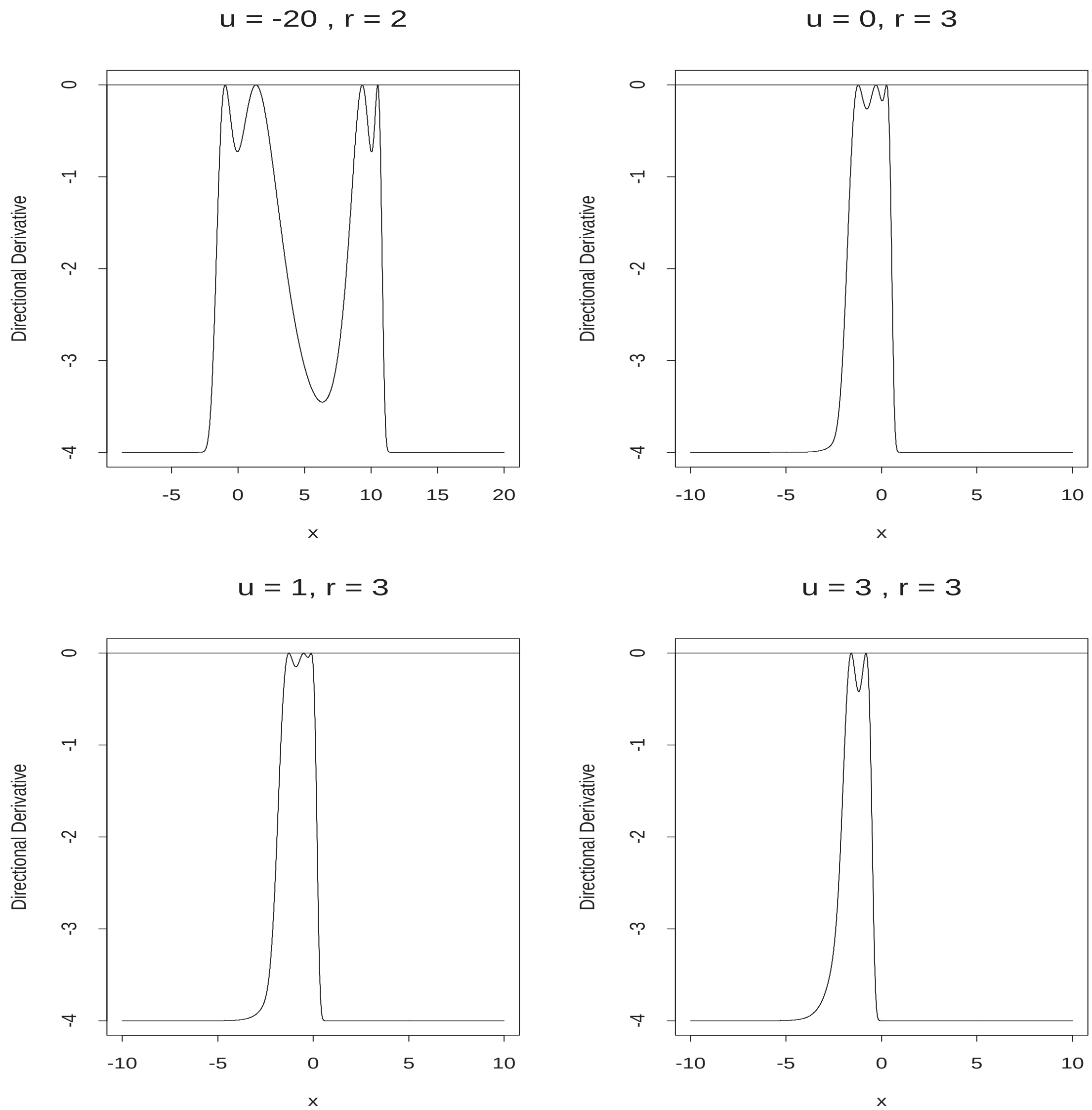

Figure 4.10: Directional derivatives of D-optimal designs for the positive-negative extreme value model for different values of μ and $r = 2, 3$.

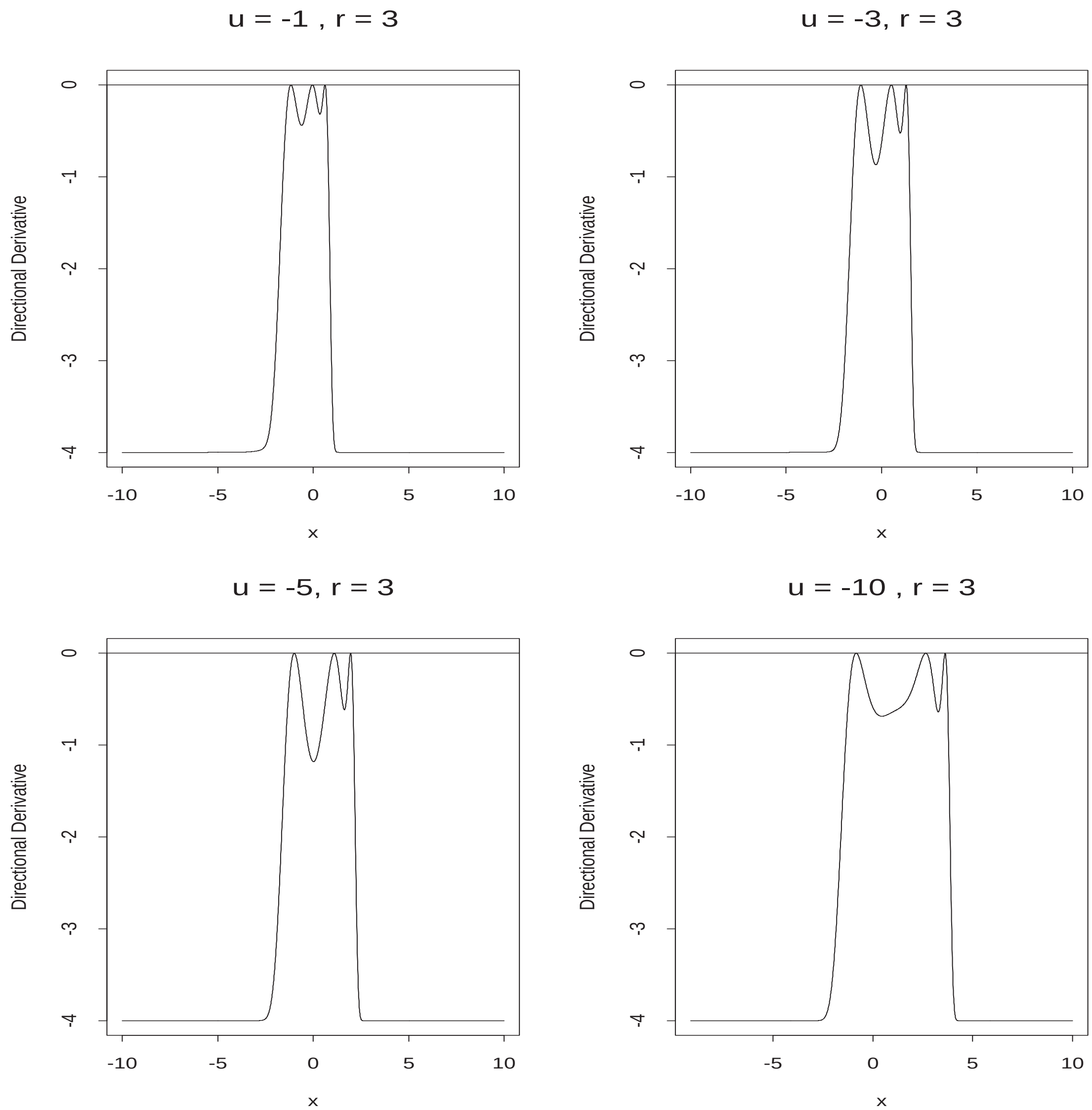

Figure 4.11: Directional derivatives of D-optimal designs for the positive-negative extreme value model for different values of μ and $r = 3$.

4.2 Equal slopes $\beta_1 = \beta_2 = \beta$

Consider the contingent response model with $F_x = F(\alpha_1 + \beta x)$ and $G_x = G(\alpha_2 + \beta x)$. Reparameterize $\mathbf{\Theta} = (\alpha_1, \beta, \alpha_2)$ to $\boldsymbol{\theta} = (\alpha_2, \beta, \mu)$, where $\mu = \alpha_1 - \alpha_2$. When $\boldsymbol{\theta} = (0, 1, \mu)$, we say we have a *canonical* (μ) *model.* Let $\boldsymbol{\xi}^*$ be the D-optimal design for the canonical (μ) model. Theorem 4.2.1 tells us that the optimal designs for all other (μ) models can be generated by a linear transformation of the canonical optimal designs.

Theorem 4.2.1 *If the design* $\boldsymbol{\xi_0}^* = \begin{pmatrix} x_1^* & \dots & xK^* \\ \xi_1^* & \dots & \xi_K^* \end{pmatrix}$ *is locally D-optimal for* $\boldsymbol{\theta_0} = (0, 1, \mu)$*, then the design* $\boldsymbol{\xi}^* = \begin{pmatrix} (\frac{x_1^* - \alpha_2}{\beta}, & \dots, & \frac{x_K^* - \alpha_2}{\beta}) \\ \xi_1^*, & \dots, & \xi_K^* \end{pmatrix}$ *is locally D-optimal for* $\boldsymbol{\theta_0} = (\alpha_2, \beta_2, \mu)$.

Proof: We show that $\det \mathbf{M}(\boldsymbol{\theta}, \boldsymbol{\xi}) = 1/\beta^2 \ \det \mathbf{M}(\boldsymbol{\xi_0}, \boldsymbol{\theta_0})$. Recall from Lemma 3.1.1 that

$$\mathbf{I}_{3\times 3}(x, \mathbf{\Theta}) = v(x, \mathbf{\Theta}) \begin{pmatrix} 1 & x & 0 \\ x & x^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} + w(x, \mathbf{\Theta}) \begin{pmatrix} 0 & 0 & 0 \\ 0 & x^2 & x \\ 0 & x & 1 \end{pmatrix}.$$

Thus

$$\det \mathbf{M}(\boldsymbol{\theta}, \boldsymbol{\xi}) \quad = \quad \det \left(\sum_i^K \xi_i \begin{pmatrix} v_i & v_i \frac{(x_i - \alpha_1)}{\beta} & 0 \\ v_i \frac{(x_i - \alpha_1)}{\beta} & (v_i + w_i) \frac{(x_i - \alpha_1)^2}{\beta^2} & w_i \frac{(x_i - \alpha_1)}{\beta} \\ 0 & w_i \frac{(x_i - \alpha_1)}{\beta} & w_i \end{pmatrix} \right)$$

$$
\begin{aligned}
&= \frac{1}{\beta^2}\,\det\sum_i^K \xi_i \begin{pmatrix} v_i & v_i(x_i-\alpha_1) & 0 \\ v_i(x_i-\alpha_1) & (v_i+w_i)(x_i-\alpha_1)^2 & w_i(x_i-\alpha_1) \\ 0 & w_i(x_i-\alpha_1) & w_i \end{pmatrix} \\
&= \frac{1}{\beta^2}[(\sum_i^K \xi_i v_i)\{\sum_i^K \xi_i w_i(\sum_i^K \xi_i(v_i+w_i)(x_i-\alpha_1)^2) \\
&\quad -(\sum_i^K \xi_i w_i(x_i-\alpha_1))^2\} \\
&\quad -(\sum_i^K \xi_i v_i(x_i-\alpha_1))[\sum_i^K \xi_i w_i \sum_i^K \xi_i v_i(x_i-\alpha_1)]] \\
&= \frac{1}{\beta^2}(\sum_i^K \xi_i v_i)(\sum_i^K \xi_i w_i)\sum_i^K \xi_i v_i x_i^2 \\
&\quad +(\sum_i^K \xi_i v_i)(\sum_i^K \xi_i w_i)\sum_i^K \xi_i v_i(-2\alpha_1 x_i+\alpha_i^2) \\
&\quad +(\sum_i^K \xi_i v_i)(\sum_i^K \xi_i w_i)\sum_i^K \xi_i w_i x_i^2 \\
&\quad +(\sum_i^K \xi_i v_i)(\sum_i^K \xi_i w_i)\sum_i^K \xi_i w_i(-2\alpha_1 x_i+\alpha_i^2) \\
&\quad -(\sum_i^K \xi_i v_i)\{\sum_i^K\sum_j^K \xi_i w_i \xi_j w_j(x_i-\alpha_1)(x_j-\alpha_1)\} - \{\sum_i^K \xi_i v_i x_i \\
&\quad -\alpha_1\sum_i^K \xi_i v_i\}\{(\sum_i^K \xi_i w_i)\sum_i^K \xi_i v_i x_i - \alpha_1(\sum_i^K \xi_i w_i)\sum_i^K \xi_i v_i\} \\
&= \frac{1}{\beta^2}(\sum_i^K \xi_i v_i)(\sum_i^K \xi_i w_i)\sum_i^K \xi_i v_i x_i^2 + (\sum_i^K \xi_i v_i)(\sum_i^K \xi_i w_i) \\
&\quad \sum_i^K \xi_i v_i(-2\alpha_1 x_i+\alpha_i^2) + (\sum_i^K \xi_i v_i)(\sum_i^K \xi_i w_i)\sum_i^K \xi_i w_i x_i^2 \\
&\quad +(\sum_i^K \xi_i v_i)(\sum_i^K \xi_i w_i)\sum_i^K \xi_i w_i(-2\alpha_1 x_i+\alpha_i^2)
\end{aligned}
$$

$$
\begin{aligned}
& -(\sum_i^K \xi_i v_i x_i) \sum_i^K \sum_i^K \xi_i w_i \xi_j w_j x_i x_j \\
& + \sum_i^K \xi_i v_i x_i) \sum_i^K \sum_i^K \xi_i w_i \xi_j w_j \{\alpha_1 x_j - \alpha_1 x_i + \alpha_i^2\} \\
& - \sum_i^K \xi_i v_i)(\sum_i^K \xi_i w_i) \sum_i^K \xi_i v_i x_i - 2\alpha_1 (\sum_i^K \xi_i v_i)(\sum_i^k \xi_i w_i) \sum_i^K \xi_i v_i x_i \\
& + \alpha_1^2 (\sum_i^K \xi_i v_i)(\sum_i^k \xi_i w_i) \sum_i^K \xi_i v_i \\
= \; & \frac{1}{\beta^2} (\sum_i^K \xi_i v_i)(\sum_i^K \xi_i w_i) \sum_i^K \xi_i v_i x_i^2 \\
& + (\sum_i^K \xi_i v_i)(\sum_i^K \xi_i w_i) \sum_i^K \xi_i w_i x_i^2 - (\sum_i^k \xi_i v_i)(\sum_i^K \xi_i w_i x_i)^2 \\
& - (\sum_i^K \xi_i v_i)(\sum_i^K \xi_i w_i) \sum_i^K \xi_i v_i x_i \\
= \; & \frac{1}{\beta^2} \det \left(\begin{pmatrix} \sum_i^K \xi_i v_i & \sum_i^K \xi_i v_i x_i & 0 \\ \sum_i^K \xi_i v_i x_i & \sum_i^K \xi_i (v_i + w_i) x_i^2 & \sum_i^K \xi_i w_i x_i \\ 0 & \sum_i^K \xi_i w_i x_i & \xi_i w_i \end{pmatrix} \right) \\
= \; & \frac{1}{\beta^2} \det \left(\sum_i^k \xi_i (\begin{pmatrix} 1 & v_i x_i & 0 \\ v_i x_i & (v_i + w_i) x_i^2 & w_i x_i \\ 0 & w_i x_i & w_i \end{pmatrix} \right) \\
= \; & \frac{1}{\beta^2} \det \mathbf{M}(\boldsymbol{\theta_0}, \boldsymbol{\xi_0}).
\end{aligned}
$$

□

Locally D-optimal designs for several canonical (μ) positive-negative extreme value models are given in Tables 4.6 and 4.7. All optimal designs were verified by the General Equivalence Theorem given by Theorem 2.3.1. Figures 4.12–4.15 show

D-optimal designs to consist of two, three and four points for small, moderate and large negative values of μ, respectively. For positive values of μ, the optimal designs consist of two points. For large negative values of μ, the optimal designs are not the same as the optimal designs for the separate $\bar{G}_x$ and $F_{\mu+rx}$ models concatenated. However, they still have four point designs:

$$\boldsymbol{\xi_{PNE2}} = \begin{pmatrix} 0.8537 & -1.0773 & (-0.8537-\mu) & (1.0773-\mu) \\ 0.2900 & 0.2100 & 0.2900 & 0.2100 \end{pmatrix}.$$

We will show in Section 6.2 that the limiting locally D-optimal designs is $\boldsymbol{\xi}^*_{PNE2}$. Having four design points is consistent with the results of Fan and Chaloner [12] for the logistic-logistic model, but it is different in having unequal weights. This may be because the extreme value function is asymmetric.

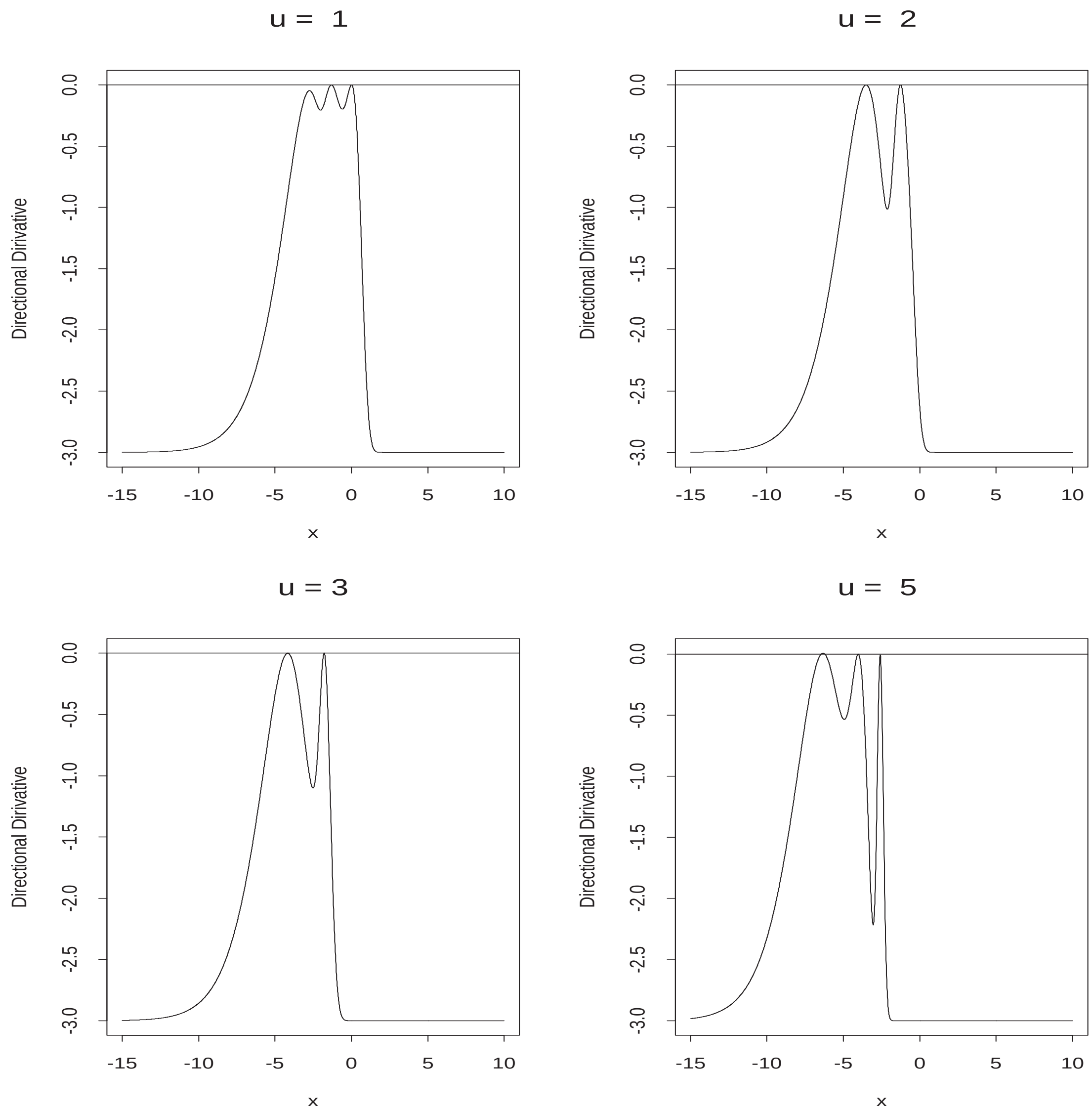

Figure 4.12: Directional derivatives for the positive-negative extreme model when $\mu = 1, 2, 3, 4$ for D-optimal designs.

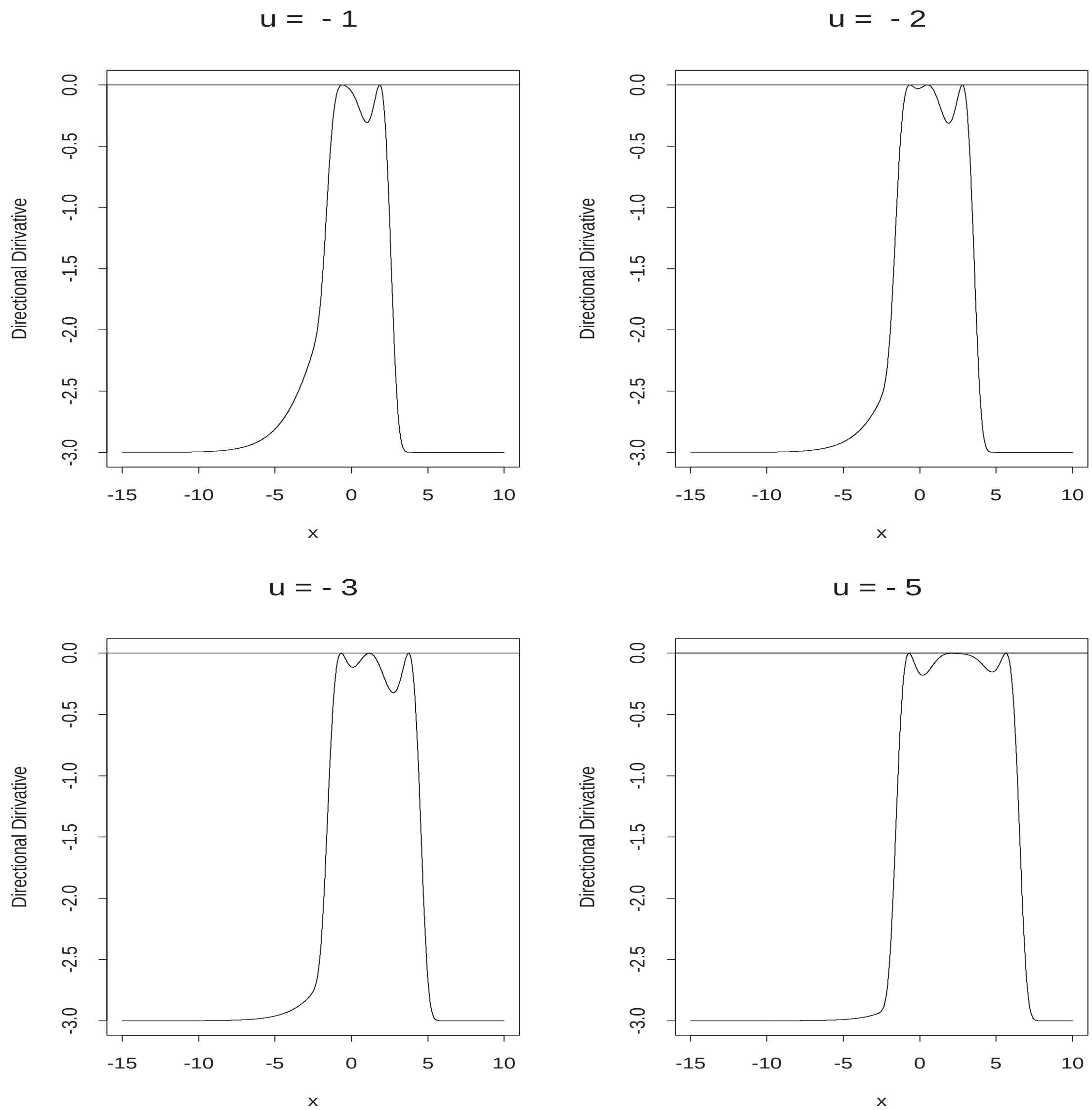

Figure 4.13: Directional derivatives for the positive-negative extreme model when $\mu = -1, -2, -3, -5$ for D-optimal designs.

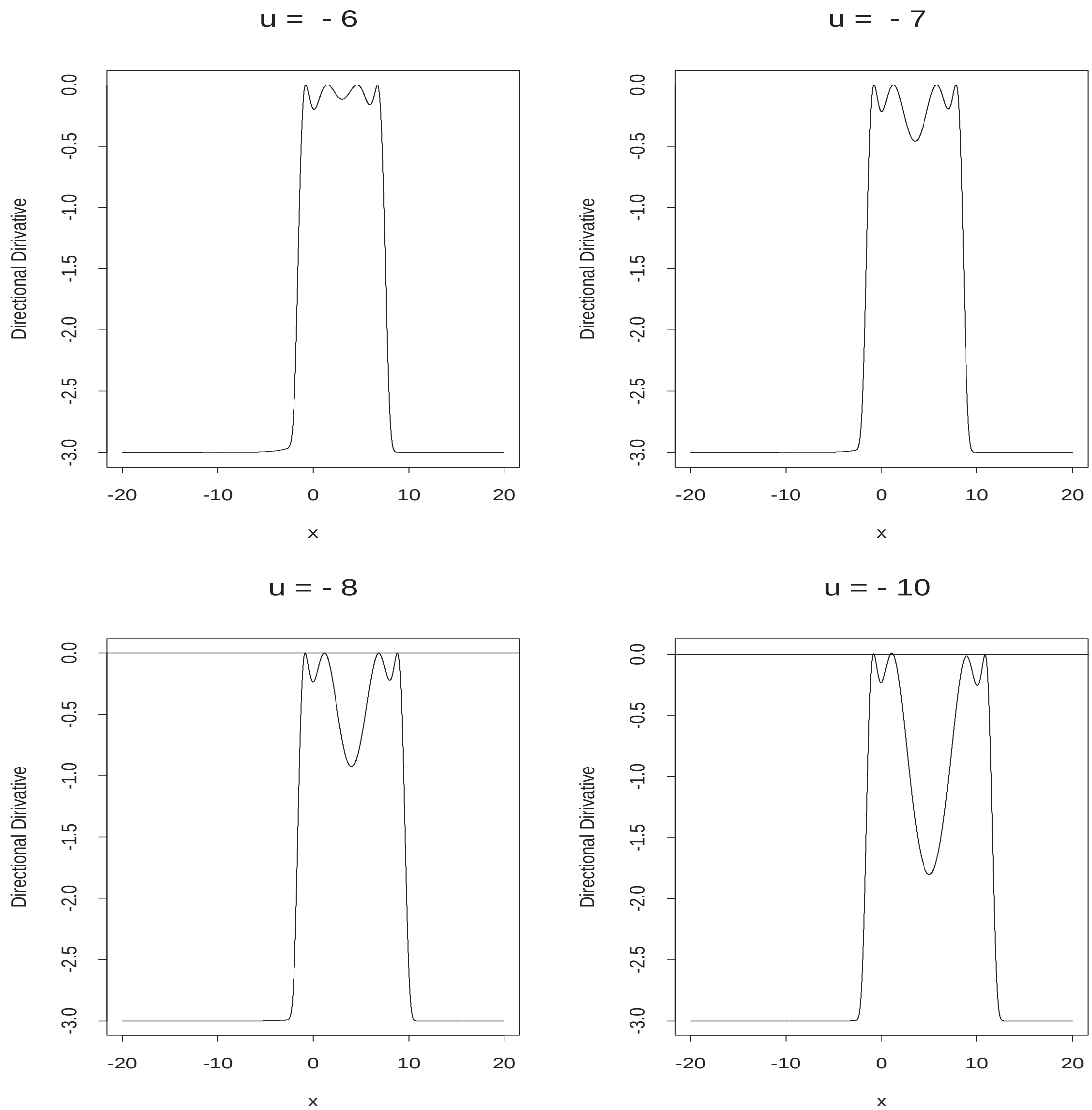

Figure 4.14: Directional derivatives for the positive-negative extreme model when $\mu = -6, -7, -8, -10$ for D-optimal designs.

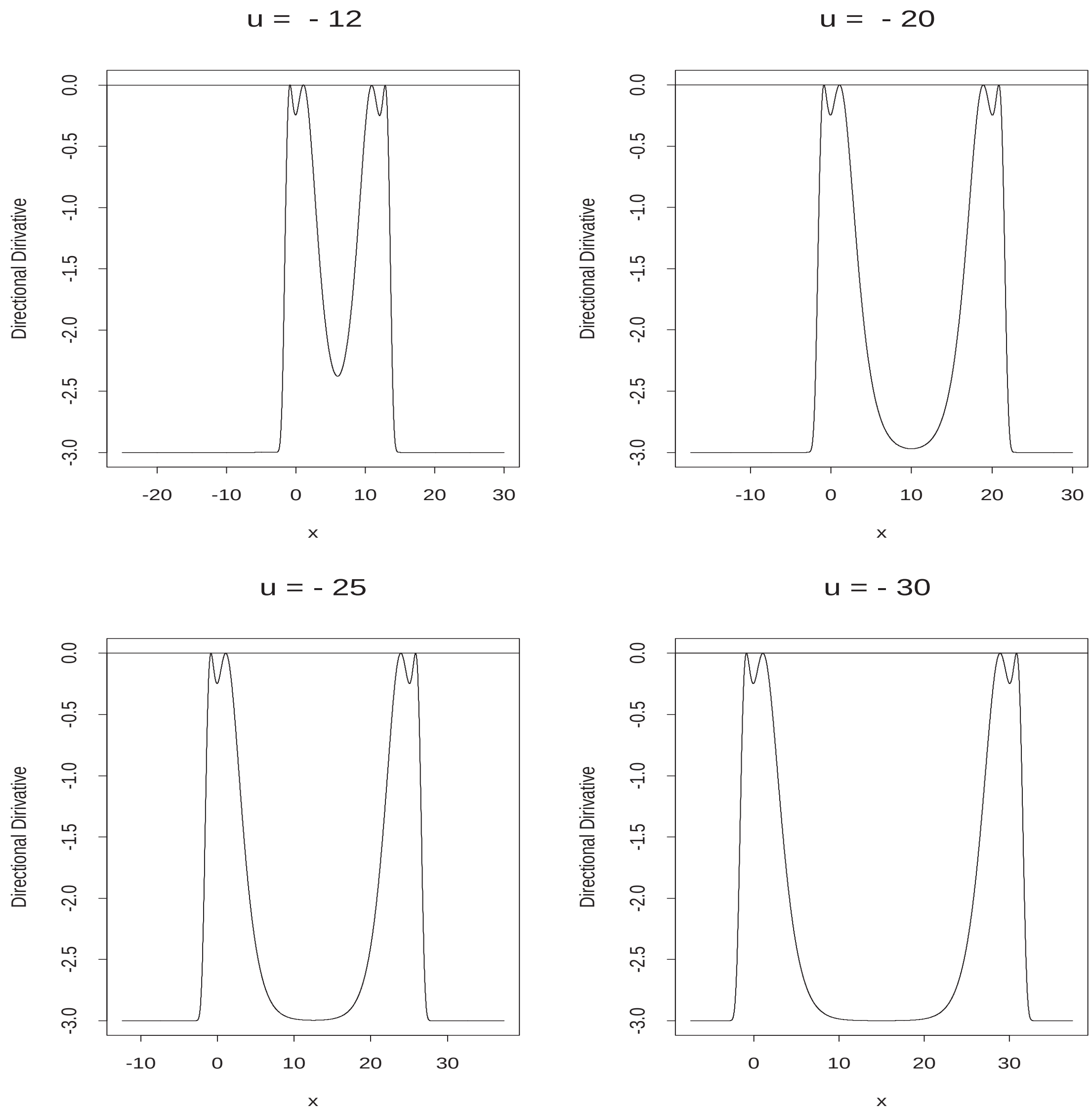

Figure 4.15: Directional derivatives for the positive-negative extreme model when $\mu = -12, -20, -25, -30$ for D-optimal designs.

4.3 Efficiency of D-Optimal Designs

In this section we investigate the efficiency of the optimal designs under several alternatives. We illustrate with two examples, one when the location parameter μ is relatively small and the other when μ large. We want to explore how efficient a design is if the points used are not the optimal ones. Following Atkinson and Donev [3], define the D-efficiency of an arbitrary design $\boldsymbol{\xi}$ as

$$D_{eff} = \{\frac{|\mathbf{M}(\boldsymbol{\xi})|}{|\mathbf{M}(\boldsymbol{\xi}^*)|}\}^{1/p},$$

where p is the number of the model parameters. The ratio of the determinants of any arbitrary design $\boldsymbol{\xi}$ and the optimal design $\boldsymbol{\xi}^*$ is proportional to the design size when raised to the power $1/p$, irrespective of the dimension of the model.

The optimal design when $\beta_1 = \beta_2$ and $\mu = -1$ is $\boldsymbol{\xi}^*_{PNE2}=\begin{pmatrix} -0.5911 & 1.8519 \\ 0.6496 & 0.3504 \end{pmatrix}$. Figure 4.16 shows the efficiency of a design when $x_1 = -0.5911$ is fixed and x_2 varies away from 1.8519. By using a point less than the optimal point 1.8519 and keeping the same weights, we see the that the efficiency decreases with the distance from 1.8519 while the probability of success increases to reach its maximum 0.297 at the point 0.5. Using the optimal dose, i.e. $x_2 = 0.5$, reduces efficiency of parameters estimators to less than 60% of the optimal design. Similarly for $\mu = -10$, the optimal design is $\boldsymbol{\xi} = \begin{pmatrix} 10.8483 & 8.9041 & 1.0914 & -.8462 \\ 0.2906 & 0.2113 & 0.2084 & 0.2897 \end{pmatrix}$. Figure 4.17 shows the efficiency of designs using a point less than the optimal $x_1 = 10.8483$, while x_2, x_3, x_4 are fixed

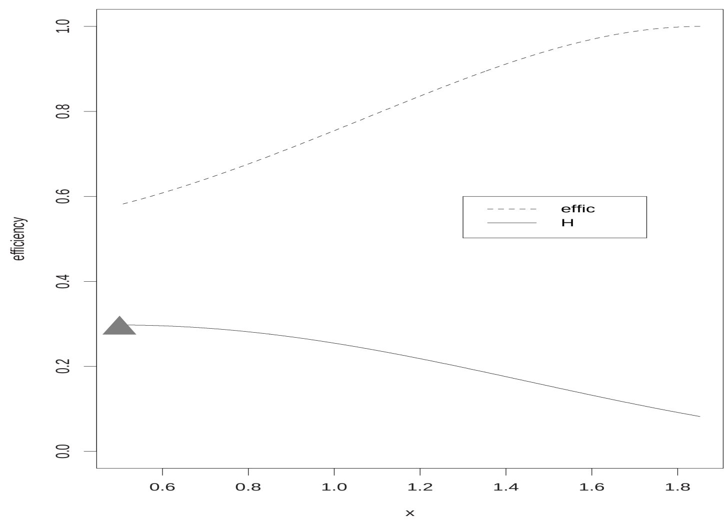

Figure 4.16: Efficiency plot for $\mu = -1$.

at the optimal levels with same weights. The efficiency decreases with the distance from 10.8483 while the probability of success increases towards its maximum 0.9866 at the point 5.

The optimal designs are not always usable. This is true for example, when the optimal designs put large positive weight on points that have high toxicity. One needs to decide how much efficiency to compromise in order to reduce the toxicity probability. Plots such as 4.16-4.18 may help in making this tradeoff.

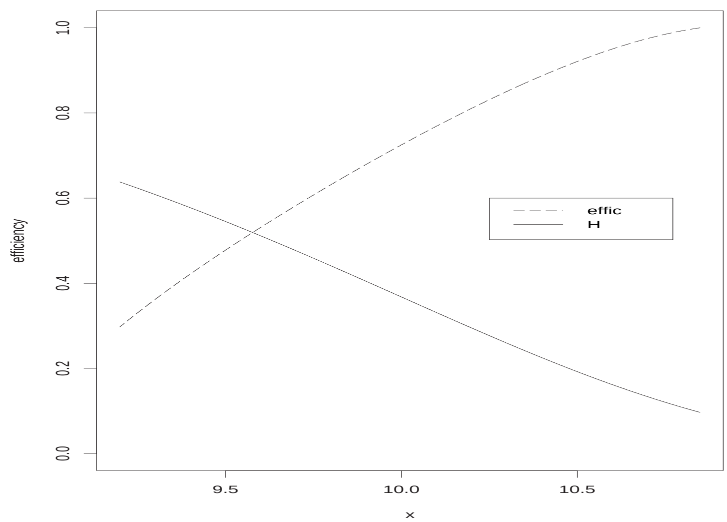

Figure 4.17: Efficiency plot for $\mu = -10$.

μ	$g(\boldsymbol{\Theta})$	x	Weights	$F(x)$	$\bar{F}(x)\bar{G}(x)$	$H(x)$
5	-2.50	-6.3369	0.3316	0.23	0.77	0.00
		-4.0204	0.3305	0.93	0.07	0.00
		-2.5862	0.3379	1.00	0.00	0.00
3	-1.50	-4.1760	0.3333	0.27	0.74	0.00
		-1.7889	0.6667	0.97	0.04	0.00
2	-1.00	-3.5131	0.3334	0.20	0.80	0.00
		-1.2554	0.6667	0.88	0.12	0.00
1	-0.50	-1.3120	0.6164	0.52	0.47	0.01
		0.0031	0.3836	0.94	0.04	0.02
-1	0.50	-0.5911	0.6496	0.18	0.68	0.13
		1.8519	0.3504	0.90	0.01	0.08
-2	1.00	-0.6450	0.4091	0.07	0.79	0.14
		0.5111	0.2675	0.20	0.36	0.44
		2.7947	0.3233	0.89	0.01	0.10
-5	2.50	-0.6986	0.3367	0.00	0.86	0.13
		2.101	0.3407	0.05	0.11	0.84
		5.6449	0.3226	0.85	0.00	0.15
-6	3	-0.7635	0.3141	0.00	0.88	0.12
		1.4694	0.2286	0.01	0.20	0.79
		4.5993	0.1539	0.22	0.01	0.77
		6.7238	0.30335	0.87	0.00	0.13
-7	3.50	-0.8060	0.3012	0.00	0.89	0.10
		1.2547	0.2178	0.00	0.25	0.75
		5.7817	0.1869	0.26	0.00	0.74
		7.7862	0.2942	0.89	0.00	0.11

Table 4.6: D-optimal designs for the positive-negative extreme model when $\beta_1 = \beta_2 = \beta$.

μ	$g(\boldsymbol{\Theta})$	x	Weights	$F(x)$	$\bar{F}(x)\bar{G}(x)$	$H(x)$
-8	4.00	-0.8299	0.2949	0.00	0.90	0.10
		1.1613	0.2140	0.00	0.27	0.73
		6.8574	0.2001	0.27	0.00	0.73
		8.8203	0.2910	0.90	0.00	0.10
-10	5.00	-0.8462	0.2897	0.00	0.90	0.10
		1.0914	0.2084	0.00	0.29	0.72
		8.9041	0.2113	0.28	0.00	0.72
		10.8483	0.2906	0.90	0.00	0.10
-12	6.0	-0.8522	0.2595	0.00	0.90	0.10
		1.080	0.2101	0.00	0.29	0.71
		10.9189	0.2107	0.29	0.00	0.71
		12.8526	0.2897	0.90	0.00	0.10
-20	10.00	-0.8537	0.28958	0.00	0.91	0.10
		1.0773	0.2105	0.00	0.29	0.71
		18.9227	0.2105	0.29	0.00	0.71
		20.8537	0.2895	0.91	0.00	0.10
-25	12.50	-0.8537	0.2895	0.00	0.91	0.10
		1.0773	0.2105	0.00	0.29	0.71
		23.9227	0.2105	0.29	0.00	0.71
		25.8537	0.2895	0.91	0.00	0.10
-30	15.00	-0.8537	0.2895	0.00	0.91	0.10
		1.0773	0.2105	0.00	0.29	0.71
		28.9223	0.2105	0.29	0.00	0.71
		30.8537	0.2895	0.91	0.00	0.10

Table 4.7: Continued D-optimal designs for the positive-negative extreme model when $\beta_1 = \beta_2 = \beta$.

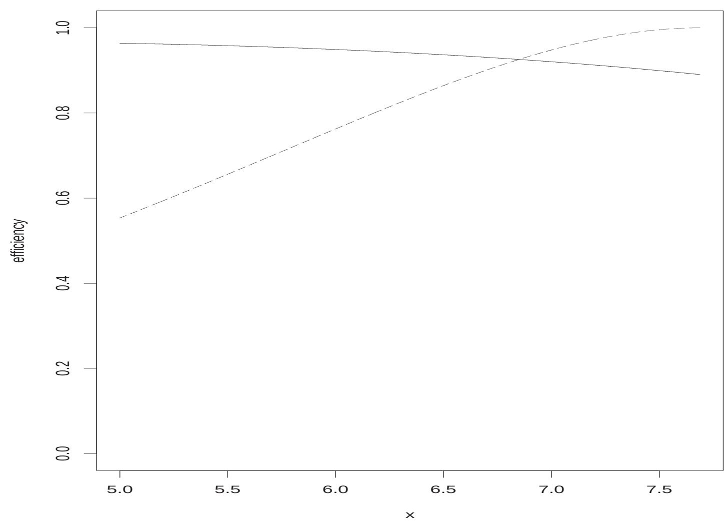

Figure 4.18: Efficiency plot for $\mu = -3, r = 0.5$.

Chapter 5

LOCALLY C-OPTIMAL DESIGNS

Our interest is to find the c-optimal design for estimating the optimal dose $\nu = g(\mathbf{\Theta})$. For non-linear models, the c-optimal criterion is to minimize

$$\phi_c(\mathbf{\Theta}, \xi) = \dot{\mathbf{g}}^{\mathbf{T}}(\mathbf{\Theta})\mathbf{M}^{-\mathbf{1}}(\xi, \mathbf{\Theta})\dot{\mathbf{g}}(\mathbf{\Theta}),$$

where $\dot{\mathbf{g}}$ is the gradient vector of $g(\mathbf{\Theta})$. The optimal designs for the positive-negative extreme model are verified using the General Equivalence Theorem. Singular optimal designs occur. In this case, Silvey's Theorem was used to verify optimality (See Silvey [37] page 16).

5.1 Unequal slopes $\beta_1 \neq \beta_2$

We find c-optimal designs for the optimal dose for the positive-negative extreme model, namely, $\nu = g(\boldsymbol{\Theta}) = (log(\beta_2/\beta_1) - \alpha_1 - \alpha_2)/(\beta_1 + \beta_2)$. The gradient vector is $\partial g(\boldsymbol{\Theta})/\partial \boldsymbol{\Theta_i} =$

$$\left(\frac{-1}{\beta_1+\beta_2}; \quad \frac{-1}{\beta_1(\beta_1+\beta_2)} - \frac{log(\beta_2/\beta_1)-\alpha_1-\alpha_2}{(\beta_1+\beta_2)^2}; \quad \frac{-1}{\beta_1+\beta_2}; \quad \frac{1}{\beta_2(\beta_1+\beta_2)} - \frac{log(\beta_2/\beta_1)-\alpha_1-\alpha_2}{(\beta_1+\beta_2)^2} \right)^T.$$

We reparameterize as in Section 4.1, that is $\boldsymbol{\Theta} = (\alpha_1, \beta_1, \alpha_2, \beta_2)$ to $\boldsymbol{\theta} = (\alpha_2, \beta_2, \mu, r)$. It follows from Theorem 5.1.1 that once the locally c-optimal designs are found for a canonical (μ, r) model, the locally c-optimal designs for all other models in the family can be found from the canonical design.

Theorem 5.1.1 $\boldsymbol{\xi_0} = \begin{pmatrix} x_1^* & \dots & x_K^* \\ \xi_1^* & \dots & \xi_K^* \end{pmatrix}$ *is locally c-optimal for* $\boldsymbol{\theta_0} = (0, 1, \mu, r)$, *then*

$$\boldsymbol{\xi} = \begin{pmatrix} \frac{x_1^*-\alpha_2}{\beta_2} & \dots & \dots, \frac{x_K^*-\alpha_2}{\beta_2}) \\ \xi_1^* & \dots & \xi_K^* \end{pmatrix}$$ *is locally c-optimal for* $\boldsymbol{\theta_0} = (\alpha_2, \beta_2, \mu, r)$.

Proof: It can be shown that $\phi_c(\boldsymbol{\theta}, \boldsymbol{\xi}) = (1/\beta_2^2)\phi_c(\boldsymbol{\theta_0}, \boldsymbol{\xi_0})$ using Maple software and simplifying. See Appendix A1. The rest of the proof is analogous to that of Theorem 4.1.1. □

Tables 5.1 and 5.2 give c-optimal designs for the canonical (μ, r) positive-negative extreme model. All the optimal designs consist of two points. For large positive values of μ the probability of success is negligible and not of interest. All optimal designs were verified by the General Equivalence Theorem and their directional derivatives are shown in Figures 5.1- 5.7. These directional derivatives are non-positive and achieve their maximum at the optimal design points.

5.2 Equal slopes $\beta_1 = \beta_2 = \beta$

If $\beta_1 = \beta_2 = \beta$, the solution to (3.2) is $\nu = g(\boldsymbol{\Theta}) = -(\alpha_1 + \alpha_2)/2\beta$ with gradient $\partial g(\boldsymbol{\Theta})/\partial \boldsymbol{\Theta_i} = \begin{pmatrix} \frac{-1}{2\beta} & \frac{\alpha_1+\alpha_2}{2\beta^2} & \frac{-1}{2\beta} \end{pmatrix}^T$. We reparameterize as in Section (4.2) on page 41, that is $\boldsymbol{\Theta} = (\alpha_2, \beta, \alpha_1)$ to $\boldsymbol{\theta}$=(α_2, β, μ). It follows from Theorem 5.2.1 that once the locally c-optimal designs are found for the canonical (μ) model, the locally c-optimal designs for all other models in the family can be found from the canonical (μ) design.

Theorem 5.2.1 *If the design* $\boldsymbol{\xi}_0^*$=$\begin{pmatrix} x_1^* & \dots & xK^* \\ \\ \xi_1^* & \dots & \xi_K^* \end{pmatrix}$ *is locally c-optimal for* $\boldsymbol{\theta_0}$= $(0, 1, \mu)$, *then the design* $\boldsymbol{\xi}^*$= $\begin{pmatrix} (\frac{x_1^*-\alpha_2}{\beta}, & \dots, & \frac{x_K^*-\alpha_2}{\beta}) \\ \\ \xi_1^*, & \dots, & \xi_K^* \end{pmatrix}$ *is locally c-optimal for* $\boldsymbol{\theta}$= (α_2, β_2, μ).

Proof: The proof is analogous to that of Theorem 5.2.2. See Appendix A2. □

Table 5.3 gives two point c-optimal designs for some canonical (μ) positive-negative extreme models. All optimal designs were verified by the General Equivalence Theorem and their direction derivatives are shown in Figures 5.1-5.4. These directional derivatives are non-positive and achieve their maximum at the optimal design points. When F_x and $\bar{G}_x$ are quite separate, the optimal design points are

r	μ	$g(\boldsymbol{\Theta})$	x	Weights	$F(x)$	$\bar{F}(x)\bar{G}(x)$	$H(x)$
0.5	0	0.46	-0.5830	0.5265	0.53	0.40	0.08
			2.3137	0.4735	0.96	0.00	0.04
	-1	1.13	0.1037	0.6005	0.32	0.40	0.28
			3.8163	0.3995	0.92	0.00	0.08
	-3	2.46	1.4519	0.7756	0.10	0.19	0.71
			6.4685	0.2244	0.72	0.00	0.28
1	1	-0.500	-1.4536	0.6030	0.47	0.52	0.01
			0.1765	0.3970	0.96	0.02	0.02
	0	0.0	-1.0323	0.5435	0.30	0.66	0.04
			1.0106	0.4565	0.94	0.02	0.05
	-1	0.50	-0.5643	0.5437	0.19	0.67	0.14
			1.7731	0.4563	0.89	0.02	0.10
	-3	1.50	0.3817	0.5890	0.07	0.46	0.47
			2.9918	0.4110	0.63	0.02	0.35
2	1	-0.57	-1.4635	0.6294	0.14	0.85	0.01
			-0.0542	0.3706	0.91	0.06	0.03
	0	-0.23	-1.2133	0.5842	0.88	0.06	0.06
			0.3745	0.4158	0.09	0.88	0.03
	-1	0.10	-0.9531	0.5489	0.05	0.87	0.07
			0.7419	0.4512	0.80	0.08	0.12
	-3	0.77	-0.5054	0.4444	0.02	0.80	0.19
			1.3595	0.5556	0.52	0.11	0.36
	-5	1.44	-0.1536	0.3206	0.01	0.69	0.31
			1.9746	0.6794	0.30	0.09	0.61
	-7	2.10	0.2305	0.2370	0.00	0.55	0.45
			2.6208	0.7631	0.16	0.06	0.78
	-9	2.77	0.7077	0.1910	0.00	0.39	0.61
			3.2801	0.8090	0.08	0.03	0.88

Table 5.1: c-optimal designs for the positive-negative extreme model for $r = 0.5, 1, 2$.

r	μ	$g(\boldsymbol{\Theta})$	x	Weights	$F(x)$	$\bar{F}(x)\bar{G}(x)$	$H(x)$
3	1	-0.5247	-1.4163	0.6336	0.04	0.95	0.02
			-0.1217	0.3664	0.85	0.10	0.05
	0	-0.2747	-1.2376	0.5923	0.024	0.95	0.03
			0.1390	0.4077	0.78	0.13	0.09
	-1	-0.0247	-1.0712	0.5415	0.02	0.93	0.05
			0.3743	0.4585	0.68	0.16	0.16
	-3	0.48	-0.8121	0.4184	0.00	0.89	0.15
			0.8361	0.5816	0.46	0.19	0.35
	-5	0.98	-0.6112	0.3081	0.00	0.84	0.16
			1.3181	0.6919	0.30	0.17	0.54
	-7	1.48	-0.4127	0.2265	0.00	0.78	0.22
			1.8120	0.7735	0.19	0.12	0.69
	-9	1.98	-0.1820	0.1704	0.00	0.70	0.30
			2.3099	0.8296	0.12	0.08	0.80
4	3	-0.88	-1.6400	0.6927	0.03	0.97	0.01
			-0.5600	0.3073	0.84	0.14	0.03
	1	-0.48	-1.3700	0.6223	0.01	0.97	0.02
			-0.1773	0.3777	0.74	0.18	0.08
	0	-0.28	-1.2425	0.5736	0.01	0.96	0.03
			0.0110	0.4264	0.65	0.22	0.13
	-1	-0.08	-1.1319	0.5185	0.00	0.95	0.05
			0.1980	0.4815	0.56	0.25	0.20
	-3	0.32	-0.9528	0.4088	0.00	0.92	0.08
			0.5817	0.5912	0.40	0.26	0.34
	-5	0.72	-0.8039	0.3156	0.00	0.89	0.10
			0.9756	0.6844	0.28	0.23	0.49

Table 5.2: c-optimal designs for the positive-negative extreme model for $r = 3, 4$.

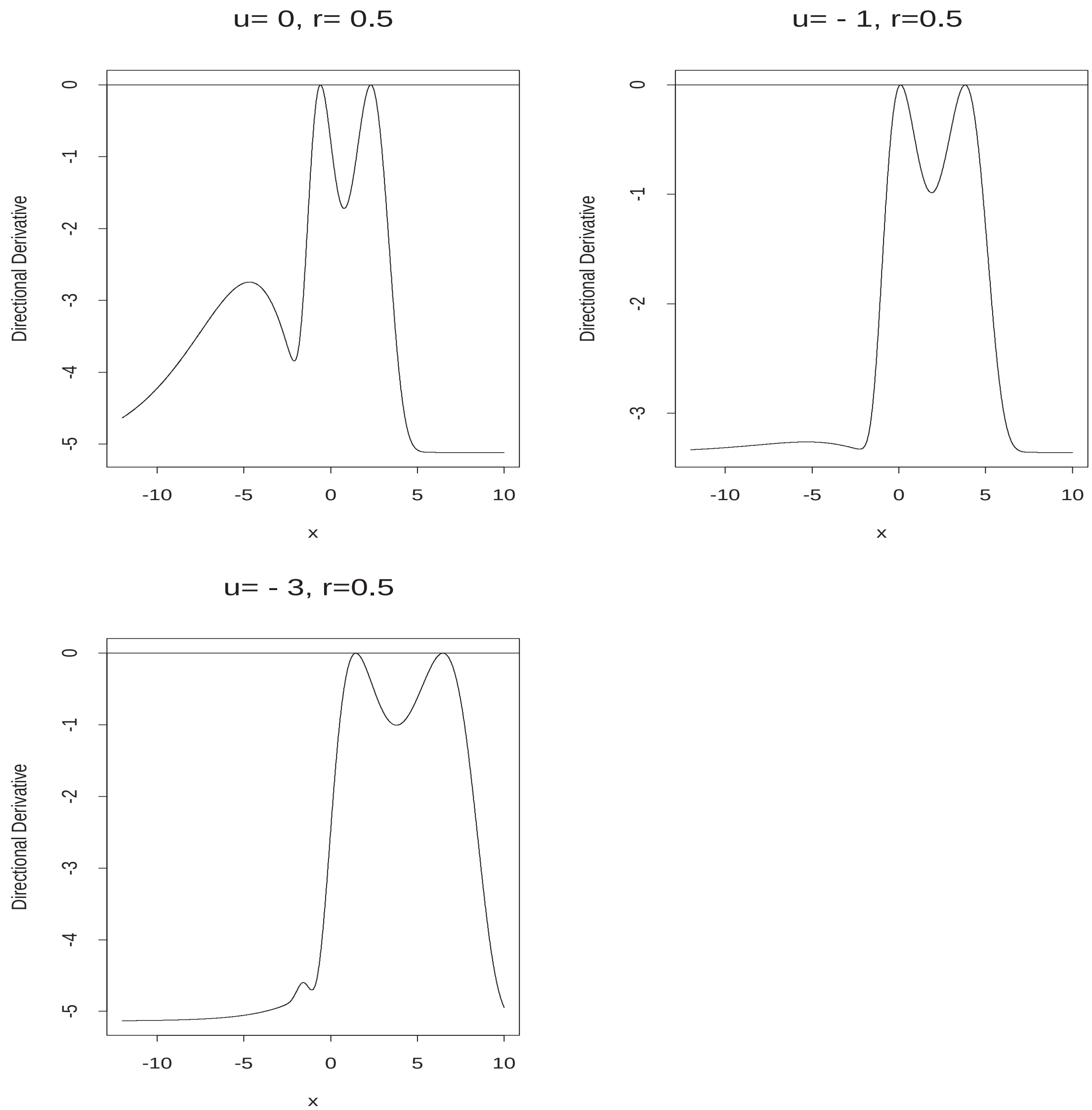

Figure 5.1: Directional derivatives c-optimal design for the positive-negative extreme model for ($r = 0.5$.)

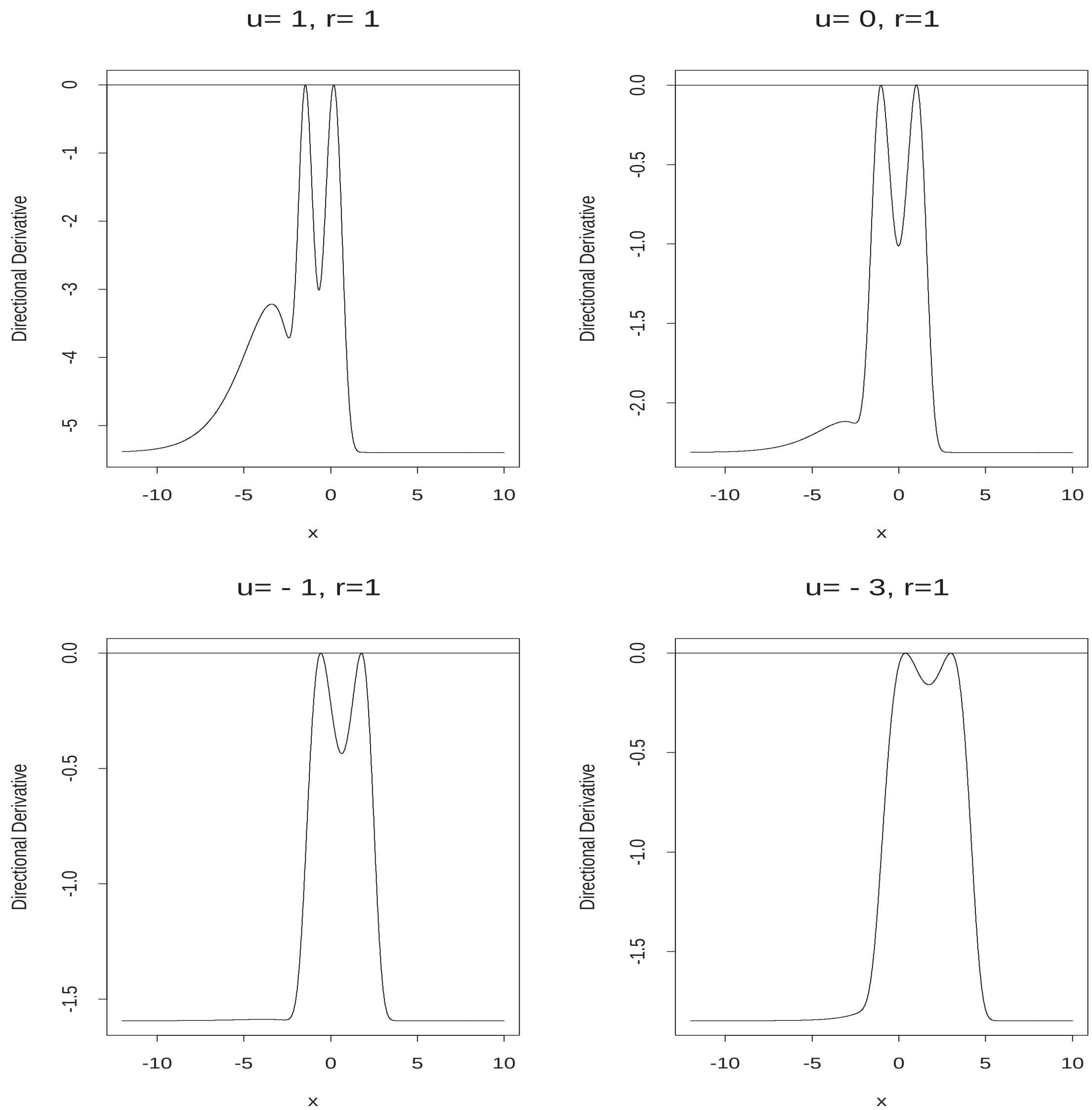

Figure 5.2: Directional derivative for c-optimal design for the positive-negative extreme model for selected values of μ and $r = 1$.

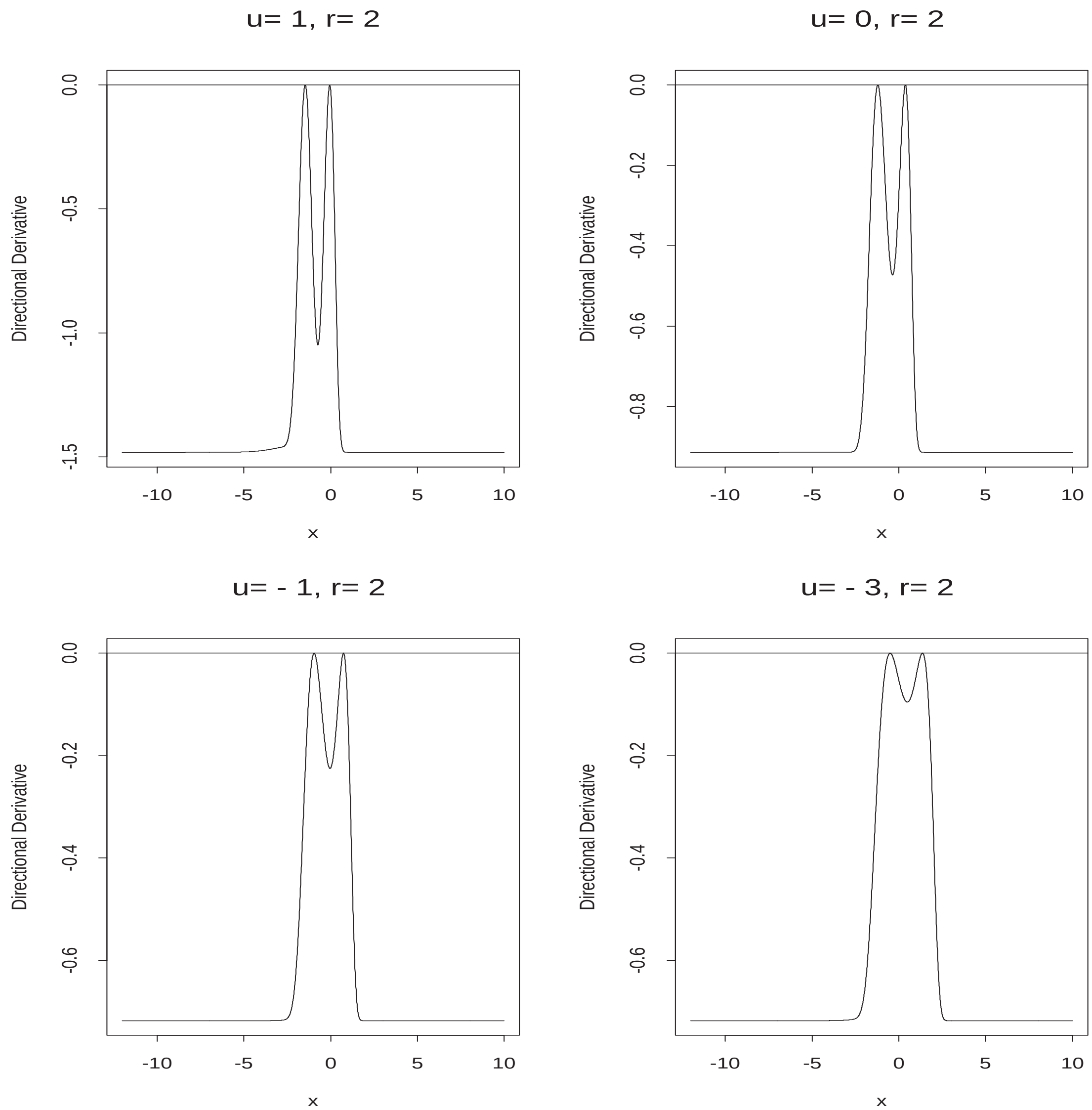

Figure 5.3: Directional derivative for the positive-negative extreme model for c-optimal design for selected values of μ and $r = 2$.

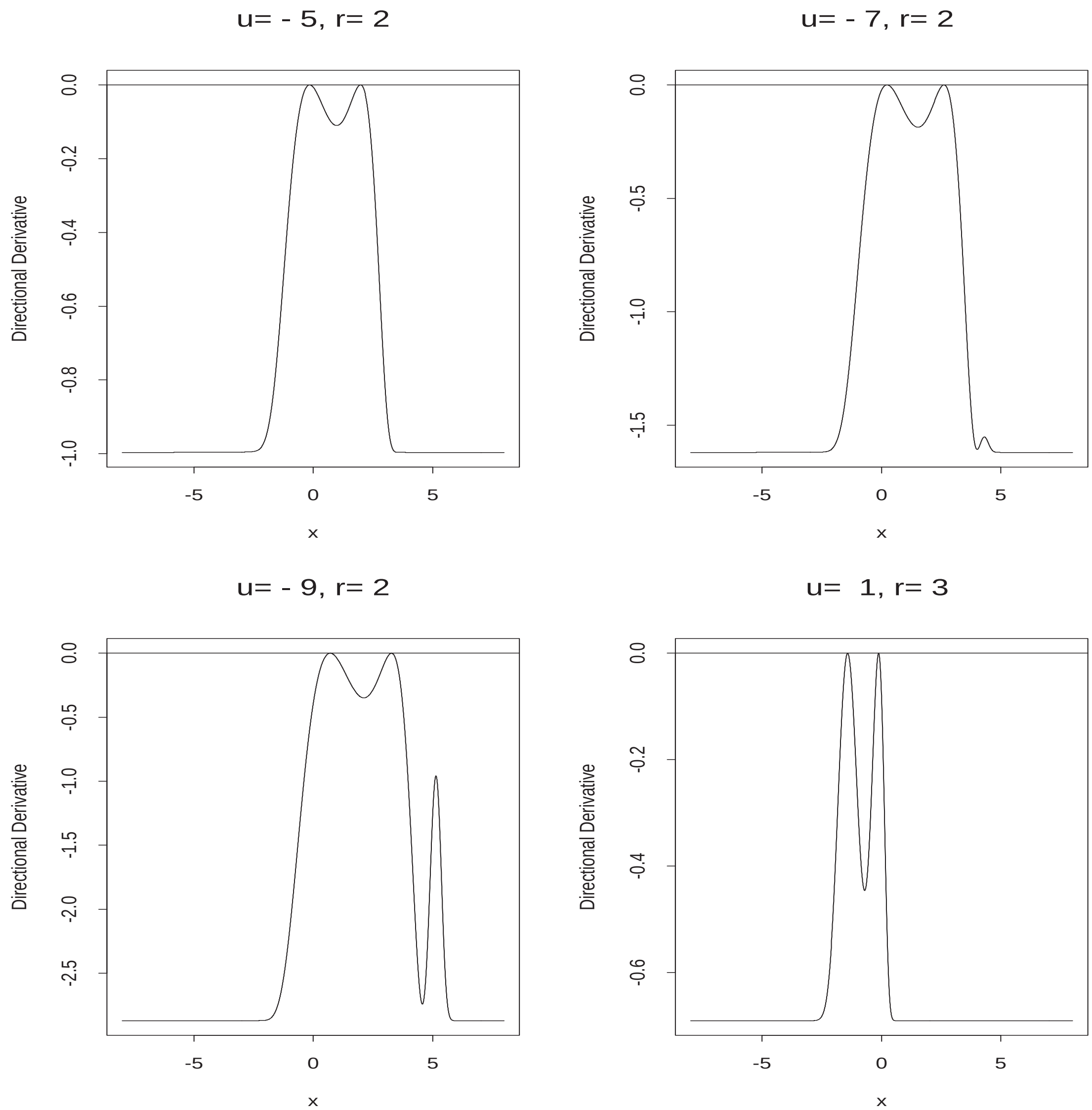

Figure 5.4: Directional derivative for the positive-negative extreme model for c-optimal design for selected values of μ and $r = 2, 3$.

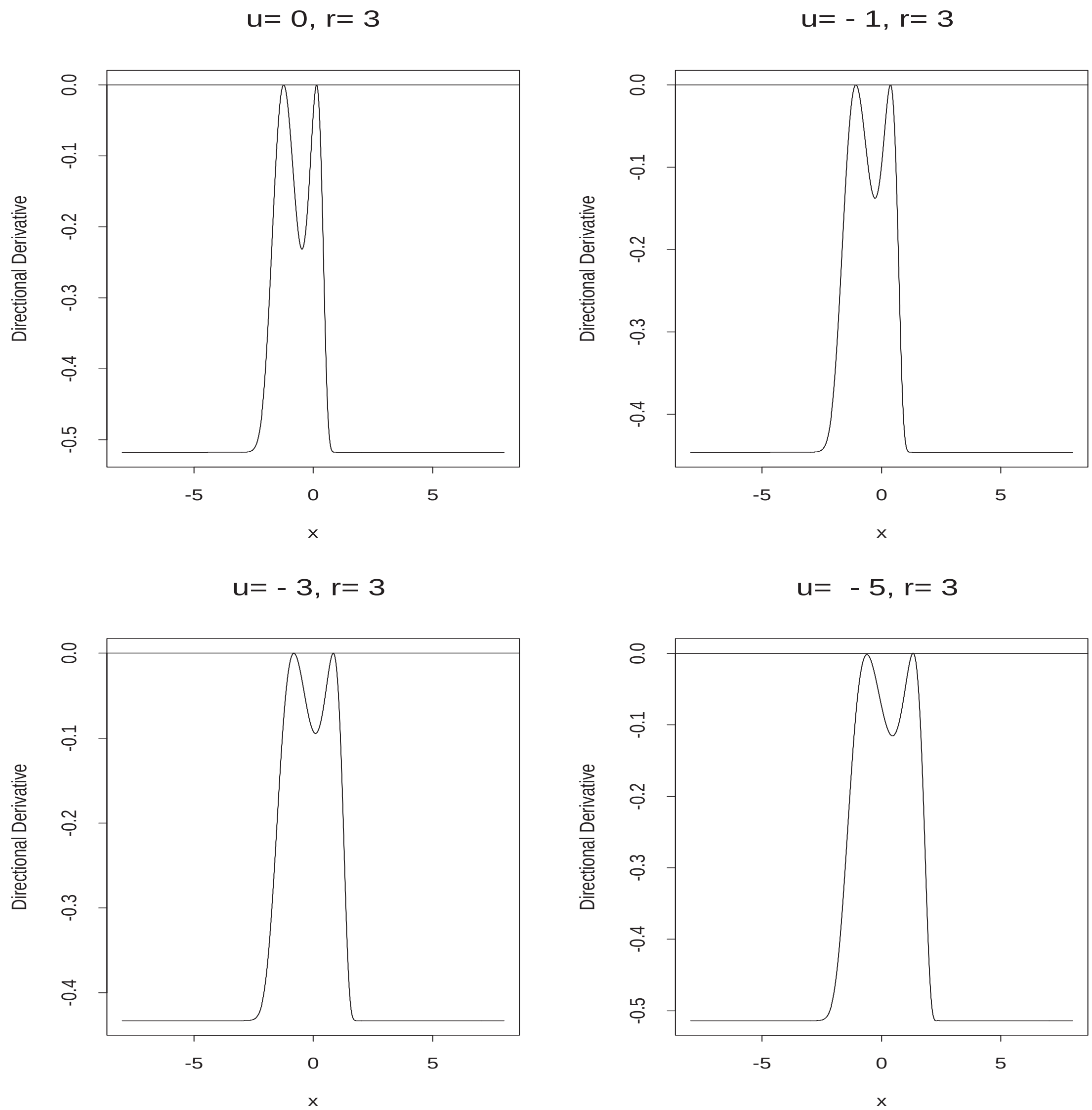

Figure 5.5: Directional derivative for the positive-negative extreme model for c-optimal design for selected values of μ and $r = 3$.

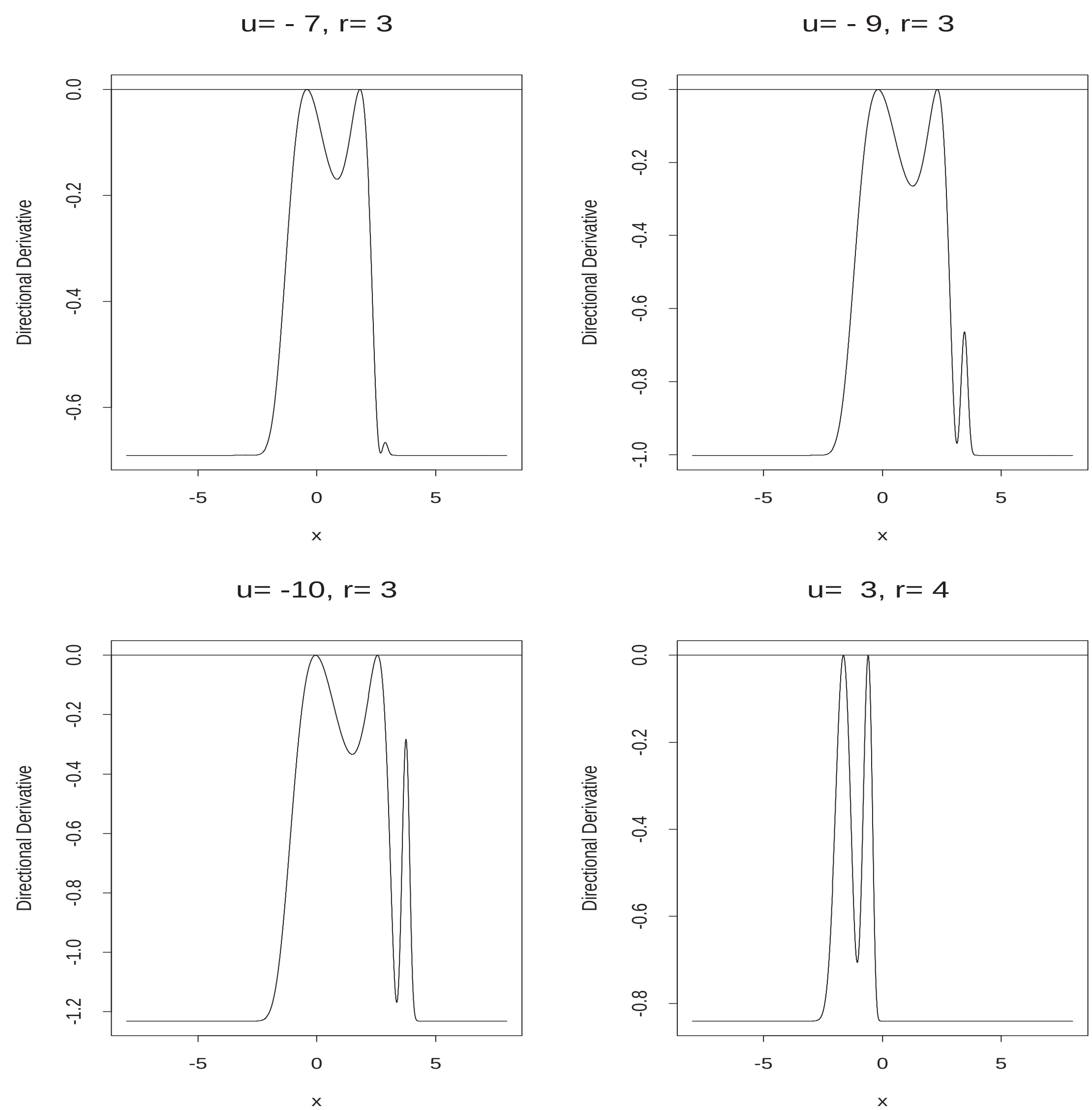

Figure 5.6: Directional derivative for c-optimal design for selected values of μ and $r = 3, 4$.

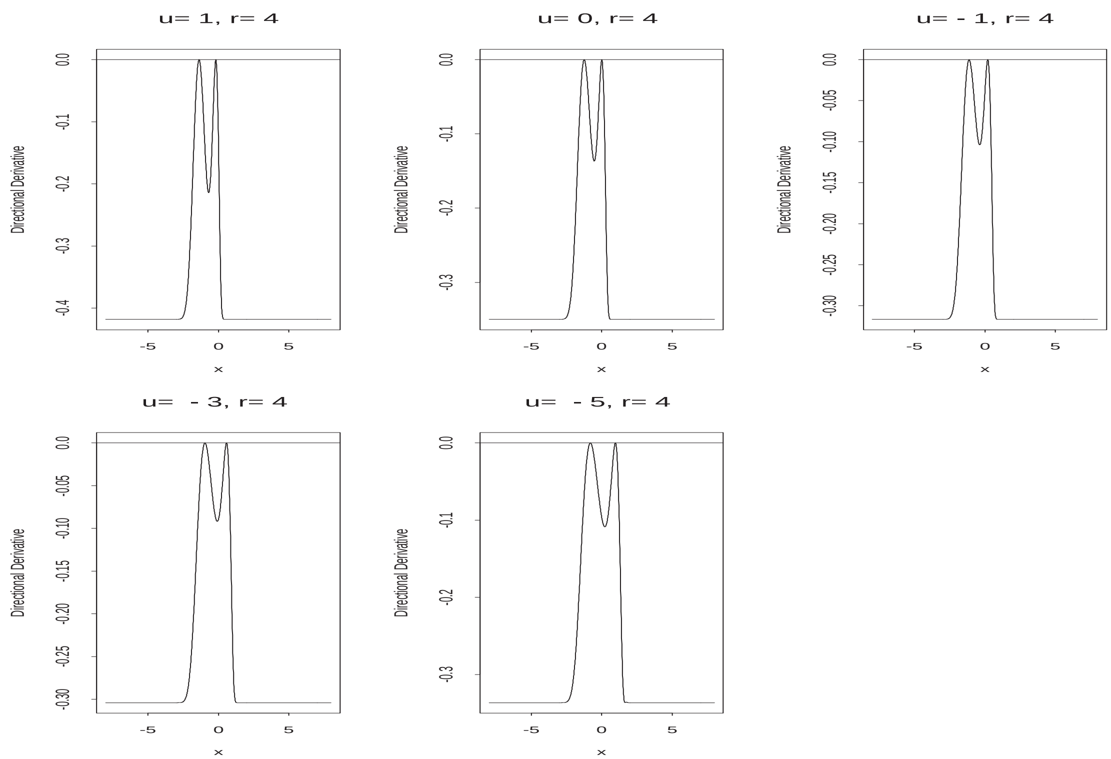

Figure 5.7: Directional derivative for the positive-negative extreme model for c-optimal design for selected values of μ and $r = 4$.

μ	$g(\mathbf{\Theta})$	x	Weights	$F(x)$	$\bar{F}(x)\bar{G}(x)$	$H(x)$
3	-1.50	-4.2617	0.1729	0.25	0.75	0.00
		-1.6981	0.9827	0.98	0.03	0.00
-1	0.50	-0.1399	0.6299	0.27	0.50	0.23
		1.5914	0.3702	0.84	0.03	0.13
-3	1.50	-0.3822	0.5162	0.033	0.74	0.22
		3.514	0.4833	0.81	0.01	0.18
-5	2.50	-0.4489	0.5024	0.00	0.79	0.21
		5.4782	0.4977	0.80	0.00	0.20
-8	4.0	-0.4647	0.5001	0.00	0.80	0.20
		8.4670	0.4999	0.80	0.00	0.20
-10	5.00	-0.4659	0.5001	0.00	0.80	0.20
		10.4663	0.5000	0.80	0.00	0.20
-12	6.0	-0.4600	0.5000	0.00	0.80	0.21
		12.4600	0.5000	0.80	0.00	0.21
-15	7.500	-0.4600	0.5000	0.00	0.80	0.21
		15.4600	0.5000	0.80	0.00	0.21

Table 5.3: c-optimal designs for the positive-negative extreme model with equal slope.

close to $(-0.465, 0.465-\mu)$ with equal weights. These points are approximately the 20^{th} percentiles for G_x and $F_{\mu+rx}$. We conjecture in Section 6.2 that the limiting locally c-optimal design consists of two equal weighted points at $(-0.465, 0.465-\mu)$. We have not been able to prove this; for details see Section 6.4.

Theorem 5.2.2 gives canonical c-optimal designs for $\boldsymbol{\theta}_0 = (0, 1, \mu) = (0, 1, 0)$ and $(0, 1, 1)$. These c-optimal designs have only one point.

Theorem 5.2.2 *For* $\boldsymbol{\theta}$= $(\alpha_2, \beta, \mu) = (0, 1, 0)$ *and* $(0, 1, 1)$*, the c-optimal design is an one point design putting mass 1 at the optimal dose* $x = 0$ *and* $x = -1/2$*, respectively.*

Proof: For these parameters, the information matrix is singular. Because ϕ_c is concave, we can verify that the design is c-optimal by finding a matrix $\mathbf{H}$ such

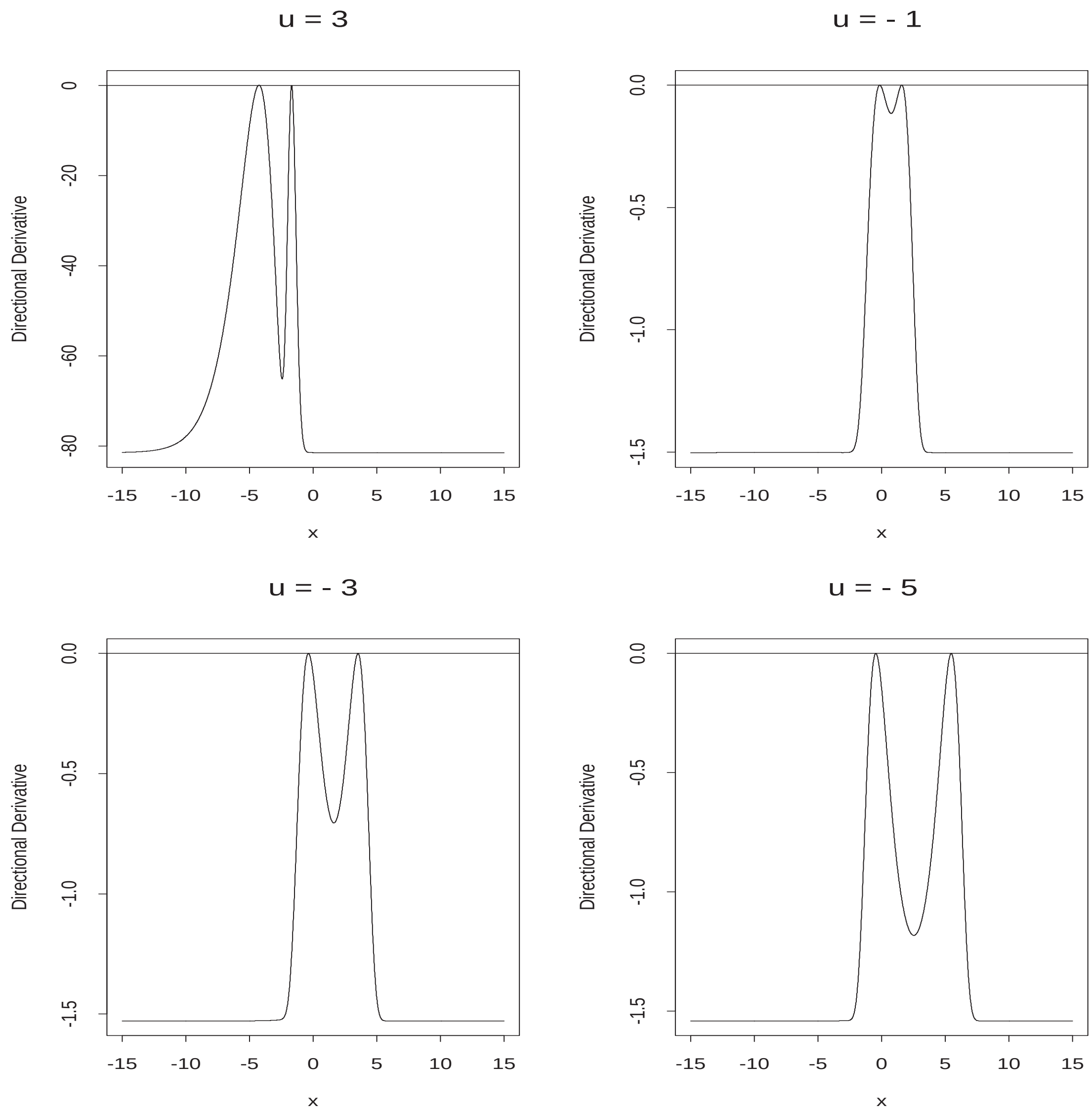

Figure 5.8: Directional derivative for the positive-negative extreme model for c-optimal design for selected values of μ.

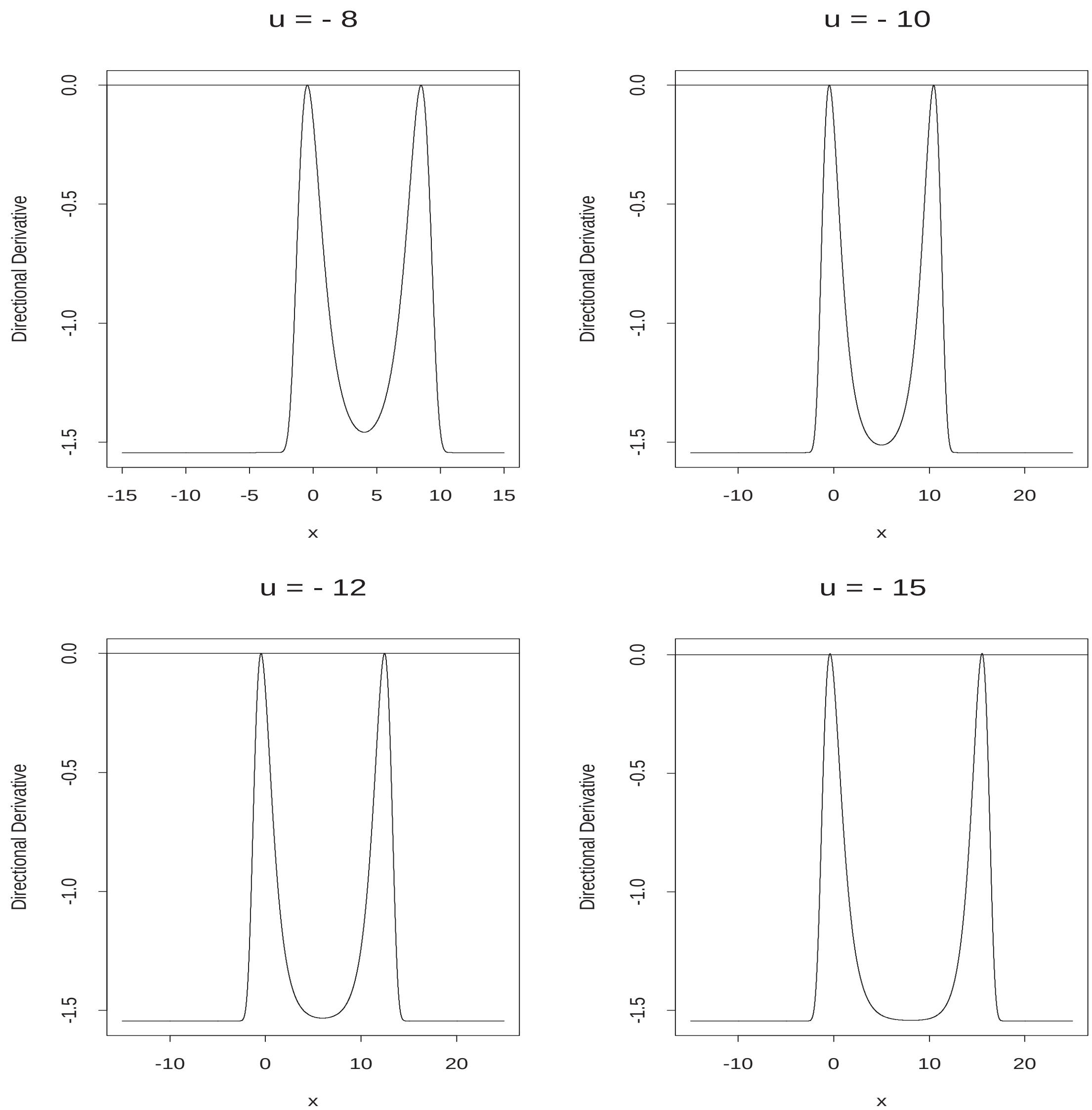

Figure 5.9: Continued directional derivative for the positive-negative extreme model for c-optimal design for selected values of μ.

that $(\mathbf{M}+\mathbf{H}\mathbf{H}^{\mathbf{T}})^{-1}$ exists and is a generalized inverse of $\mathbf{M} = \mathbf{M}(\boldsymbol{\xi},\boldsymbol{\theta})$ such that $F_\phi(\mathbf{M}+\mathbf{H}\mathbf{H}^T, \mathbf{I}(x,\boldsymbol{\theta})) \leq 0$ for all x and the Sup $F_\phi(\mathbf{M}+\mathbf{H}\mathbf{H}^{\mathbf{T}}, \mathbf{I}(x,\boldsymbol{\theta})) = 0$; F_ϕ is the directional derivative of ϕ as defined in [36]. The Fisher's information when $\boldsymbol{\theta} = (\alpha_2, \beta, \mu) = (0,1,0)$ is

$$\mathbf{M}(\boldsymbol{\xi},\boldsymbol{\theta}) = \begin{pmatrix} 0.5819767 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0.2140973 \end{pmatrix}.$$

Take $\mathbf{H} = \begin{pmatrix} \frac{100}{59\sqrt{3}} & \frac{-2}{\sqrt{3}} & 0 \end{pmatrix}^T$. Then

$$\mathbf{G} = (\mathbf{M}+\mathbf{H}\mathbf{H}^{\mathbf{T}})^{-1} = \begin{pmatrix} 1.718282 & 1.456171 & 0 \\ 1.456171 & 1.984043 & 0 \\ 0 & 0 & 4.670774 \end{pmatrix}.$$

Fisher's information when $\boldsymbol{\theta} = (\alpha_2, \beta, \mu) = (0,1,1)$ is

$$\mathbf{M}(\boldsymbol{\xi},\boldsymbol{\theta}) = \begin{pmatrix} 0.6471598 & -0.3235799 & 0 \\ -0.3235799 & 0.1929014 & -0.062223 \\ 0 & -0.062223 & 0.124446 \end{pmatrix}.$$

Let $\mathbf{H} = \begin{pmatrix} \frac{100}{49\sqrt{3}} & \frac{-2}{\sqrt{3}} & 0 \end{pmatrix}^T$. Then

$$\mathbf{G} = (\mathbf{M}+\mathbf{H}\mathbf{H}^{\mathbf{T}})^{-1} = \begin{pmatrix} 7.222648 & 8.135674 & 4.067837 \\ 8.135674 & 9.832958 & 4.916479 \\ 4.067837 & 4.916479 & 10.493853 \end{pmatrix}.$$

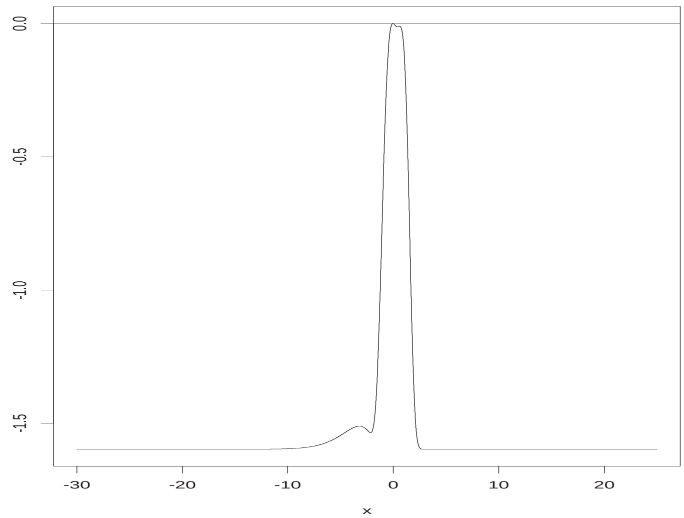

Figure 5.10: Directional derivative for Theorem 5 ($\mu = 0$)

Figures 5.10 and 5.11 show that the directional derivatives were found to be non-positive and the Sup $F_\phi(\mathbf{M} + \mathbf{H}\mathbf{H}^\mathbf{T}, \mathbf{I}(x, \boldsymbol{\theta})) = 0$. □

5.3 Efficiency of Locally c-optimal Designs

In this section we investigate how efficient the c-optimal designs are. We illustrate one example when the slopes are equal and the location parameter is $\mu = -3$. The optimal design is $\boldsymbol{\xi} = \begin{pmatrix} -0.3822 & 3.514 \\ 0.5162 & 0.4838 \end{pmatrix}$. Figure 5.12 shows the efficiency of a design when $x_1 = -0.3822$ is fixed and x_2 varies away from 3.514 and the success function H. By using a point less than the optimal point 3.514 and keeping the same

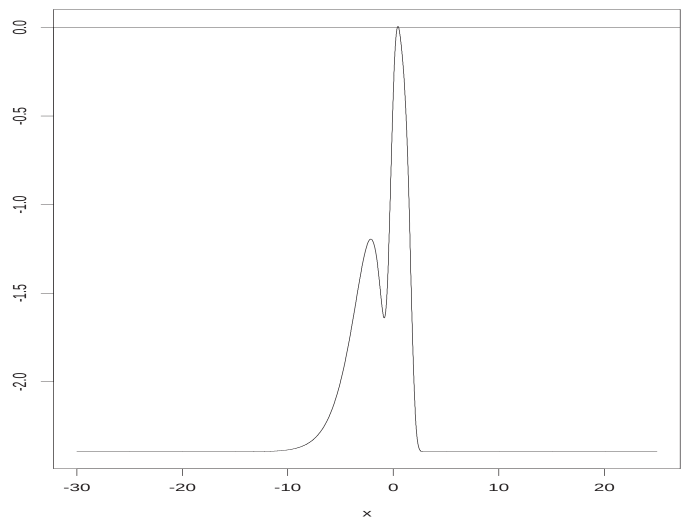

Figure 5.11: Directional derivative for Theorem 5 ($\mu = 1$)

weights, we see the that the efficiency decreases with the distance from 3.514 while the probability of success increases to reach 0.6. Using the optimal dose, i.e., $x_2 = 2.004$, reduces efficiency of parameters estimators to about 0.41 % of the optimal design.

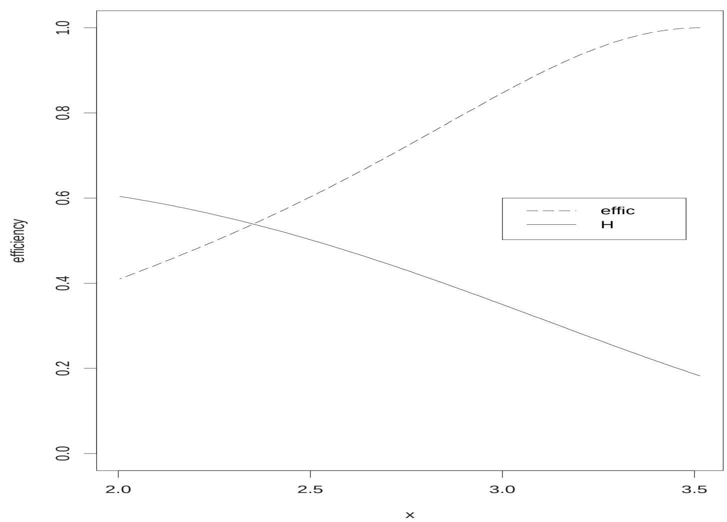

Figure 5.12: Efficiency Plot for c-optimal design when $(\mu = -3, r = 1.)$

Chapter 6

Limiting Locally Optimal Designs

Recall from Chapter 4 and 5 that optimal designs for the contingent response model were found numerically because there is no closed form for these optimal designs. It can be seen from Table 4.1 that when the two response functions $F_{\mu+rx}$ and $\bar{G}_x$ are quite separate (see Figure 1.1) the optimal designs consist of four approximately equally weighted design points. These design points are the optimal designs for $F_{\mu+rx}$ and $\bar{G}_x$ separately. These optimal designs can be expressed as solutions to closed form equations and used as an approximation for the optimal designs when the two functions $F_{\mu+rx}$ and $\bar{G}_x$ are not quite separate. These approximate optimal designs are called the *limiting optimal designs*. The *limiting optimal designs* were introduced by Fan and Chaloner [12].

Limiting optimal designs for the contingent response model often can be surmised from the knowledge of the optimal designs for each individual response $F_{\mu+rx}$ and G_x. Such a guess can be used as starting values in numerical searches for the optimal designs and as initial treatments in a sequential designs that approximate the optimal ones. Fan and Chaloner [12] studied the logistic-logistic model and found the limiting optimal designs for this model. They found these limiting optimal designs to be very efficient approximations of the true optimal designs for the models they studied. This suggests one can use these designs in place of the optimal design.

We study the limiting optimal designs for the canonical positive-negative extreme value model. We start by defining the limiting optimal designs as given in Fan and Chaloner [12].

Definition 6.0.1 *For a concave criterion ϕ on a set of design measures H, a sequence of designs, $\{\eta_i, i \in R\}$, is called a sequence of limiting ϕ-optimal designs for a sequence of prior distributions, $\{\pi_i, i \in R,\}$ if*

$$Sup_{\eta \in H} F_\phi(\eta_i, \eta, \pi_i) \to 0 \; as \; i \to \infty.$$

A design of this sequence, η_i, is called a limiting ϕ-optimal design for π_i.

The index i in the definition serves as $-\mu$ in the canonical model. That is, we find the limiting optimal design for a sequence of designs as $-\mu$ goes to ∞.

6.1 Limiting Locally D-optimal Designs for Models with Unequal Slopes

Recall that for the canonical positive-negative extreme model with $\beta_1 \neq \beta_2$ and $\boldsymbol{\theta} = (0, 1, \mu, r)$ that a candidate limiting locally D-optimal design was found in

Section 4.1 on page 20 to consist of four equally weighted points at $[x_1 = 1.3377, x_2 = -0.9796, (-x_1 - \mu)/r, (-x_2 - \mu)/r]$. These are the optimal design points for $F_{\mu+rx}$ and $\bar{G}_x$ concatenated. Define

$v(t) = \exp(2\mu + 2rt)\exp(-\exp(\mu + rt))/(1 - \exp(-\exp(\mu + rt)))$

and

$w(t) = \exp(-\exp(\mu + rt))\exp(-2rt)\exp(-\exp(-rt))/(1 - \exp(-\exp(-rt)))$.

In Theorem 6.1.1, we prove that the directional derivative of the locally limiting D-optimal designs for the positive-negative extreme model is asymptotically less than or equal to the sum of the two directional derivative functions for the locally D-optimal designs of the positive and negative extreme models when each is considered separately. To prove Theorem 6.1.1, we use Lemma 6.1.1 from Fan and Chaloner [13]. We will use $\boldsymbol{\xi}_*$ to represent the limiting optimal design, where $\boldsymbol{\xi}^*$ is the optimal design.

Lemma 6.1.1 *Consider two symmetric* 2×2 *nonsingular matrices* $\mathbf{P}, \mathbf{Q}$ *and suppose that* $\mathbf{P} = \mathbf{Q} + \mathbf{R}$. *Let* q_{ij} *and* r_{ij} *be the* $ij^{th} i, j = 1, 2$ *elements of* $\mathbf{Q}$ *and* $\boldsymbol{R}$, *respectively. Then for any symmetric* 2×2 *matrix* $\boldsymbol{I} = \begin{pmatrix} i_{11} & i_{12} \\ i_{21} & i_{22} \end{pmatrix}$,

$$\begin{aligned} Tr(\boldsymbol{I}(\boldsymbol{P}^{-1} - \boldsymbol{Q}^{-1})) &= i_{22}(\frac{-q_{11}}{\det \boldsymbol{Q}} + \frac{q_{11} + r_{11}}{\det \boldsymbol{Q} + q_{22}r_{11} - 2q_{12}r_{12} - r_{12}^2 + q_{11}r_{22} + r_{11}r_{22}}) \\ &+ 2i_{12}(\frac{q_{12}}{\det \boldsymbol{Q}} + \frac{-q_{12} - r_{12}}{\det \boldsymbol{Q} + q_{22}r_{11} - 2q_{12}r_{12} - r_{12}^2 + q_{11}r_{22} + r_{11}r_{22}}) \\ &+ i_{11}(\frac{-q_{22}}{\det \boldsymbol{Q}} + \frac{q_{22} + r_{22}}{\det \boldsymbol{Q} + q_{22}r_{11} - 2q_{12}r_{12} - r_{12}^2 + q_{11}r_{22} + r_{11}r_{22}}). \end{aligned}$$

Theorem 6.1.1 *The locally limiting D-optimal Design is given by*

$$\boldsymbol{\xi}_* = \begin{pmatrix} (-0.9796 & 1.3377 & (0.9796-\mu)/r & (-1.3377-\mu)/r) \\ 0.25 & 0.25 & 0.25 & 0.25 \end{pmatrix}.$$

Proof: The Fisher's information (See Lemma 3.1 on page 16)at $\boldsymbol{\xi}_*$ is given by

$$\mathbf{M}(\boldsymbol{\xi}_*, \boldsymbol{\theta}) = 0.25 \begin{pmatrix} \mathbf{M_A}(\boldsymbol{\xi}_*, \boldsymbol{\theta}) & \mathbf{0} \\ \mathbf{0} & \mathbf{M_B}(\boldsymbol{\xi}_*, \boldsymbol{\theta}) \end{pmatrix},$$

that is,

$$\mathbf{M_A}(\boldsymbol{\xi}_*, \boldsymbol{\theta}) = \Sigma_i^2 v(x_i) \begin{pmatrix} 1 & x_i \\ x_i & x_i^2 \end{pmatrix} + v\left(\frac{-x_i-\mu}{r}\right) \begin{pmatrix} 1 & \frac{-x_i-\mu}{r} \\ \frac{-x_i-\mu}{r} & \frac{(-x_i-\mu)^2}{r^2} \end{pmatrix}$$

and

$$\mathbf{M_B}(\boldsymbol{\xi}_*, \boldsymbol{\theta})) = \Sigma_i^2 w(x_i) \begin{pmatrix} 1 & x_i \\ x_i & x_i^2 \end{pmatrix} + w\left(\frac{-x_i-\mu}{r}\right) \begin{pmatrix} 1 & \frac{-x_i-\mu}{r} \\ \frac{-x_i-\mu}{r} & \frac{(-x_i-\mu)^2}{r^2} \end{pmatrix}$$

Rewrite the sub-matrices of Fisher's Information in terms of the optimal design for the component models separately :

$$\mathbf{M_A}(\boldsymbol{\xi}_*, \boldsymbol{\theta}) = 0.5\mathbf{M_A}^* + \mathbf{R_{A^*}},$$

$$\mathbf{M_B}(\boldsymbol{\xi}_*, \boldsymbol{\theta}) = 0.5\mathbf{M_B}^* + \mathbf{R_{B^*}}'$$

$\mathbf{M_{A^*}}$is Fishers information for a single extreme value function $F(\mu + rx)$ evaluated at the optimal points $\frac{-x_1-\mu}{r}$ and $\frac{-x_2-\mu}{r}$, where

$$\mathbf{M_{A^*}} = \Sigma_i^2 0.5v\left(\frac{-x_i-\mu}{r}\right) \begin{pmatrix} 1 & \frac{-x_i-\mu}{r} \\ \frac{-x_i-\mu}{r} & \frac{(-x_i-\mu)^2}{r^2} \end{pmatrix}$$

and

$$\mathbf{R}_{\mathbf{A}_*} = \Sigma_i^2 0.25 v(x_i) \begin{pmatrix} 1 & x_i \\ x_i & x_i^2 \end{pmatrix};$$

$\mathbf{M}_{\mathbf{B}^*}$ is Fishers information for a single extreme value function $\bar{G(x)}$ evaluated at the optimal points x_1 and x_2, where

$$\mathbf{M}_{\mathbf{B}^*} = \Sigma_i^2 \frac{0.5 w(x_i)}{\exp(-\exp(\mu + r x_i))} \begin{pmatrix} 1 & x_i \\ x_i & x_i^2 \end{pmatrix},$$

and

$$\begin{aligned} \mathbf{R}_{\mathbf{B}_*} &= \Sigma_i^2 0.25 w(x_i)(1 - \frac{1}{\exp(-\exp(\mu + r x_i))} \begin{pmatrix} 1 & x_i \\ x_i & x_i^2 \end{pmatrix} \\ &\quad + \Sigma_i^2 0.25 w\left(\frac{-x_i - \mu}{r}\right) \begin{pmatrix} 1 & \frac{-x_i-\mu}{r} \\ \frac{-x_i-\mu}{r} & \frac{(-x_i-\mu)^2}{r^2} \end{pmatrix}. \end{aligned}$$

Denote Fisher's information for a single point x from $F(\mu + rx)$ and $\bar{G}(x)$ by $\mathbf{I}_F$ and $\mathbf{I}_G$, respectively. Note that in Lemma 3.1.1 on page 16, $\mathbf{A}(x, \boldsymbol{\theta}) = \mathbf{I}_F$ and $\mathbf{B}(x,\boldsymbol{\theta})= [1 - F(\mu + rx)]\mathbf{I}_G$.

Now as $-\mu$ approaches ∞ or μ approaches $-\infty$, $v(x_i)$ and $w(\frac{-x_i-\mu}{r})$, $i = 1, 2$ are close to zero and hence so is any polynomial of μ multiplied by $(v(x_i))$ and $w(\frac{-x_i-\mu}{r})$. Therefore, $\mathbf{R}_{\mathbf{A}_*}$ and $\mathbf{R}_{\mathbf{B}_*}$ go to zero, and hence $\mathbf{M}_{\mathbf{A}}{}^{-1}(\boldsymbol{\xi}_*, \boldsymbol{\theta})$ and $\mathbf{M}_{\mathbf{B}}{}^{-1}(\boldsymbol{\xi}_*, \boldsymbol{\theta})$ can be approximated by $2\mathbf{M}_{\mathbf{A}^*}^{-1}$ and $2\mathbf{M}_{\mathbf{B}^*}^{-1}$, respectively. The directional derivative for the locally limiting D-optimal design ($\boldsymbol{\xi}_*$) is

$$\begin{aligned}
F_D &= Trace(\mathbf{I}(x,\boldsymbol{\theta}))\mathbf{M}^{-1}(\boldsymbol{\xi}_*,\boldsymbol{\theta}) - 4 \\
&= Trace(\mathbf{A}(x,\boldsymbol{\theta})\mathbf{M}_{\mathbf{A}}^{-1}(\boldsymbol{\xi}_*,\boldsymbol{\theta})) + Trace(\mathbf{B}(x,\boldsymbol{\theta})\ \mathbf{M}_{\mathbf{B}}^{-1}(\boldsymbol{\xi}_*,\boldsymbol{\theta})) - 4
\end{aligned}$$

which can be rewritten and bounded as

$$\begin{aligned}
F_D \quad &= \quad Tr(\mathbf{A}(x,\boldsymbol{\theta})\mathbf{M}_{\mathbf{A}}^{-1}(\boldsymbol{\xi}_*,\boldsymbol{\theta})) + Tr(\mathbf{B}(x,\boldsymbol{\theta}){\mathbf{M}_{\mathbf{B}}}^{-1}(\boldsymbol{\xi}_*,\boldsymbol{\theta})) \\
&+ \quad 2Tr\mathbf{I}_F\mathbf{M}_{\mathbf{A}^*}^{-1} - 2Tr\mathbf{I}_F\mathbf{M}_{\mathbf{A}^*}^{-1} + 2Tr\mathbf{I}_G\mathbf{M}_{\mathbf{B}^*}^{-1} - 2Tr\mathbf{I}_G\mathbf{M}_{\mathbf{B}^*}^{-1} - 4 \\
&= \quad Tr(\mathbf{A}(x,\boldsymbol{\theta})\mathbf{M}_{\mathbf{A}}^{-1}(\boldsymbol{\xi}_*,\boldsymbol{\theta})) + Tr(\mathbf{B}(x,\boldsymbol{\theta})\ \mathbf{M}_{\mathbf{B}}^{-1}(\boldsymbol{\xi}_*,\boldsymbol{\theta})) + 2Tr\mathbf{I}_F\mathbf{M}_{\mathbf{A}^*}^{-1} \\
&- \quad 2Tr\mathbf{I}_F\mathbf{M}_{\mathbf{A}^*}^{-1} + 2Tr\mathbf{I}_G\mathbf{M}_{B^*}^{-1} \\
&\quad -2Tr(\mathbf{B}(x,\boldsymbol{\theta}) + [1-\exp(-\exp(\mu+rx))]\mathbf{I}_G)\mathbf{M}_{\mathbf{B}^*}^{-1} - 4 \\
&< \quad Tr(\mathbf{A}(x,\boldsymbol{\theta})\mathbf{M}_{\mathbf{A}}^{-1}(\boldsymbol{\xi}_*,\boldsymbol{\theta})) + Tr(\mathbf{B}(x,\boldsymbol{\theta})\mathbf{M}_{\mathbf{B}}^{-1}(\boldsymbol{\xi}_*,\boldsymbol{\theta})) + 2Tr\mathbf{I}_F\mathbf{M}_{\mathbf{A}^*}^{-1} \\
&- \quad 2Tr\mathbf{I}_F\mathbf{M}_{\mathbf{A}^*}^{-1} + 2Tr\mathbf{I}_G\mathbf{M}_{\mathbf{B}^*}^{-1} - 2Tr(\mathbf{B}(x,\boldsymbol{\theta}))\mathbf{M}_{\mathbf{B}^*}^{-1} - 4 \\
&= \quad Tr(\mathbf{A}(x,\boldsymbol{\theta})\mathbf{M}_{\mathbf{A}}^{-1}(\boldsymbol{\xi}_*,\boldsymbol{\theta})) - 2Tr\mathbf{I}_F\mathbf{M}_{\mathbf{A}^*}^{-1} + Tr(\mathbf{B}(x,\boldsymbol{\theta})\mathbf{M}_{\mathbf{B}}^{-1}(\boldsymbol{\xi}_*,\boldsymbol{\theta})) \\
&- \quad 2Tr(\mathbf{B}(x,\boldsymbol{\theta}))\mathbf{M}_{\mathbf{B}^*}^{-1} + 2Tr\mathbf{I}_F\mathbf{M}_{\mathbf{A}^*}^{-1} + 2Tr\mathbf{I}_G\mathbf{M}_{\mathbf{B}^*}^{-1} - 4 \\
&= \quad Tr\mathbf{I}_F\mathbf{M}_{\mathbf{A}^*}^{-1} + Tr\mathbf{I}_G\mathbf{M}_{\mathbf{B}^*}^{-1} - 4 + Tr\mathbf{I}_F\mathbf{M}_{\mathbf{A}^*}^{-1} + Tr\mathbf{I}_G\mathbf{M}_{\mathbf{B}^*}^{-1} \\
&+ \quad Tr(\mathbf{A}(x,\boldsymbol{\theta})(\mathbf{M}_{\mathbf{A}}^{-1}(\boldsymbol{\xi}_*,\boldsymbol{\theta}) - 2\mathbf{M}_{\mathbf{A}^*}^{-1})) + Tr(\mathbf{B}(x,\boldsymbol{\theta})(\mathbf{M}_{\mathbf{B}}^{-1}(\boldsymbol{\xi}_*,\boldsymbol{\theta}) - 2\mathbf{M}_{\mathbf{B}^*}^{-1})).
\end{aligned} \tag{6.1}$$

We now show that the last two terms in (6.1) go to zero as $-\mu$ goes to ∞ using Lemma 6.1.1 and the remaining terms are non-positive with supremum zero. That is, the directional derivative asymptotically is equal to

$$F_D = 2Tr\mathbf{I}_F\mathbf{M}_{\mathbf{A}^*}^{-1} + 2Tr\mathbf{I}_G\mathbf{M}_{\mathbf{B}^*}^{-1} - 4 \leq 0,$$

which will complete the proof.

The first three terms in (6.1) are the sum of the directional derivatives evaluated at the optimal design points for $F_{\mu+rx}$ and $\bar{G}_x$ separately. To understand this consider the single negative extreme model $F_{\mu+rx}$ with location-scale parameters μ and r. $\mathbf{M}_{\mathbf{A}_*}$ is the Fisher's information for the local D-optimal design for this model. By the General Equivalence Theorem, the $Tr\mathbf{I}_F\mathbf{M}_{\mathbf{A}_*}^{-1} - 2$ is the directional derivative of the locally D-optimal design for this model and it is non-positive for all possible values of x. Similarly, consider the single positive extreme model $\bar{G}_x$ with location-scale parameters 0 and 1. $\mathbf{M}_{\mathbf{B}_*}$ is the Fisher's information for the locally D-optimal design for this model. By the General Equivalence Theorem, $Tr\mathbf{I}_G\mathbf{M}_{\mathbf{B}_*}^{-1} - 2$ is the directional derivative of the locally D-optimal design for this model and it is non-positive for all possible values of x.

It can be seen from Figure 6.1a that as $\mu \to -\infty$ $TrI_F\mathbf{M}_{\mathbf{A}_*}^{-1}$ asymptotes toward zero as x decreases from $-\mu/2r$ and $Tr\mathbf{I}_G\mathbf{M}_{\mathbf{B}_*}^{-1}$ asymptotes toward zero as x increases from $-\mu/2r$. Therefore $2Tr\mathbf{I}_F\mathbf{M}_{\mathbf{A}^*}^{-1} + 2Tr\mathbf{I}_G\ \mathbf{M}_{\mathbf{B}^*}^{-1} - 4$ is non-positive on x and asymptotes towards zero as $\mu \to -\infty$. An example is given in Figure 6.1b when $r = 2$ as μ varies over $-10, -15, -25$ the supremum of the directional derivative goes to 0. See Appendix A.3

To evaluate $Tr(\mathbf{B}(x,\boldsymbol{\theta})(\mathbf{M_B}^{-1}(\boldsymbol{\xi}_*,\boldsymbol{\theta}) - 2\mathbf{M}_{\mathbf{B}_*}^{-1})$, let $\mathbf{P} = \mathbf{M_B}(\boldsymbol{\xi}_*,\boldsymbol{\theta}), \mathbf{Q} = 0.5\mathbf{M}_{\mathbf{B}}^*$

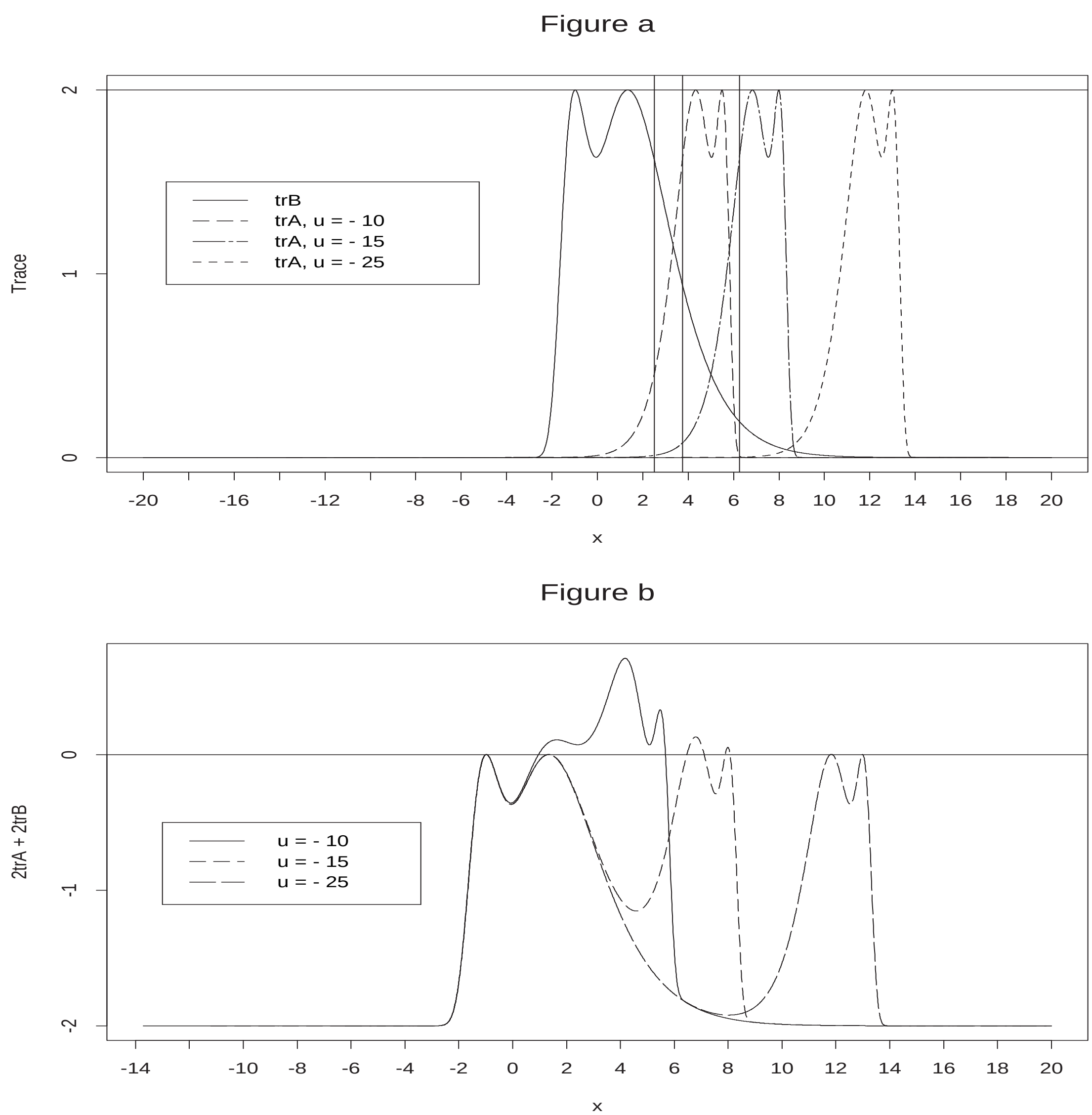

Figure 6.1: (Figure a: $trB = TrI_G M_{B_*}^{-1}$ and $trA = TrI_F M_{A_*}^{-}1$ for $\mu = -10, -15, -25$. Figure b: trA+trB $= 2TrI_F M_{A_*}^{-}1 + 2TrI_G M_{B_*}^{-1} - 4$)

and $\mathbf{I} = \mathbf{B}(x, \boldsymbol{\theta})$ in Lemma 6.1.1. Then $\mathbf{R}$ in the Lemma 6.1.1 equals $\mathbf{R}_{\mathbf{B}_*}$. Recall that $\mathbf{R}_{\mathbf{B}_*}$ goes to a matrix of zeros as μ goes $-\infty$. The elements of $\mathbf{I} = \mathbf{B}(\boldsymbol{\xi}_*, \boldsymbol{\theta})$ are bounded functions and $\mathbf{M}^*_{\mathbf{B}}$ does not depend on μ. This implies that each element of $\mathbf{B}(x, \boldsymbol{\theta})(\mathbf{M}^{-1}_{\mathbf{B}}(\boldsymbol{\xi}_*, \boldsymbol{\theta}) - 2\mathbf{M}^{-1}_{\mathbf{B}_*})$ goes to zero as μ goes to $-\infty$ and so the trace of this matrix goes to zero.

Now to evaluate $Tr(\mathbf{A}(x, \boldsymbol{\theta})(\mathbf{M}^{-1}_{\mathbf{A}}(\boldsymbol{\xi}_*, \boldsymbol{\theta}) - 2\mathbf{M}^{-1}_{A_*})$ let $\mathbf{P} = \mathbf{M}_A(\boldsymbol{\xi}_*, \boldsymbol{\theta}), \mathbf{Q} = 0.5\mathbf{M}^*_{\mathbf{A}}$ and $\mathbf{I} = \mathbf{A}(x, \boldsymbol{\theta})$ in Lemma 6.1.1. Then $\mathbf{R}$ in the Lemma 6.1.1 equals $\mathbf{R}_{\mathbf{A}_*}$. Let $d = q_{22}r_{11} - 2q_{12}r_{12} - r^2_{12} + q_{11}r_{22} + r_{11}r_{22}$. Because $\mathbf{R}_{\mathbf{A}_*}$ goes to a matrix of zeros as μ goes $-\infty$, each element of d, which is a polynomial of μ times an exponential of μ, goes to zero too. The determinant of $\mathbf{Q}$ is independent of μ. Rewrite $Tr(\mathbf{A}(x, \boldsymbol{\theta})(\mathbf{M}^{-1}_{\mathbf{A}}(\boldsymbol{\xi}_*, \boldsymbol{\theta}) - 2\mathbf{M}^{-1}_{\mathbf{A}_*})$ as

$$\begin{aligned} Tr(\mathbf{A}(x, \boldsymbol{\theta})(\mathbf{M}^{-1}_{A}(\boldsymbol{\xi}_*, \boldsymbol{\theta}) - 2\mathbf{M}^{-1}_{\mathbf{A}_*})) &= i_{22}(\frac{-q_{11}d}{\det \mathbf{Q} + (\det \mathbf{Q} + d)} + \frac{r_{11}}{\det \mathbf{Q} + d} \\ &+ 2i_{12}(\frac{q_{12}d}{\det \mathbf{Q} + (\det \mathbf{Q} + d)} + \frac{-r_{12}}{\det \mathbf{Q} \;\; + d} \\ &+ i_{11}(\frac{-q_{22}d}{\det \mathbf{Q} + (\det \mathbf{Q} + d)} + \frac{r_{22}}{+ \det \mathbf{Q} + d}. \end{aligned} \tag{6.2}$$

We will show that each term of the right hand side of (6.2) goes to zero as μ goes to $-\infty$. We show this first for the second term by showing $i_{12}q_{12}d$ and $i_{12}r_{12}$ go to zero as μ goes to $-\infty$. Consider the following:

$$\begin{aligned} 1)\, |i_{12}| &= |xv(x)| = |xv(x) + \frac{\mu}{r}v(x) - \frac{\mu}{r}v(x)| \\ &\leq |\frac{\mu + rx}{r}v(x)| + |\frac{\mu}{r}v(x)| \end{aligned}$$

$$\leq \quad \frac{\grave{K}}{r} - \frac{\mu}{r}K, \tag{6.3}$$

where $\grave{K} = max_x(\mu + rx)v(x)$ and $K = max_x v(x)$.

$$\begin{aligned}
2)\,|i_{12}r_{12}| &= 0.25|i_{12}||x_1 v(x_1) + x_2 v(x_2)| \\
&= 0.25|i_{12}|\exp(2\mu)|x_1 \exp(2rx_1)\frac{\exp(-\exp(\mu + rx_1))}{1 - \exp(-\exp(\mu + rx_1))} \\
&+ x_2 \exp(2rx_2)\frac{\exp(-\exp(\mu + rx_2))}{1 - \exp(-\exp(\mu + rx_2))}| \\
&< 0.25|i_{12}|\exp(2\mu)|x_1 \exp(2rx_1) + x_2 \exp(2rx_2)| \\
&\leq 0.25[\frac{\grave{K}}{r} - \frac{\mu}{r}K]|\exp(2\mu)|x_1 \exp(2rx_1) + x_2 \exp(2rx_2)|.
\end{aligned}$$

The last quantity goes to zero as μ goes to $-\infty$, and therefore, so does $i_{12}r_{12}$. It follows from the same argument that $i_{12}q_{12}d$ goes to zero as μ goes to $-\infty$, since $q_{12}d$, which is a polynomial of μ multiplied by an exponential function of μ, goes to zero as μ goes $-\infty$.

To show that the first term in 6.2 goes to zero as μ goes to $-\infty$, we show that $i_{22}r_{11}$ and $q_{11}d$ each go to zero as μ goes to $-\infty$. Consider

$$\begin{aligned}
1)\,|i_{22}| &= x^2 v(x) = \frac{(\mu + rx)^2}{r^2}v(x) - \frac{\mu^2}{r^2}v(x) - 2\frac{\mu r x v(x)}{r^2} \\
&\leq \frac{(\mu + rx)^2}{r^2}v(x) - 2\frac{\mu r x v(x)}{r^2} \\
&\leq \frac{\grave{\grave{K}}}{r^2} + 2|\frac{\mu}{r}i_{12}| \\
&\leq \frac{\grave{\grave{K}}}{r^2} - 2\frac{\mu}{r}[\frac{\grave{K}}{r} - \frac{\mu}{r}K] \\
&= \frac{1}{r^2}\{\grave{\grave{K}} - 2\mu(\grave{K} - \mu K)\},
\end{aligned}$$

where $\grave{\grave{K}} = max_x(\mu+rx)^2 v(x)$. Hence

$$\begin{aligned}
2)\,|i_{22}r_{11}| &= 0.25|i_{22}||v(x_1)+v(x_2)| \\
&\leq 0.25\frac{1}{r^2}\{\grave{\grave{K}} - 2\mu(\grave{K}-\mu K)\}\exp(\mu) \\
&\quad |\exp(2rx_1)\frac{\exp(-\exp(\mu+rx_1))}{1-\exp(-\exp(\mu+rx_1))} \\
&\quad +\exp(2rx_2)\frac{\exp(-\exp(\mu+rx_2))}{1-\exp(-\exp(\mu+rx_2))}| \\
&\leq 0.25\frac{1}{r^2}\{\grave{\grave{K}} - 2\mu(\grave{K}-\mu K)\}\exp(\mu)\{\exp(rx_1)+\exp(rx_2)\}.
\end{aligned}$$

This quantity goes to zero as μ goes to $-\infty$. With a similar argument it is easy to prove that $i22q_{11}d$ goes to zero as μ goes to $-\infty$. Hence the first term in (6.2) goes to zero as μ goes $-\infty$.

Finally it is straight forward to show that the third term goes to zero as μ goes to $-\infty$. Since i_{11} is a bounded function and each of $q_{22}d$ and r go to zero as μ goes $-\infty$ (using the same above argument). We conclude that $Tr(\mathbf{A}(x,\boldsymbol{\theta})(\mathbf{M}_{\mathbf{A}}^{-1}(\boldsymbol{\xi}_*,\boldsymbol{\theta}) - \mathbf{M}_{\mathbf{A}_*})$ goes to zero as μ goes to $-\infty$. □

6.2 Limiting Locally D-Optimal Designs for Models with Equal slopes

Recall from Section 4.2 on page 41 for the canonical positive-negative model with $\beta_1 = \beta_2 = \beta$ and $\boldsymbol{\theta}$=$(0,1,\mu)$, the candidate limiting locally D-optimal design was $\boldsymbol{\xi}_* = \begin{pmatrix} -0.8537 & 1.0773 & (0.8537-\mu) & (-1.0773-\mu) \\ 0.2900 & 0.2100 & 0.2900 & 0.2100 \end{pmatrix}$ when μ goes to $-\infty$. Although this design consists of four design points, these points are not the optimal design points obtained from concatenating the optimal designs points of the separate positive and negative extreme value models.

Theorem 6.2.1 *The locally limiting D optimal design is given by*
$$\boldsymbol{\xi}_* = \begin{pmatrix} -0.8537 & 1.0773 & (0.8537-\mu) & (-1.0773-\mu) \\ 0.2900 & 0.2100 & 0.2900 & 0.2100 \end{pmatrix}.$$

Proof: First we find the design that maximizes *log* det Fisher's information among all designs of the form $\boldsymbol{\xi} = \begin{pmatrix} (a & b & (-a-\mu) & (-b-\mu)) \\ \xi_1 & \xi_2 & \xi_3 & \xi_4 \end{pmatrix}$ when μ goes to $-\infty$. Then we show that the design is the locally D-optimal design using the General Equivalence Theorem. Let

$$v(t) = \exp(2t+2\mu)\exp(-\exp(xt+\mu))/(1-\exp(-\exp(t+\mu)))$$

and

$$w(t) = \exp(-\exp(t+\mu))\exp(-2t)\exp(-\exp(-t))/(1-\exp(-\exp(-t))).$$

Define the following:

$$v_1 = v(a) = \frac{\exp(2\mu+2a)\exp(-\exp(\mu+a))}{(1-\exp(-\exp(\mu+a)))},$$

$$v_2 = v(-a-\mu) = \frac{\exp(-2a)\exp(-\exp(-a))}{(1-\exp(-\exp(-a)))},$$

$$v_3 = v(b) = \frac{\exp(2\mu + 2b)\exp(-\exp(\mu + b))}{(1 - \exp(-\exp(\mu + b)))},$$

$$v_4 = v(-b - \mu) = \frac{\exp(-2b)\exp(-\exp(-b))}{(1 - \exp(-\exp(-b)))},$$

$$w_1 = w(a) = \frac{\exp(-\exp(\mu + a))\exp(-2a)\exp(-\exp(-a))}{(1 - \exp(-\exp(-a)))},$$

$$w_2 = w(-a - \mu) = \frac{\exp(-\exp(-a))\exp(2a + 2\mu)\exp(-\exp(a + \mu))}{(1 - \exp(-\exp(a + \mu)))}$$

$$w_3 = w(b) = \frac{\exp(-\exp(\mu + b))\exp(-2b)\exp(-\exp(-b))}{(1 - \exp(-\exp(-b)))},$$

$$w_4 = w(|(-b - \mu)) = \frac{\exp(-\exp(-b))\exp(2b + 2\mu)\exp(-\exp(b + \mu))}{(1 - \exp(-\exp(b + \mu)))}.$$

Evaluating Fisher's information given in Lemma 3.1.1 on page 16 for the above design we get

$$\begin{aligned}
\mathbf{M}(\boldsymbol{\xi}, \boldsymbol{\theta}) &= \xi_1 v_1 \begin{pmatrix} 1 & a & 0 \\ a & a^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \xi_1 w_1 + \begin{pmatrix} 0 & 0 & 0 \\ 0 & a^2 & a \\ 0 & a & 1 \end{pmatrix} \\
&\quad + \xi_2 v_2 \begin{pmatrix} 1 & -a-\mu & 0 \\ -a-\mu & (-a-\mu)^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \xi_2 w_2 \begin{pmatrix} 0 & 0 & 0 \\ 0 & (-a-\mu)^2 & -a-\mu \\ 0 & -a-\mu & 1 \end{pmatrix}
\end{aligned}$$

$$
+\xi_3 v_3 \begin{pmatrix} 1 & b & 0 \\ b & b^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \xi_3 w_3 + \begin{pmatrix} 0 & 0 & 0 \\ 0 & b^2 & b \\ 0 & b & 1 \end{pmatrix}
$$

$$
+\xi_4 v_4 \begin{pmatrix} 1 & -b-\mu & 0 \\ -b-\mu & (-b-\mu)^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \xi_4 x w_4 \begin{pmatrix} 0 & 0 & 0 \\ 0 & (-b-\mu)^2 & -b-\mu \\ 0 & -b-\mu & 1 \end{pmatrix}.
$$

Now when μ is negatively large, w_1 and w_3 are approximated by v_2 and v_4, respectively. Also v_1, v_3, w_2, w_4 tend to zero. Hence Fisher's information matrix can be approximated by

$$
\begin{aligned}
\dot{\mathbf{M}}(\boldsymbol{\xi}_*, \boldsymbol{\theta}) \quad &= \quad \xi_1 v_2 \begin{pmatrix} 0 & 0 & 0 \\ a & a^2 & a \\ 0 & a & 1 \end{pmatrix} + \xi_3 v_4 \begin{pmatrix} 0 & 0 & 0 \\ 0 & b^2 & b \\ 0 & b & 1 \end{pmatrix} \\
&+\xi_2 v_2 \begin{pmatrix} 1 & -a-\mu & 0 \\ -a-\mu & (-a-\mu)^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \xi_4 v_4 \begin{pmatrix} 1 & -b-\mu & 0 \\ -b-\mu & (-b-\mu)^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}.
\end{aligned}
$$

$$
= \begin{pmatrix} \xi_2 v_2 + \xi_4 v_4 & v_2 \xi_2(-a-\mu) + \xi_4 v_4(-b-\mu) & 0 \\ v_2 \xi_2(-a-\mu) + \xi_4 v_4(-b-\mu) & m22 & \xi_1 v_2 a + \xi_3 v_4 b \\ 0 & \xi_1 v_2 a + \xi_3 v_4 b & \xi_1 v_2 + \xi_3 v_4 \end{pmatrix},
$$

where $m22 = \xi_1 v_2 a^2 + \xi_3 v_2(-a-\mu)^2 + \xi_3 v_4 b^2 + \xi_4 v_4(-b-\mu)^2$.

Maximizing the determinant of $\dot{\mathbf{M}}(\boldsymbol{\xi}_*, \boldsymbol{\theta})$ using NPSOL yields $a = -0.8536657, b = 1.077288, \xi_1 = \xi_2 = 0.2895051$, and $\xi_3 = \xi_4 = 0.2104949$.

Now we show that the directional derivative at this design is non-positive and its maximum is zero. The directional derivative at this design is

$$\begin{aligned} F_D(x, \boldsymbol{\xi}_*, \boldsymbol{\theta}) &= Trace(\mathbf{I}(x, \boldsymbol{\theta})\mathbf{M}^{-1}(\boldsymbol{\xi}_*, \boldsymbol{\theta})) - 3 \\ &= Trace(\mathbf{I}(x, \boldsymbol{\theta})\dot{\mathbf{M}}^{-1}(\boldsymbol{\xi}_*, \boldsymbol{\theta})) \\ &\quad + Trace(\mathbf{I}(x, \boldsymbol{\theta})[\mathbf{M}^{-1}(\boldsymbol{\xi}_*, \boldsymbol{\theta}) - \dot{\mathbf{M}}^{-1}(\boldsymbol{\xi}_*, \boldsymbol{\theta}))]) - 3. \end{aligned}$$

Recall that when μ is negatively large $\mathbf{M}(\boldsymbol{\xi}_*, \boldsymbol{\theta}) \doteq \dot{\mathbf{M}}(\boldsymbol{\xi}_*, \boldsymbol{\theta})$ and each element of $I(x, \boldsymbol{\theta})$ is a bounded function. This implies that $Trace(\mathbf{I}(x, \boldsymbol{\theta})[\mathbf{M}^{-1}(\boldsymbol{\xi}_*, \boldsymbol{\theta}) - \dot{\mathbf{M}}^{-1}(\boldsymbol{\xi}_*, \boldsymbol{\theta}))])$ $\to 0$ as $\mu \to -\infty$. We need to show now that $Tr = Trace(\mathbf{I}(x, \boldsymbol{\theta})\dot{\mathbf{M}}^{-1}(\boldsymbol{\xi}_*, \boldsymbol{\theta})) - 3 \leq 0$ as $\mu \to -\infty$. Hence the directional derivative at this design F_D is non-positive as $\mu \to -\infty$ and the proof will be complete.

Denote $v(x)$ and $w(x)$ by v_x and w_x, respectively. Let $g_x = g(x) = \exp(-2x)\exp(-\exp(-x))/(1 - \exp(-\exp(-x)))$. Note that $w_x \leq g_x$ for all values of $\mu \leq 0$. Using Maple software, (see Appendix A4), the trace (Tr) is found to be

$$\begin{aligned} Tr &= v_x[-2.091898121\mu + 3.023967019\mu^2 + 4.732702573] \\ &\quad +2v_x x[-1.045949058 + 3.023967013\mu] + 3.023967012 v_x x^2 \\ &\quad +3.023967012 w_x x^2 + 2.091898118 w_x x + 4.732702591 w_x \end{aligned}$$

Case when $\mathbf{x} < -\mu/\mathbf{2}$:

Assuming different large negative values of μ, for each sequence such that

$x < -\mu/2$ the following statements were verified by S-plus: $v(x) < v(-\mu/2), -(x + \mu)v(x) < (-\mu/2)v(-\mu/2)$ and $(x+\mu)^2 v(x) < (\mu^2/4)v(-\mu/2)$. Assuming these inequalities hold and keeping four significant digits, we rearrange Tr as

$$\begin{aligned}
Tr &= 3.0240(x+\mu)^2 v_x - 2.0919(x+\mu)v_x + 4.7327v_x \\
&\quad +(3.0240x^2 w_x + 2.0919xw_x + 4.7327w_x) \\
&< 3.0240(\mu^2/4)v(-\mu/2) + 2.0919(-\mu/2)v(-\mu/2) + 4.7327v(-\mu/2) \\
&\quad +(3.0240x^2 g_x + 2.0919xg_x + 4.7327g_x) \\
&\leq 3.0240(\mu^2/4)v(-\mu/2) + 2.0919(-\mu/2)v(-\mu/2) + 4.7327v(-\mu/2) \\
&\quad +Max(3.0240x^2 g_x + 2.0919xg_x + 4.7327g_x) \\
&= 3.0240(\mu^2/4)v(-\mu/2) + 2.0919(-\mu/2)v(-\mu/2) + 4.7327v(-\mu/2) + 3.
\end{aligned}$$

Case when $\mathbf{x} \geq (-\mu/\mathbf{2})$:

For each sequence of x values such that $x < -\mu/2$ with different series of large negative values of μ the following statements hold. They were verified by S-plus: $w_x \leq g_x$, $g(x) < g(-\mu/2), xg(x) < (-\mu/2)g(-\mu/2)$ and $x^2 g(x) < (\mu^2/4)g(-\mu/2)$. Assuming these inequalities hold and keeping four significant digits, we rearrange Tr as

$$\begin{aligned}
Tr &= 3.0240(x+\mu)^2 v_x - 2.0919(x+\mu)v_x + 4.7327v_x \\
&\quad +(3.0240x^2 w_x + 2.0919xw_x + 4.7327w_x) \\
&\leq (3.0240(x+\mu)^2 v_x - 2.0919(x+\mu)v_x + 4.7327v_x) \\
&\quad +(3.0240x^2 g_x + 2.0919xg_x + 4.7327g_x)
\end{aligned}$$

$$
\begin{aligned}
&< \quad (3.0240(x+\mu)^2 v_x - 2.0919(x+\mu)v_x + 4.7327v_x) \\
&\quad +3.0240(\mu^2/4)g(-\mu/2) + 2.0919(-\mu/2)g(-\mu/2) + 4.7327g(-\mu/2) \\
&\leq \quad Max(3.0240(x+\mu)^2 v_x - 2.0919(x+\mu)v_x + 4.7327v_x) \\
&\quad +3.0240(\mu^2/4)g_x|(-\mu/2) + 2.0919(-\mu/2)g(-\mu/2) + 4.7327g(-\mu/2) \\
&< \quad 3 + 3.0240(\mu^2/4)g(-\mu/2) + 2.0919(-\mu/2)g(-\mu/2) + 4.7327g(-\mu/2.)
\end{aligned}
$$

Note that $v(-\mu/2) = g(-\mu/2) = \exp(u)\exp(-\exp(\mu/2))/(1-\exp(-exp(\mu/2))) \to 0$ as $\mu \to -\infty$, which implies that $Tr = Trace(\mathbf{I}(x,\boldsymbol{\theta})\dot{\mathbf{M}}^{-1}(\boldsymbol{\xi}_*,\boldsymbol{\theta}))$ goes to 3. Therefore, the directional derivative F_D is non-positive and the maximum is zero as μ goes to $-\infty$ and the proof of Theorem 6.2.1 is complete. □

6.3 Efficiency of Limiting Locally D-optimal Designs

The limiting D-optimal Designs found in the previous sections are found to be very efficient with respect to the exact optimal designs. The limiting optimal designs were found when μ takes negative large values and in applications μ might sufficiently large. The definition of efficiency we use is given by

Definition 6.3.1 *Efficiency is given by the sample size needed for an experiment using the optimal design to reach the same criterion value as an experiment using the limiting design with sample size one.*

The sample size needed for limiting design $\boldsymbol{\xi}_{\mathbf{L}}$ to be as efficient at $\boldsymbol{\xi}^*$ when $\beta_1 \neq \beta_2$ is calculated as follows:

$$\begin{aligned}
log\det(n\mathbf{M}(\boldsymbol{\xi}^*,\boldsymbol{\theta}) &= log\det\mathbf{M}(\boldsymbol{\xi}_{\mathbf{L}},\boldsymbol{\theta})) \\
logn^4\det(\mathbf{M}(\boldsymbol{\xi}^*,\boldsymbol{\theta}) &= log\det\mathbf{M}(\boldsymbol{\xi}_{\mathbf{L}},\boldsymbol{\theta})) \\
4logn + log\det(\mathbf{M}(\boldsymbol{\xi}^*,\boldsymbol{\theta}) &= log\det\mathbf{M}(\boldsymbol{\xi}_{\mathbf{L}},\boldsymbol{\theta})) \\
logn &= [log\det\mathbf{M}(\boldsymbol{\xi}_{\mathbf{L}},\boldsymbol{\theta})) - log\det(\mathbf{M}(\boldsymbol{\xi}^*,\boldsymbol{\theta})]/4 \\
n &= \exp[log\det\mathbf{M}(\boldsymbol{\xi}_{\mathbf{L}},\boldsymbol{\theta})) - log\det(\mathbf{M}(\boldsymbol{\xi}^*,\boldsymbol{\theta})]/4.
\end{aligned}$$

In a similar way when $\beta_1 = \beta_2$ the required sample size is given by

$$n = \exp[log\det\mathbf{M}(\boldsymbol{\xi}_{\mathbf{L}},\boldsymbol{\theta})) - log\det(\mathbf{M}(\boldsymbol{\xi}^*,\boldsymbol{\theta})]/3.$$

Figures 6.2 -6.6 display plots of required sample size n versus μ when $\beta_1 \neq \beta_2$ for different values of r. The efficiencies for large negative values of μ approaches 0.99 with both large and small values of r. For small to moderate values of $|\mu|$, the efficiencies vary with the value of r, and they range from $71\% - 91\%$. It can be seen that the efficiencies are higher for $r = 0.5$ and $r = 1$ than larger values of r, but they are still reasonable. In general, as the values of $|\mu/2r|$ get larger and larger the efficiency approaches 1.

Figure 6.7 shows the efficiency plot n versus μ when $\beta_1 = \beta_2$. These efficiencies are all higher than 92% and they approaches 96% for large values of $|\mu|$.

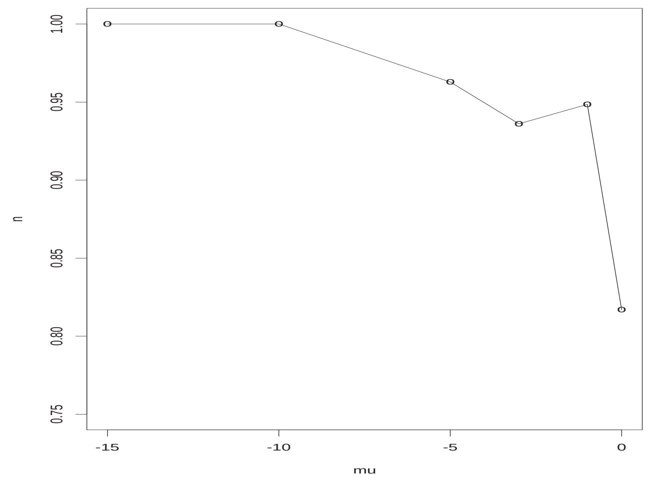

Figure 6.2: Efficiency plot for limiting D-optimal design when $\beta_1 \neq \beta_2$ and $r = 0.5$

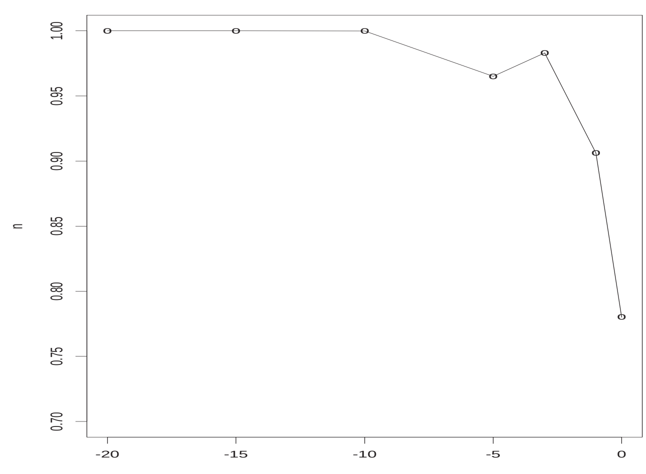

Figure 6.3: Efficiency plot for limiting D-optimal design when $\beta_1 \neq \beta_2$ and $r = 1$

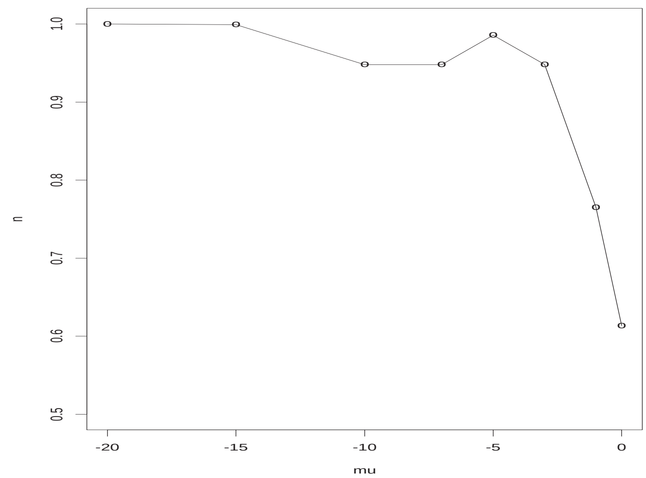

Figure 6.4: Efficiency plot for limiting D-optimal design when $\beta_1 \neq \beta_2$ and $r = 2$

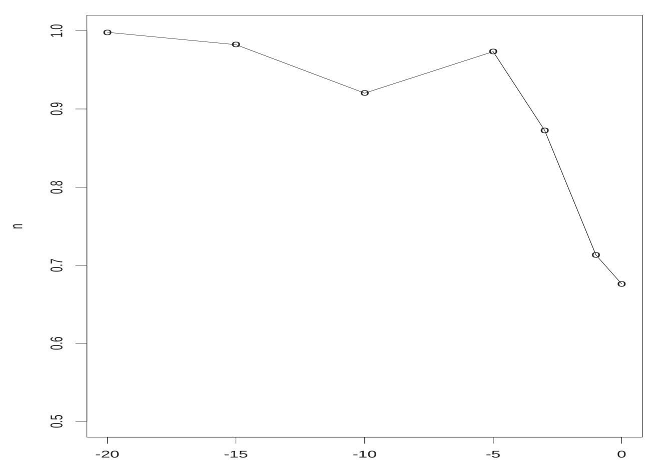

Figure 6.5: Efficiency plot for limiting D-optimal design when $\beta_1 \neq \beta_2$ and $r = 3$

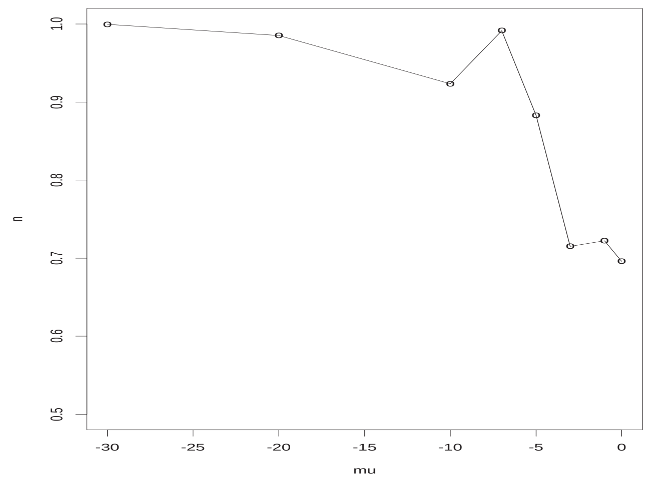

Figure 6.6: Efficiency plot for limiting D-optimal design when $\beta_1 \neq \beta_2$ and $r = 4$

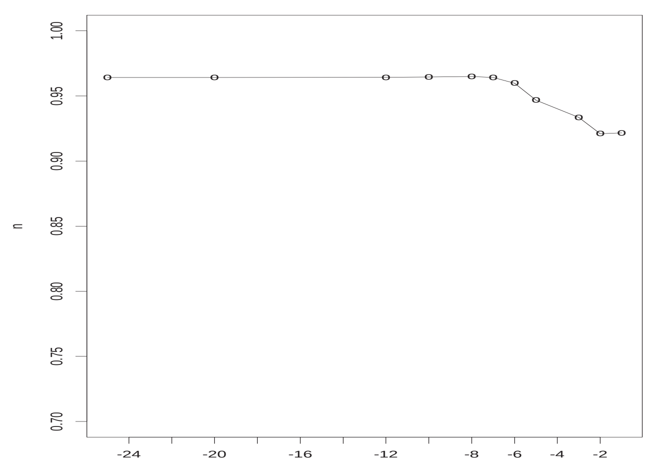

Figure 6.7: Efficiency plot for limiting D-optimal design when $\beta_1 = \beta_2$

6.4 Limiting Locally c-optimal Designs

Unfortunately for c-optimal designs we were not able to find theorems similar to those found in the previous sections to D-optimal designs. When the two slopes are not equal with $\boldsymbol{\theta}=(\alpha_2, \beta_2, \mu, r)$, the locally c-optimal designs consist of two design points with different weights depending on the different values of r, μ. See Tables 5.1 and 5.2. No pattern was observed. When the two slopes are equal with $\boldsymbol{\theta}=(\alpha_2, \beta_2, \mu)$, the locally c-optimal designs consist of two points with different weights. When μ assumes negative large values the weights become equal. (See Table 5.3). One can suggest that a candidate limiting locally c-optimal design consist of two equally weighted at $-0.47, 0.47-\mu$. It is interesting to note that the maximum of the function $(G')^2/(G(1-G))$ is 0.6476 attained at $x = -0.47$.

But unfortunately a similar theorem to theorems proved in the previous sections was not found. It was difficult to prove that the limiting locally c-optimal consists of a two point design equally weighted at -0.47, and $0.47 - \mu$ as μ goes to $-\infty$. The computation of the directional derivative for this candidate design was too complicated and the Maple output was about 40 pages with many terms on each page that were very hard to simplify.

We conjecture that the limiting locally c-optimal consist of two point design equally weighted at -0.47, and $0.47 - \mu$ as μ goes to $-\infty$. Motivated by the high efficiencies we found (see Figure 6.9), we show graphically that this conjecture holds. Figure 6.8 shows the directional derivatives for $\mu = -10, -15$. The directional derivatives are nonpositive and achieve the maximum at the candidate design points. Computational problems arise when μ is a very large negative large value.

The efficiency for c-optimal designs is defined as in Section 6.3 on page 90. The efficiencies of this candidate designs are higher than 0.965 as it can be seen from Figure 6.9 .

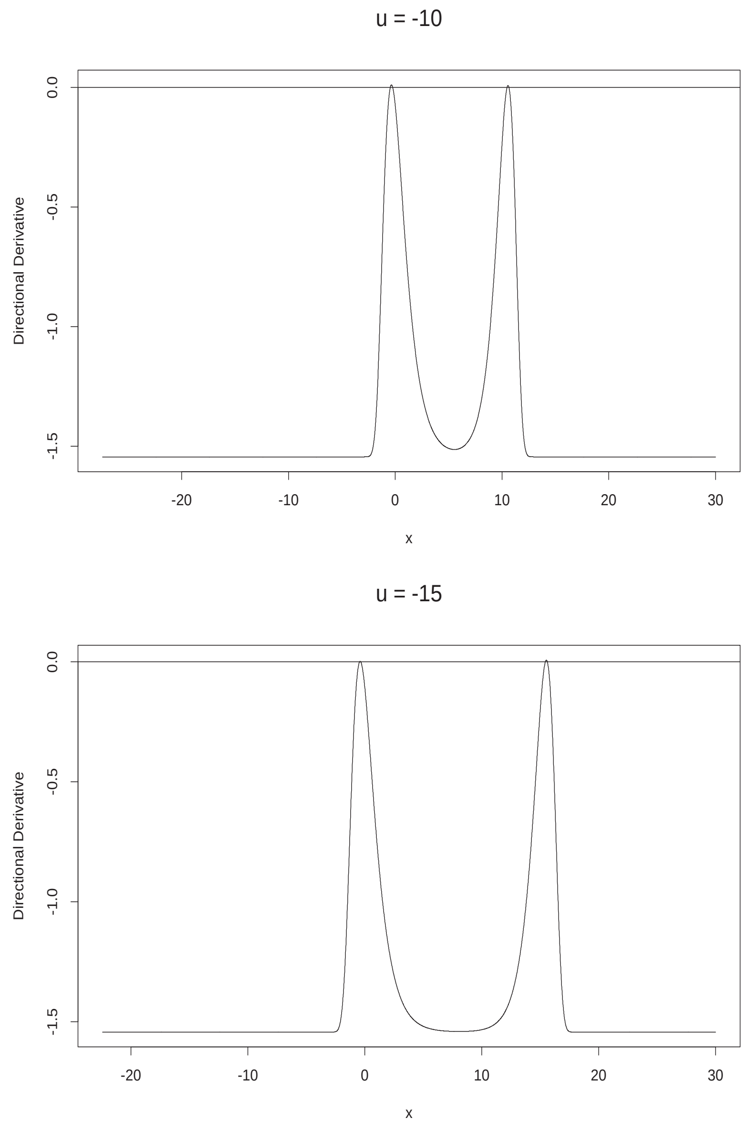

Figure 6.8: Directional derivative for limiting c-optimal design when $\mu = -10, -15$.

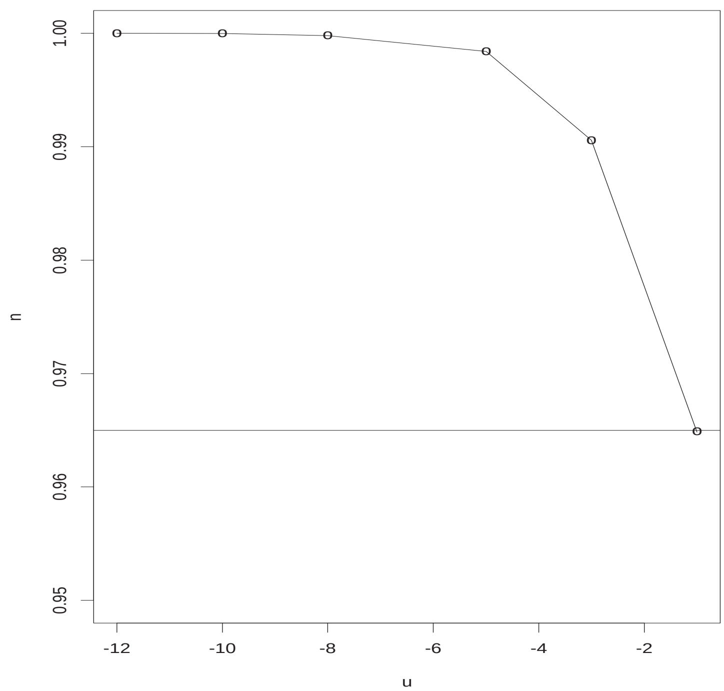

Figure 6.9: Efficiency for limiting c-optimal design.

Chapter 7

UP–AND–DOWN PROCEDURES

Optimal design criteria typically are functions of Fisher's information matrix. In nonlinear models, Fisher's information depends on the unknown parameters, so for the contingent response model, c-and-D-optimal designs serve as unattainable *gold standards* for efficient estimation. To implement these designs, one can put a prior distribution on the parameters and average the locally optimal designs with respect to this prior distribution. Such designs are called Bayesian (*cf.* Chaloner and Verdinelli [4]). Sequential designs provide another approach to this problem. We focus on the use of Markovian up-and-down designs and discuss some other alternative procedures.

7.1 Up-and-Down Procedures

Markovian up-and-down procedures have been constructed to cluster treatments around the unknown quantals of an increasing response function [See Durham and Flournoy ([7], [8]), Derman [6], and Giovagnoli and Pintacuda [20]]. Giovagnoli and Pintacuda [21] suggest using Markovian up–and–down designs to cluster points around optimal design points. For the contingent response model, we follow this suggestion using Theorem 1 of Durham and Flournoy [7] for random walks rules that states, under mild conditions, that if the P(increasing the dose) decreases while P(decreasing the dose) increases, then the asymptotic treatment distribution will be unimodal. Durham and Flournoy's Theorem 1 also provides insight for adjusting treatment allocation procedures to shift the treatment mode into a close neighborhood of the optimal design points, namely, treatments will cluster unimodally in the neighborhood of the point for which the P(increasing the dose) equals the P(decreasing the dose). We will say a procedure *targets* a dose x if it produces a unimodal treatment distribution with mode x_m such that $x_{m-1} \leq x < x_{m+1}$.

Let treatments belong to a finite set $\Omega_x = \{x_1, \ldots, x_K\}$ and let $N_j(n)$ denote the number of subjects treated at x_j up to and including the n^{th} subject. Then we call $\mathbf{N}(n)/n = \{N_1(n)/n, N_2(n)/n, \ldots, N_K(n)/n\}$ the *treatment distribution.* Let $X(n)$ denote the treatment for the n^{th} subject. It is not hard to show Proposition 7.1.1.

Proposition 7.1.1 *(cf. Karlin and Taylor, [26]). Assume* $P_{ij} = P(X(n+1) = x_j | X(n) = x_i)$ *are first order Markovian transition probabilities with* $P_{ij} = 0$ *for* $|j - i| > 1$*. If* $P_{i,\,i-1} > 0, i = 2, \ldots, K, P_{i,\,i+1} > 0, i = 1, \ldots, K-1, P_{11} > 0,$ *and* $P_{KK} > 0,$ *a limiting treatment distribution will exist:*

$$\lim_{n \to \infty} \frac{\mathbf{N}(n)}{n} = \{\pi_1, \pi_2, \ldots, \pi_K\} = \boldsymbol{\pi},$$

where $\pi_j = \pi_1 \Pi_{i=2}^{j} P_{i-1,\,i}/P_{i,\,i-1},\ ,\ j = 2, \ldots, K,$ *and* $\Sigma_{j=1}^{K} \pi_j = 1.$

A Markovian up–and–down procedure can be made to target an optimal design point by, for example, the use of biased coins. Define four Bernoulli random variables:

$$B_i = \begin{cases} 1 & \text{with probability } b_i,\ i = 1, \ldots, 4. \\ \\ 0 & \text{else} \end{cases}$$

It is natural to assume the dose will not be increased following a toxic failure, nor will it be decreased following disease failure. Thus consider rules for changing dose levels that can be described by a first order Markov chain with transition probabilities

$$\begin{aligned} P_{i,\,i-1} &= P(X(n) = x_{i-1})|X(n-1) = x_i) \\ &= b_1 F(x_i) + b_4(1-b_2)H(x_i) \\ P_{i,\,i} &= P(X(n) = x_i)|X(n-1) = x_i) \\ &\quad = b_2 H(x_i) + (1-b_1)F(x_i) + (1-b_3)\bar{F}(x_i)\bar{G}(x_i) \\ P_{i,\,i+1} &= P(X(n) = x_{i+1})|X(n-1) = x_i) \\ &= b_3 \bar{F}(x_i)\bar{G}(x_i) + (1-b_4)(1-b_2)H(x_i) \\ P_{i,\,j} &= 0,\ |i-j| > 1 \end{aligned} \tag{7.1}$$

with $P_{i,\,i-1} + P_{i,\,i} + P_{i,\,i+1} = 1$. Then following directly Durham and Flournoy Theorem 1 [7], we have

Theorem 7.1.1 *A Markovian up–and–down procedure meeting the conditions of Proposition 7.1.1 with probability transitions given by (7.1) and with any* $b_1, \ldots, b_4 \in [0, 1]$

such that $P_{i,i-1} = P_{i,i+1}$, *i.e.,*

$$b_1 F(x) + b_4(1-b_2)H(x) = b_3 \bar{F}(x)\bar{G}(x) + (1-b_4)(1-b_2)H(x), \tag{7.2}$$

targets the dose x.

The asymptotic design resulting from such up–and–down procedure is $\boldsymbol{\xi} = \begin{pmatrix} x_1 & \dots & x_K \\ \\ \pi_1 & \dots & \pi_K \end{pmatrix}$, where the $\{\pi_i\}$ are calculated according to the formulae in Proposition 7.1.1.

Corollary 7.1.1 *As a practical special case, set* $b_2 = 1$. *This prescribes that one will treat with the same dose again following a success. Then procedure will target the point* x *for which*

$$F(x)/\bar{F}(x)\bar{G}(x) = b_3/b_1. \tag{7.3}$$

In Section 7.1 we specify an up–and–down procedure motivated by Theorem 7.1.1 and optimal design theory that causes treatments to cluster around the optimal design points. Like the optimal designs, the specification of this procedure depends on the unknown parameters, so they too are not directly implementable. However, they have value as "*gold standard*" up–and–down designs that are useful for comparison purposes.

In Section 7.1.2 we generalize an ad hoc up–and–down procedure proposed by Flournoy[15]. We characterize this procedure using Proposition 7.1.1 and Theorem 7.1.1 These procedures form the basis of current experiments. In Section 7.1.3 we describe an alternative up–and–down procedure studied by Kpamegan and Flournoy (2000) that causes treatments to cluster around the optimal dose. In Section 7.2, we compare these procedures.

7.1.1 An Up–and–Down Procedure Targeting the Optimal Design Points

If the n^{th} subject is treated at x_k, the most straightforward treatment allocation rule satisfying (7.1) and (7.2) is to set $b_2 = 1$ and treat the $(n+1)st$ subject as follows:

$$X_{n+1} = \begin{cases} x_{k-1} & \text{if } (Y_1 = 1, B_1 = 1) \\ \\ x_k & \text{if } (Y_1 = 0, Y_2 = 1, B_3 = 0) \text{ or } (Y_1 = 1, B_1 = 0) \text{ or } (Y_1 = 0, Y_2 = 0) \\ \\ x_{k+1} & \text{if } (Y_1 = 0, Y_2 = 1, B_3 = 1). \end{cases} \tag{7.4}$$

Define $\Gamma_{Tj} = F(x_j^*)$ and $\Gamma_{Dj} = \bar{F}(x^*)\bar{G}(x_j^*), j = 1, \ldots, P$ where x_j^* is an optimal design point and P is the number of optimal design points. Setting $b_3/b_1 = F(x_j^*)/\bar{F}(x^*)\bar{G}(x_j^*) = \Gamma_{Tj}/\Gamma_{Dj}$ causes the treatment distribution to target the j^{th} optimal design point. Therefore, as subjects arrive, assign them to the treatment sequence targeting x_j^* with probability w_j. Denote the asymptotic design that results from this procedure by $\boldsymbol{\xi}_{x^*}$.

For example, the c-optimal design points for the $(-3, 2)$ canonical model are found in Table 5.2 to be ($x_1^* = -0.5054, x_2^* = 1.3595$) with weights ($w_1 = 0.4444, w_2 = 0.5556$). So

$$\Gamma_{T1}/\Gamma_{D1} = 0.0180/0.7949 = 0.023;$$

$$\Gamma_{T2}/\Gamma_{D2} = 0.5300/0.1064 = 4.981.$$

Since this optimal design has two points, we need two up–and–down procedures with $b_3/b_1 = 0.023$ and 4.981, respectively. Therefore, as subjects arrive, assign them to the treatment sequences targeting x_1^* and x_2^* with probability 0.4444 and 0.5556, respectively (see Appendix B.3). Figure 7.1 shows the canonical $(-3, 2)$ model and the treatment distribution that results from using this up–and–down procedure. The treatment modes resulting from the two procedures are -0.5 and 1.5 respectively. As predicted by random walk design theory, these are the largest possible treatments that are not greater than x_1^* and x_2^*, respectively.

This procedure is still "local" because it depends on the parameters (μ, r).

7.1.2 Up–and–Down Procedure Balancing Failure Rates

An appealing ad hoc procedure in the contingent response setting is to decrease the dose after each toxicity and increase the dose after each disease failure. This was proposed by Flournoy [15], used as the basis of a procedure simulated by Gooley [22] et al. and currently used in practice.

Toxicity and disease failures may not be of equal significance. For, example, if disease failures are preferred over toxicities, then one wants a design that will

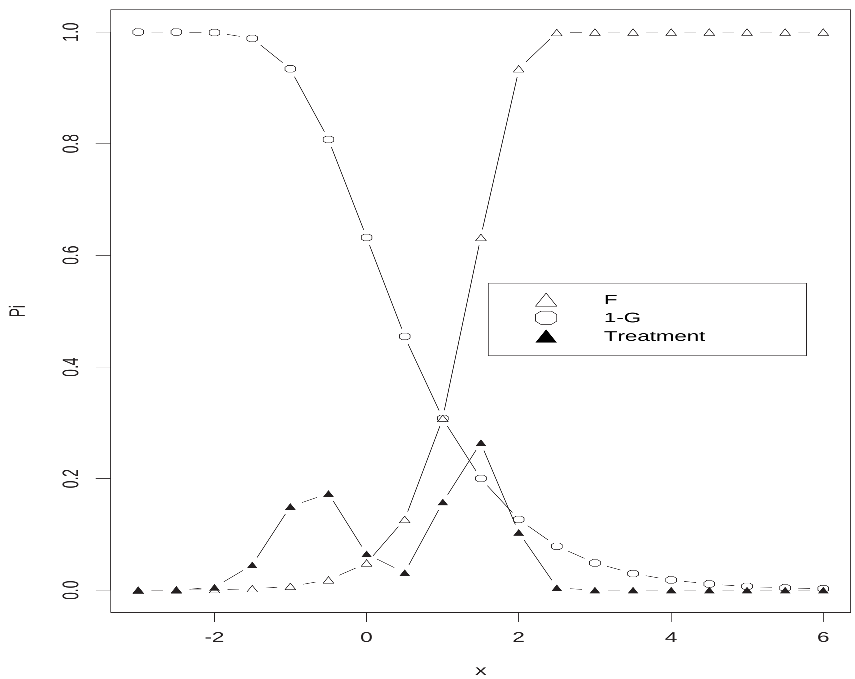

Figure 7.1: Treatment Distribution Targeting the Optimal Design Points of the Canonical Model $(-3, 2)$.

increase the dose for given disease failure more often than it decreases the dose given toxicity. In this case, one wants a design in which the ratio of the rate of failure due to disease failure to the rate of failure due to toxicity is greater than one. This can be accomplished by using the up–and–down procedure given by (7.4) with $b_3/b_1 > 1$. Although the procedure is not based on any optimality properties, it is implementable in that it does not depend on the model parameters, and scientists have suggested it to an author. Denote the design that is produced from this procedure by $\boldsymbol{\xi}_\rho$, where $\rho = b_3/b_1$.

7.1.3 Kpamegan and Flournoy's Up–and–Down Procedure

An up–and–down design procedure was studied by Kpamegan and Flournoy [28] with the ethical goal of clustering treatments around the optimal dose, rather the optimal design points. Consider Ω_x, the set of possible dosages, with $\Delta = x_j - x_{j-1}$ and let $X(n)$ be the midpoint of the dose interval for the n^{th} pair of subjects. Then the optimizing up–and–down for selecting the dose with maximum success probability is defined by the following algorithm with adjustments at the treatment boundaries.

If the n^{th} pair of subjects has been treated at $X(n) - \frac{\Delta}{2}$ and $X(n) + \frac{\Delta}{2}$, the midpoint of the (n+1)st pair is

$$X(n+1) = X(n) + \Delta V(n),$$

where

$$V_n = \begin{cases} -1 & \text{if the treatment at } X(n) - \frac{\Delta}{2} \text{ results in success} \\ & \text{and the treatment at } X(n) + \frac{\Delta}{2} \text{ in failure;} \\ 0 & \text{if the } n^{th} \text{ pair of treatment results in two success} \\ & \text{or two failure;} \\ 1 & \text{if the treatment at } X(n) - \frac{\Delta}{2} \text{ results in failure} \\ & \text{and the treatment at } X(n) + \frac{\Delta}{2} \text{ in success.} \end{cases}$$

This procedure produces a design we denote by $\boldsymbol{\xi}_{KF}$. As opposed to the procedure described in Section 7.1, this procedure is independent of the model parameters.

7.2 Comparisons

In this section we examine efficiency of the up–and–down designs discussed in the Section 7.1.3. All comparisons assume the positive-negative extreme value contingent response model. We define the efficiency of two designs for estimating the optimal dose $\mathbf{g}(\boldsymbol{\Theta})$ to be the ratio of the asymptotic variances of their maximum likelihood estimators. That is, if $\boldsymbol{\xi}_1$ is one design and $\boldsymbol{\xi}_2$ is some other design, then

$$E(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2) = \frac{\dot{\mathbf{g}}^{\mathbf{T}}(\boldsymbol{\Theta})\mathbf{M}^{-1}(\boldsymbol{\xi}_1, \boldsymbol{\Theta})\dot{\mathbf{g}}(\boldsymbol{\Theta})}{\dot{\mathbf{g}}^{\mathbf{T}}(\boldsymbol{\Theta})\mathbf{M}^{-1}(\boldsymbol{\xi}_2, \boldsymbol{\Theta})\dot{\mathbf{g}}(\boldsymbol{\Theta})}.$$

In Table 7.2, the canonical $(-3, 2)$ positive-negative extreme value model is assumed true. $E(\boldsymbol{\xi}_{x^*}, \boldsymbol{\xi}^*) = 0.85$. That is, the up–and–down design targeting the optimal design points is 85% efficient compared with the locally optimal design. This up–and–down design was characterized with $\Omega_x = \{-3, 2.5, \ldots, 5.5, 6, \}$, so $K = 19$

and the interval between doses is $\Delta = 0.5$.

Table 7.1 shows, for a variety of canonical positive-negative extreme value response models, that decreasing Δ and increasing K over the same span of points increases the efficiency of up–and–down designs targeting the optimal design points. This is expected because the variance of a Markovian up–and–down design is inversely proportional to Δ. So decreasing Δ causes the treatment distribution to converge to the optimal design.

Returning attention to Table 7.2, efficiencies relative to the locally optimal design also are given for designs denoted by $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_9$. These are asymptotic designs resulting from using the Markovian up–and–down procedures (7.1) operating on the same treatment space Ω_x as $\boldsymbol{\xi}_{x^*}$. The true response function for all efficiency calculations in this table is the canonical $(-3, 2)$ positive-negative extreme value response model. Since determining $\boldsymbol{\xi}_{x^*}$ requires specification of b_3/b_1 which requires specification of the optimal design points, it is of interest to see how model misspecification influences the efficiency of estimating the optimal dose $g(\boldsymbol{\Theta})$. Thus b_3/b_1 is calculated assuming the values of r and μ given for $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_9$. b_2 is again set to 1. So the procedures producing these designs differ in that they target the wrong design points. Note that, efficiency decreases as μ decreases for $r = 2, 3$ and 0.5, that is, as the response functions become further apart. This decrease in efficiency is accelerated as the ratio of misspecified scale parameters deviates from the true $r = 2$. If one considers $\boldsymbol{\xi}_{x^*}$ the "gold standard" for Markovian up–and–down designs targeting

r	μ	$\Delta = 0.5$	$\Delta = 0.25$
0.5	-1	0.88	0.94
0.5	-3	0.93	0.94
2	-1	0.85	0.93
2	-3	0.75	0.87
2	-5	0.89	0.94
2	-7	0.82	0.94
3	-1	0.74	0.86
3	-3	0.80	0.90
3	-5	0.83	0.91
3	-7	0.85	0.92

Table 7.1: Efficiency's dependence on Δ

the optimal design points with $\Delta = 0.5$, then model specification doesn't hurt much until it becomes rather extreme.

As described in Section 7.1.2, it is natural to consider decreasing the dose after a toxic failure and increasing it after a dose failure. The authors have also heard investigators suggest modifying this up–and–down procedure to give toxicity and disease failure different weights. We found the asymptotic treatment distribution assuming procedure (7.4) for six given values of b_3/b_1 and a variety of positive-negative extreme value models. The efficiencies resulting from such procedures are shown in Tables 7.3 and 7.4, together with the mode of the treatment distribution and the expected rates of toxicity, disease failure and success. The expectations are calculated as $E(F(x)) = \sum \pi_i F(x_i)$, $E(\bar{F}(x)\bar{G}(x)) = \sum \pi_i \bar{F}(x_i)\bar{G}(x_i)$ and $E(H(x)) = \sum \pi_i \bar{F}(x_i) G(x_i)$.

Thus, for example, one can see from Table 7.4, for the canonical $(-3, 3)$positive-negative extreme model, that if one considers toxicity twice as serious as disease failure, and hence sets $b_3/b_1 = 2$, efficiency will only be 15% relative to $\boldsymbol{\xi}_{x^*}$, the dose targeting the optimal design points, and it will do even worse relative to the locally optimal design. In general the efficiencies of these designs, defined by heuristically fixing b_3/b_1, perform very poorly and should avoided.

For each model evaluated in Tables 7.3 and 7.4, the Kpamegan-Flournoy design $\boldsymbol{\xi}_{KF}$ is also evaluated. Cross checking with Table 5.2, one sees that the treatment mode resulting from these procedures is close to the optimal dose $g(\boldsymbol{\Theta})$. Because of this, the efficiency is rather low, but not nearly so low when b_3/b_1 is heuristically fixed. Surprisingly, the expected numbers of successes are often not increased compared with $\boldsymbol{\xi}_{x^*}$.

Comparing Tables 7.3 and 7.4 with Table 7.2, we conclude that, if up–and–down designs are used for the contingent response model, better estimates of the optimal dose are expected using one that targets the optimal design points, even through it will most certainly be misspecified, than heuristically weighting the failures by fixing b_3/b_1 or by targeting the optimal dose as done by Kpamegan and Flournoy.

7.3 Conclusion

Clearly more work is needed to find efficient, ethical designs for the contingent response model.

Other sequential procedures for approximating optimal designs include using

				b_3/b_1	
	r	μ	Efficiency		
ξ^*	2	-3	1		
ξ_{x^*}	2	-3	0.85	0.023	4.981
ξ_1	2	-1	0.78	0.061	10.728
ξ_2	2	-5	0.76	0.007	3.230
ξ_3	2	-7	0.62	0.016	2.680
ξ_4	3	-1	0.83	0.016	4.210
ξ_5	3	-3	0.67	0.005	2.400
ξ_6	3	-5	0.49	0.001	1.793
ξ_7	3	-7	0.22	0.000	1.545
ξ_8	0.5	-1	0.16	0.797	502.466
ξ_9	0.5	-3	0.21	0.519	1637.870

Table 7.2: Efficiencies with model misspecification

sequential maximum likelihood estimates (White, [39], McLeish and Tosh, [30], Fedorov, [14] p.186, and Wynn, [40]), the directed walk of Hardwick, Meyer and Stont [24], a sequential Bayesian approach (Haines, Perevoskaya, and Rosenberger, [23], urn designs by Flournoy [34] and Mungo, Zhu, and Rosenberger [32]. Application of the sequential maximum likelihood and Bayesian procedures to the contingent response model is, conceptually, straightforward. In sequential and optimal Bayesian designs, after each outcome is observed, the optimality criterion is re-estimated by applying the method of maximum likelihood or by finding a new posterior model, respectively. The criterion is then optimized to obtain a new estimate of the optimal design, which forms the basis for allocating the next subject, or group of subjects, to treatments.

The sequential maximum likelihood approach is difficult to apply in practice, and difficult to analyze, because of well known problems with the existence of estimators for small samples. For use in a single response function model (e.g. logistic),

	Mode	$E(\boldsymbol{\xi}_{x^*}, \boldsymbol{\xi})$	$EF(x)$	$E(\bar{F}(x)\bar{G}(x))$	$EH(x)$
$\mu = -1, r = 0.5$					
$\boldsymbol{\xi}^*$		1.14	0.56	0.25	0.20
$\boldsymbol{\xi}_{\boldsymbol{x}^*}$	0.0, 4.0	1	0.47	0.32	0.21
$\boldsymbol{\xi}_1$	0.5	0.16	0.36	0.36	028
$\boldsymbol{\xi}_{1.5}$	0.5	0.19	0.41	0.27	0.32
$\boldsymbol{\xi}_2$	1	0.22	0.44	0.22	0.34
$\boldsymbol{\xi}_{2.5}$	1	0.22	0.47	0.19	0.35
$\boldsymbol{\xi}_3$	1	0.22	0.49	0.16	0.35
$\boldsymbol{\xi}_{3.5}$	1.5	0.20	0.51	0.15	0.35
$\boldsymbol{\xi}_{KF}$	1.25	0.69	0.55	0.27	0.18
$\mu = -3, r = 0.5$					
$\boldsymbol{\xi}^*$		1.08	0.24	0.15	0.62
$\boldsymbol{\xi}_{\boldsymbol{x}^*}$	1.5, 6.0	1	0.22	0.16	0.62
$\boldsymbol{\xi}_1$	2	0.09	0.01	0.13	0.74
$\boldsymbol{\xi}_{1.5}$	2.5	0.12	0.15	0.10	0.75
$\boldsymbol{\xi}_2$	2.5	0.13	0.16	0.08	0.76
$\boldsymbol{\xi}_{2.5}$	2.5	0.13	0.17	0.07	0.76
$\boldsymbol{\xi}_3$	2.5	0.12	0.18	0.06	0.76
$\boldsymbol{\xi}_{3.5}$	3	0.11	0.19	0.06	0.75
$\boldsymbol{\xi}_{KF}$	2.25	0.70	0.20	0.14	0.66
$\mu = -1, r = 3$					
$\boldsymbol{\xi}^*$		1.36	0.32	0.58	0.10
$\boldsymbol{\xi}_{\boldsymbol{x}^*}$	-1.5, 0.5	1	0.17	0.77	0.60
$\boldsymbol{\xi}_1$	0	0.37	0.41	0.41	0.19
$\boldsymbol{\xi}_{1.5}$	0	0.26	0.49	0.33	0.18
$\boldsymbol{\xi}_2$	0.5	0.20	0.56	0.28	0.16
$\boldsymbol{\xi}_{2.5}$	0.5	0.17	0.61	0.24	0.15
$\boldsymbol{\xi}_3$	0.5	0.13	0.64	0.22	0.14
$\boldsymbol{\xi}_{3.5}$	0.5	0.12	0.68	0.19	0.13
$\boldsymbol{\xi}_{KF}$	-0.25	0.35	0.62	0.34	0.04

Table 7.3: Efficiency and expected number of toxicity, disease failure and success for for canonical $(\mu, r) = (-1, 0.5) = (-3, 0.5) = (-1, 3)$.

$\mu = -3, r = 3$					
	Mode	$E(\boldsymbol{\xi}_{x^*}, \boldsymbol{\xi})$	$EF(x)$	$E(\bar{F}(x)\bar{G}(x))$	$EH(x)$
$\boldsymbol{\xi}^*$		1.25	0.27	0.48	0.25
$\boldsymbol{\xi}_{\boldsymbol{x}^*}$	-1.0, 1.0	1	0.12	0.76	0.13
$\boldsymbol{\xi}_1$	0.5	0.28	0.32	0.32	0.35
$\boldsymbol{\xi}_{1.5}$	0.5	0.20	0.40	0.27	0.33
$\boldsymbol{\xi}_2$	1.0	0.15	0.46	0.23	0.31
$\boldsymbol{\xi}_{.5}$	1.0	0.12	0.51	0.20	0.29
$\boldsymbol{\xi}_3$	1.0	0.10	0.55	0.18	0.27
$\boldsymbol{\xi}_{3.5}$	1.0	0.09	0.58	0.17	0.25
$\boldsymbol{\xi}_{KF}$	0.25	0.47	0.50	0.38	0.12
$\mu = -5, r = 3$					
$\boldsymbol{\xi}^*$		1.21	0.21	0.37	0.42
$\boldsymbol{\xi}_{\boldsymbol{x}^*}$	-1.0, 1.5	1	0.08	0.72	0.20
$\boldsymbol{\xi}_1$	1.0	0.23	0.24	0.24	0.52
$\boldsymbol{\xi}_{1.5}$	1.0	0.17	0.31	0.20	0.49
$\boldsymbol{\xi}_2$	1.0	0.13	0.36	0.18	0.47
$\boldsymbol{\xi}_{2.5}$	1.5	0.11	0.40	0.16	0.44
$\boldsymbol{\xi}_3$	1.5	0.09	0.44	0.15	0.42
$\boldsymbol{\xi}_{3.5}$	1.5	0.08	0.47	0.13	0.40
$\boldsymbol{\xi}_{KF}$	0.75	0.61	0.36	0.37	0.27
$\mu = -7, r = 3$					
$\boldsymbol{\xi}^*$		1.18	0.15	0.27	0.58
$\boldsymbol{\xi}_{\boldsymbol{x}^*}$	-0.5, 1.5	1	0.05	0.67	0.28
$\boldsymbol{\xi}_1$	1.5	0.21	0.17	0.17	0.67
$\boldsymbol{\xi}_{1.5}$	1.5	0.16	0.22	0.14	0.64
$\boldsymbol{\xi}_2$	2.0	0.13	0.26	0.13	0.62
$\boldsymbol{\xi}_{2.5}$	2.0	0.10	0.29	0.12	0.59
$\boldsymbol{\xi}_3$	2.0	0.09	0.32	0.11	0.57
$\boldsymbol{\xi}_{3.5}$	2.0	0.08	0.35	0.10	0.55
$\boldsymbol{\xi}_{KF}$	1.25	0.73	0.23	0.31	0.46

Table 7.4: Efficiency and expected number of toxicity, disease failure and success for canonical $(\mu, r) = (-3, 3) = (-5, 3) = (-7, 3)$.

existence criteria are known explicitly (Silvapule, [35]). This is an open problem for the contingent response model. With a Bayesian prior, this problem is resolved. But one must select a prior treatment distribution, and updating the posterior model poses a significant computational burden on the practitioner. For both approaches optimizing the design criterion, for each subject, requires tinkering with optimization software. At this time, few practitioners will undertake such labor intensive approaches. However, it would be good to see the Bayesian approach developed.

Both Flournoy's [16] and Mungo, Zhu, and Rosenberger's [32] urn designs have the biased coin design of Durham and Flournoy [7] as a special degenerate case. Parameters in the urn models could be re-set so that treatments will cluster around optimal design points based on Durham and Flournoy [7] in the same way as is done in the biased coin design presented in Section 3.1 of this paper. Urn designs should be considered when full randomization is important for an application. However, with more randomization will come increased variation in the design, and hence also in the parameter estimates. So the efficiency loss in adopting an urn design in place of a biased coin design typically will be significant. A new approach to urn designs is needed if there are to be be useful completely randomized designs for the contingent response model.

Hardwick et al. [24] examine the performance of *directed walks* combined with smoothed shape constrained curve fitting techniques in the context of competing failures. They do not, however, adopt the assumption that observation of disease failures is contingent on toxicity. In this more general context the probability of success is not necessarily a unimodal function of dose. Directed walks are up–and–down designs in that the dose for the next subject is no more than one step away from the current dose. Various curve fitting techniques are used to fit the $P(Y_1 = 1|x)$ and $P(Y_2 = 1|x)$. Then the next dose is one step in the direction of the dose with highest estimated probability of success. If this prescribes repeating the same dose, exploration is under taken with decreasing probability as the experiment continues. Directed walks are not Markovian, and therefore, characterizations are simulation not theoretical. Directed walks may be considered generalizations of the Kpamegan-Flournoy procedure; they too attempt to cluster treatments around the optimal dose rather than the optimal design points.

Hardwick et al. evaluated 8 varieties of directed walks, the Kpamegan-Flournoy procedure and equal allocation assuming three different underlying response models. In addition to not assuming contingent responses, the estimation procedure they used was not maximum likelihood, and therefore, their results are not directly comparable with the results presented here. Also, their evaluation criterion for estimation degenerates in the case of asymptotic maximum likelihood estimation. However, it is worth noting that they found equal allocation to perform very badly for all models. Using their evaluation criteria, all procedures except equal allocation do a good job of estimating $g(\mathbf{\Theta})$ with most performing a bit better than the Kpamegan-Flournoy procedure. In addition, the directed walk procedures tend to cluster treatments some-what more closely around the optimal dose. Thus it would be very interesting to see a simulated evaluation of the directed walk procedures incorporating the contingent response assumption compared to the c-optimal design.

Chapter 8

FUTURE WORK

In this dissertation D- and c-optimal designs for the contingent response model were found. Since no closed form was found for the optimal designs, there is a need to investigate other contingent response models. The probit is of interenst. It is also of interest to extend what is known about the location-scale contingent response models to more complex settings. For example, the four parameter logistic models are models of interest.

Limiting optimal designs for the positive–negative extreme value model were found and proved to be efficient. Fan and Chaloner findings for the continuation-ratio model were almost similar with the finding of the positive–negative extreme value model studied in this dissertation. I want to look more for common properties in these optimal designs and make more generalizations for the contingent response model. One goal is to see if there are closed forms for the limiting optimal designs for these new models.

As it is well known, the frequentist approach assumes that $\boldsymbol{\Theta}$, the parameters underlying the model, is fixed. For nonlinear models Fisher's information matrix depends on Θ. That means we have to estimate or guess $\boldsymbol{\Theta}$, which could be a disaster if our estimate is far away from the true value of $\boldsymbol{\Theta}$. One way to solve this problem is to put a prior distribution on $\boldsymbol{\Theta}$, and average the optimality criterion of interest over the prior distribution. I am very interested in studying the Bayesian optimal designs for the contingent response model, in particular for the positive–negative extreme value model.

Another idea of some interest and value is to find restricted optimal designs. Restricted optimal designs allow a compromise between optimality and some constraints, like sample size, ethical concerns, design region, cost and others.

We saw in Chapter 4 and 5 that optimal designs have large probability of toxicity and small probability of efficacy. From ethical viewpoint patients in clinical trials should not be assigned to highly toxic or non-effective doses. In clinical trials usually the level of toxicity and effective doses are predetermined by the clinicians in pharmaceutical companies, physicians and other agencies. Restricted optimal designs takes into consideration the maximum tolerated dose and the minimum effective dose. That means find the optimal designs within an interval between these doses.

A challenging and open problem that still needs to be addressed is how to implement these optimal designs.

APPENDIX A

MAPLE SOFTWARE CODE

A.1 Code for Theorem 5.1

1) Find the objective function for $\Theta = (\alpha_2, \beta_2, \mu, r)$.

The location-scale parameters are denoted by (a1, a2) and (b1, b2), respectively. c[i] is the weight for the design point x[i].

Define Fisher's Information (M)

```
> M := matrix([[Sum(c[i]*v[i],i = 1 .. k),
Sum(c[i]*v[i]*x[i]/b2-a2*c[i]*v[i]/b2,i = 1 .. k), 0, 0],
[Sum(c[i]*v[i]*x[i]/b2-a2*c[i]*v[i]/b2,i = 1 .. k),
Sum(c[i]*v[i]*(x[i]^2-2*a2*x[i]+a2^2)/b2^2,i = 1 .. k), 0, 0], [0, 0,
Sum(c[i]*w[i],i = 1 .. k), Sum(c[i]*w[i]*x[i]/b2-a2*c[i]*w[i]/b2,i = 1
.. k)], [0, 0, Sum(c[i]*w[i]*x[i]/b2-a2*c[i]*w[i]/b2,i = 1 .. k),
Sum(c[i]*w[i]*(x[i]^2-2*a2*x[i]+a2^2)/b2^2,i = 1 .. k)]])
```

Find the inverse matrix of M

```
> MMI:=inverse(M);
```

Simplify the inverse matrix

```
> MI:=simplify(MMI);
```

Define the gradient vector gdot and its transpose

```
gt := matrix([[-1/b2/(r+1),
-1/b2^2/r/(r+1)+(ln(r)+u+a2*(r+1))/b2^2/(r+1)^2, -1/b2/(r+1),
1/b2^2/(r+1)+(ln(r)+u+a2*(r+1))/b2^2/(r+1)^2]])
```

```
g := matrix([[-1/b2/(r+1)],
[-1/b2^2/r/(r+1)+(ln(r)+u+a2*(r+1))/b2^2/(r+1)^2], [-1/b2/(r+1)],
[1/b2^2/(r+1)+(ln(r)+u+a2*(r+1))/b2^2/(r+1)^2]])
```

Define the objective function (the c optimality criteria)

```
> objj:=multiply(gt,MI,g);
> Simplify the objective function
> phi:=simplify(objj);
```

2) Find the objective function for $\Theta = (0, 1, \mu, r)$.

Define Fisher's Information (M_0)

```
M0 := matrix([[Sum(c[i]*v[i],i = 1 .. k), Sum(c[i]*v[i]*x[i],i = 1 ..
k), 0, 0], [Sum(c[i]*v[i]*x[i],i = 1 .. k), Sum(c[i]*v[i]*x[i]^2,i = 1
.. k), 0, 0], [0, 0, Sum(c[i]*w[i],i = 1 .. k), Sum(c[i]*w[i]*x[i],i =
1 .. k)], [0, 0, Sum(c[i]*w[i]*x[i],i = 1 .. k),
Sum(c[i]*w[i]*x[i]^2,i = 1 .. k)]])
```

Find the inverse matrix

```
> MMI0:=inverse(M0);
```

Simplify the inverse matrix

```
> MI0:=simplify(MMI0);
```

Define the gradient vector gdot for theta (0, 1, u, r)

```
gt0 := matrix([[-1/(r+1), -1/r/(r+1)+(ln(r)+u)/(r+1)^2, -1/(r+1),
1/(r+1)+(ln(r)+u)/(r+1)^2]])

g0 := matrix([[-1/(r+1)], [-1/r/(r+1)+(ln(r)+u)/(r+1)^2], [-1/(r+1)],
[1/(r+1)+(ln(r)+u)/(r+1)^2]])
```

Define the objective function (the c optimality criterion)

```
> cc:= multiply(gt0,MI0,g0);

> Simplify the objective function

> phi0:= simplify(cc);
```

A.2 Code for Theorem 5.2

> **1) Find the objective function for** $\Theta = (\alpha_2, \beta_2, \mu)$.

Define Fisher's information (M)

```
M := matrix([[Sum(c[i]*v[i],i = 1 .. k),
Sum(c[i]*v[i]*x[i]/b-a*c[i]*v[i]/b,i = 1 .. k), 0],
[Sum(c[i]*v[i]*x[i]/b-a*c[i]*v[i]/b,i = 1 .. k),
```

```
Sum(c[i]*(v[i]+w[i])*(x[i]^2-2*a*x[i]+a^2)/b^2,i = 1 .. k),
Sum(c[i]*w[i]*x[i]/b-a*c[i]*w[i]/b,i = 1 .. k)], [0,
Sum(c[i]*w[i]*x[i]/b-a*c[i]*w[i]/b,i = 1 .. k), Sum(c[i]*w[i],i = 1 ..
k)]])
```

Find the inverse of M

```
> MI:=inverse(M);
```

Define the gradient vector gdot

```
gt := matrix([[-1/2/b, 1/2*(2*a+u)/b^2, -1/2/b]])
```

```
g := matrix([[-1/2/b], [1/2*(2*a+u)/b^2], [-1/2/b]])
```

Simplify the inverse matrix of M

```
> Ms:= simplify(MI);
```

Find the objective function (the c optimality criterion)

```
> obj:=multiply(gt,MI,g);
```

```
> copt:= simplify(obj);
```

2) Find the objective function for $\Theta = (0, 1, \mu)$.

Define Fisher's information (M0)

```
M0 := matrix([[Sum(c[i]*v[i],i = 1 .. k), Sum(c[i]*v[i]*x[i],i = 1 ..
k), 0], [Sum(c[i]*v[i]*x[i],i = 1 .. k), Sum(c[i]*(v[i]+w[i])*x[i]^2,i
= 1 .. k), Sum(c[i]*w[i]*x[i],i = 1 .. k)], [0, Sum(c[i]*w[i]*x[i],i =
1 .. k), Sum(c[i]*w[i],i = 1 .. k)]])
```

Find the inverse of M0

```
>  MM0:=inverse(M0);
```

Simplify the inverse of M0

```
>  Ms0:=simplify(MM0);
```

Define the gradient vector gdot and its transpose

```
gt0 := matrix([[-1/2, 1/2*u, -1/2]])
```

```
g0 := matrix([[-1/2], [1/2*u], [-1/2]])
```

Find the objective function (the c optimality criterion)

```
>  obj0:=multiply(gt0,Ms0,g0);

>  Simplify the objective function

>  copt0:=simplify(obj0);
```

A.3 Code of the trace function for the example used in Chapter 6 (Section 1)

Define Fisher's information components at the optimal design points x1,x2.

```
x1:=convert(1.3377, fraction)

>  x2:=convert(-.9796,fraction);

>  f1:= exp(-2*x1)*exp(-exp(-x1))/(1-exp(-exp(-x1)));

>  f2:= exp(-2*x2)*exp(-exp(-x2))/(1-exp(-exp(-x2)));
```

```
f1 :=
exp(-13377/5000)*exp(-exp(-13377/10000))/(1-exp(-exp(-13377/10000)))


f2 := exp(2449/1250)*exp(-exp(2449/2500))/(1-exp(-exp(2449/2500)))
```

Mlo is Fisher's information for Gbar(x) at x1,x2

```
> Mlo:=array([[.5*f1+.5*f2,.5*x1*f1+.5*x2*f2], [.5*x1*f1+.5*x2*f2,.5*x1^
2*f1+.5*x2^2*f2]]);
```

Mup is Fisher's information for F(u + r x) at (-x1-u)/r,(x2-u)/r

```
> Mup:=array([[.5*f1+.5*f2,.5*((-x1-u)/r)*f1+.5*((-x2-u)/r)*f2],[.5*((-
x1-u)/r)*f1+.5*((-x2-u)/r)*f2,.5*((-x1-u)/r)^2*f1+.5*((-x2-u)/r)^2*f2]
]);
```

Define Fisher's information for a single point

```
> f:=exp(2*r*x+2*u)*exp(-exp(r*x+u))/(1-exp(-exp(r*x+u)));

> Iup:=array([[f,x*f],[x*f,x^2*f]]);


f := exp(4*x-50)*exp(-exp(2*x-25))/(1-exp(-exp(2*x-25)))


Iup := matrix([[exp(4*x-50)*exp(-exp(2*x-25))/(1-exp(-exp(2*x-25))),
x*exp(4*x-50)*exp(-exp(2*x-25))/(1-exp(-exp(2*x-25)))],
[x*exp(4*x-50)*exp(-exp(2*x-25))/(1-exp(-exp(2*x-25))),
x^2*exp(4*x-50)*exp(-exp(2*x-25))/(1-exp(-exp(2*x-25)))]])

      g:=exp(-2*x)*exp(-exp(-x))/(1-exp(-exp(-x)))

> Ilo:=array([[g,x*g],[x*g,x^2*g]]);
```

```
g := exp(-2*x)*exp(-exp(-x))/(1-exp(-exp(-x)))


Ilo := matrix([[exp(-2*x)*exp(-exp(-x))/(1-exp(-exp(-x))),
x*exp(-2*x)*exp(-exp(-x))/(1-exp(-exp(-x)))],
[x*exp(-2*x)*exp(-exp(-x))/(1-exp(-exp(-x))),
x^2*exp(-2*x)*exp(-exp(-x))/(1-exp(-exp(-x)))]])
```

Find the inverse of Mup, Mlo

```
> MIup:=inverse(Mup);
> MIlo:=inverse(Mlo);
```

Find the trace: Tr(I_F* M^-1_F) and Tr(I_G *M^-1_G when u = -10, r =2

```
> u:-10;
> tup:=convert(trace(multiply(Iup,MIup)),float);


tup :=
1487.575983*exp(4.*x-50.)*exp(-1.*exp(2.*x-25.))/(1.-1.*exp(-1.*exp(2.
*x-25.)))-234.9528400*x*exp(4.*x-50.)*exp(-1.*exp(2.*x-25.))/(1.-1.*ex
p(-1.*exp(2.*x-25.)))+9.293731502*x^2*exp(4.*x-50.)*exp(-1.*exp(2.*x-2
5.))/(1.-1.*exp(-1.*exp(2.*x-25.)))
> tlo:=convert(trace(multiply(Ilo,MIlo)),float);


tlo :=
2.811031575*exp(-2.*x)*exp(-1.*exp(-1.*x))/(1.-1.*exp(-1.*exp(-1.*x)))
+1.304776227*x*exp(-2.*x)*exp(-1.*exp(-1.*x))/(1.-1.*exp(-1.*exp(-1.*x
)))+2.323432836*x^2*exp(-2.*x)*exp(-1.*exp(-1.*x))/(1.-1.*exp(-1.*exp(
-1.*x)))
```

Find the trace: Tr(I_F* M^-1_F) and Tr(I_G *M^-1_G when u =-15, r =2

```
> u:=-15;

> tup:=convert(trace(multiply(Iup,MIup)),float);

> tlo:=convert(trace(multiply(Ilo,MIlo)),float);
```

Find the trace: Tr(I_F* M^-1_F) and Tr(I_G *M^-1_G when u =-25, r =2

```
> u:=-25;

> tup:=convert(trace(multiply(Iup,MIup)),float);

> tlo:=convert(trace(multiply(Ilo,MIlo)),float);
```

A.4 Code of the Trace function of Theorem 6.2

Define the optimal design weights

```
> a1:=convert(.2895,fraction);

> a2:=convert(.2105,fraction);

> a3:=convert(.2895,fraction);

> a4:=convert(.2105,fraction);
```

Define the optimal design points

```
> x1:= convert(-.8537,fraction);

> x2:=convert(1.0773,fraction);
```

Define Fisher's information MM at the optimal design points

```
> f1:=exp(-2*x1)*exp(-exp(-x1))/(1-exp(-exp(-x1)));

> f2:=exp(-2*x2)*exp(-exp(-x2))/(1-exp(-exp(-x2)));
```

```
> M1b:=array([[f2*a4,a4*f2*(-x2-u),0],[a4*f2*(-x2-u),a4*f2*(-x2-u)^2+a2
*f2*x2^2,a2*f2*x2],[0,a2*f2*x2,a2*f2]]);
```

```
> M1a:=array([[f1*a3,a3*f1*(-x1-u),0],[a3*f1*(-x1-u),a3*f1*(-x1-u)^2+a1
*f1*x1^2,a1*f1*x1],[0,a1*f1*x1,a1*f1]]);
```

```
> MM:=evalm(M1a+M1b);
```

Find the inverse of MM

```
> MI:=convert(inverse(MM),float);
```

Define v_x = ff1

```
> ff1:=exp(2*x+2*u)*exp(-exp(x+u))/(1-exp(-exp(x+u)));
```

Define w_x = ff2

```
> ff2:=exp(-exp(x+u))*exp(-2*x)*exp(-exp(-x))/(1-exp(-exp(-x)));
```

Define Fisher's information for a single pint

```
> FI:=array([[ff1,ff1*x,0],[ff1*x,ff1*x^2+ff2*x^2,ff2*x],[0,ff2*x,ff2]]
);
```

Define the trace function Tr(I M^-1)

```
> tt:=trace(multiply(FI,MI));
```

APPENDIX B

S-plus Code

B.1 The Directional Derivatives for D-optimality Criterion

Define function F(negative extreme value)and return Fbar.

```
 ff<-function(x,d1,b1)
 {
 f<-exp(-exp(b1*x+d1)) return(f)}
```

Define function G(positive extreme value) and return G

```
 gg<-function(x,d,b) {
  g<-exp(-exp(-b*x-d))
return(g) }
```

Define Fisher matrix:(a: the weights vector, y: the design points vector)

```
 FM<-function(a,y,d1,b1,d,b){
   s11<-NULL
   s12<-NULL
   s22<-NULL
   s33<-NULL
   s34<-NULL
   s44<-NULL
   k<-NULL
   h<-NULL
for(i in 1:length(y)) { k[i]<-ff(y[i],d1,b1)
  h[i]<-gg(y[i],d,b)
  s1<-NULL
  s1[i]<-((a[i])*(exp(2*(b1*y[i]+d1)))*(k[i]))/(1- k[i])
  s11<-c(s11,s1[i])
  s11<-sum(s11)
  s13<-NULL
  s13[i]<-((a[i])*(y[i])*(exp(2*(b1*y[i]+d1)))*(k[i]))/(1-
   k[i])
  s12<-c(s12,s13[i])
  s12<-sum(s12)
  s2<-NULL
  s2[i]<-((a[i])*(y[i]^2)*(exp(2*(b1*y[i]+d1)))*(k[i]))/(1-
   k[i])
  s22<-c(s22,s2[i])
  s22<-sum(s22)
  s3<-NULL
  s3[i]<-((a[i])*(exp(-2*(b*y[i]+d)))*(k[i])*(h[i]))/(1- h[i])
  s33<-c(s33,s3[i])
  s33<-sum(s33)
  s31<-NULL
  s31[i]<-((a[i])*(y[i])*(exp(-2*(b*y[i]+d)))*(k[i])*(h[i]))/(1-
   h[i])
  s34<-c(s34,s31[i])
  s34<-sum(s34)
  s4<-NULL
  s4[i]<-((a[i])*(y[i]^2)*(exp(-2*(b*y[i]+d)))*(k[i])*(h[i]))/(1-
   h[i])
  s44<-c(s44,s4[i])
  s44<-sum(s44)}
  A<-matrix(ncol=4,nrow=4)
  A[1,]<-c(s11,s12,0,0)
  A[2,]<-c(s12,s22,0,0)
```

```
  A[3,]<-c(0,0,s33,s34)
  A[4,]<-c(0,0,s34,s44)
  B<-solve(A) return(B) }
```

BB is the inverse of Fisher information evaluated at the optimal design points.

```
BB<-FM(a,y,d1,b1,d,b)
```

Define Fisher information for a single point y.

```
MF1<-function(y,d1,b1,d,b) {
  f<-exp(-exp(b1*y+d1))
  g<-exp(-exp(-b*y-d))
  s11<-exp(2*(b1*y+d1))*f/(1-f)
  s12<-y*exp(2*(b1*y+d1))*f/(1-f)
  s22<-(y^2)*exp(2*(b1*y+d1))*f/(1-f)
  s33<-((exp(-2*(b*y+d)))*(f)*(g))/(1-g)
  s34<-((y)*(exp(-2*(b*y+d)))*(f)*(g))/(1-g)
  s44<-((y^2)*(exp(-2*(b*y+d)))*(f)*(g))/(1-g)
  A<-matrix(ncol=4,nrow=4)  A[1,]<-c(s11,s12,0,0)
  A[2,]<-c(s12,s22,0,0)  A[3,]<-c(0,0,s33,s34) A[4,]<-c(0,0,s34,s44)
  return(A)}
```

Define the directional derivative: (tr(I*M^-1)-p, p is the no. of parameters.

I is fisher for a single point and tr is the trace function.

Define the trace of 2 matrices -p.

```
tr<-function(I,A) {
  p<- 4
  d<- sum(diag(I *A))-p
  return(d)}
```

Find the directional derivative for different values of x.

```
 dir<-function(B) {
    tx<- seq(-15,15,.01)
```

```
    dd<- NULL
    dd_c(1:length(tx))
    for(i in 1:length(tx))
    dd[i]<-tr(MF1(tx[i],d1,b1,d,b),B)
    return(dd)}
```

The directional derivative at the inverse of Fisher's information evaluated at the optimal design points.

```
y<-dir(BB)
plot(  tx,y,type="l", xlab="x", ylab  ="Directional Derivative")
abline( h = 0 )
```

B.2 The Directional derivative for c-Optimality Criterion

Define function F(negative extreme value)and return Fbar.

```
 ff<-function(x,d1,b1)
   {f<-exp(-exp(b1*x+d1))
   return(f)}
```

Define function G(positive extreme value) and return G

```
 gg<-function(x,d,b) {
   g<-exp(-exp(-b*x-d))
   return(g)}
```

Define Fisher information matrix at(a: the weights vector, y: the design point vector)

```
 FM<-function(a,y,d1,b1,b,d){
   s11<-NULL
   s12<-NULL
   s22<-NULL
```

```
 s33<-NULL
 s34<-NULL
 s44<-NULL
 k<-NULL
 h<-NULL
for(i in 1:length(y))}
k[i]<-ff(y[i],d1,b1)
h[i]<-gg(y[i])
s1<-NULL
s1[i]<-((a[i])*(exp(2*(b1*y[i]+d1)))*(k[i]))/(1- k[i])
s11<-c(s11,s1[i])
s11<-sum(s11)
s13<-NULL
s13[i]<-((a[i])*(y[i])*(exp(2*(b1*y[i]+d1)))*(k[i]))/(1-
        k[i])
s12<-c(s12,s13[i])
s12<-sum(s12)
s2<-NULL
s2[i]<-((a[i])*(y[i]^2)*(exp(2*(b1*y[i]+d1)))*(k[i]))/(1-
        k[i])
s22<-c(s22,s2[i])
s22<-sum(s22)
s3<-NULL
s3[i]<-((a[i])*(exp(-2*(b*y[i]+d)))*(k[i])*(h[i]))/(1-
       h[i])
s33<-c(s33,s3[i])
s33<-sum(s33)
s31<-NULL
s31[i]<-((a[i])*(y[i])*(exp(-2*(b*y[i]+d)))*(k[i])*(h[i]))/(1-
        h[i])
s34<-c(s34,s31[i])
s34<-sum(s34)
s4<-NULL
s4[i]<-((a[i])*(y[i]^2)*(exp(-2*(b*y[i]+d)))*(k[i])*(h[i]))/(1-
       h[i])
s44<-c(s44,s4[i])
s44<-sum(s44)}
A<-matrix(ncol=4,nrow=4)
A[1,]<-c(s11,s12,0,0)
A[2,]<-c(s12,s22,0,0)
A[3,]<-c(0,0,s33,s34)
A[4,]<-c(0,0,s34,s44) B<-solve(A)
```

```
return(B) }
```

BB is the inverse of Fisher's information matrix at the optimal designs points.

```
BB<-FM(a,y,d1,b1)
```

Return the transpose of a matrix.

```
tran<-function(A,n,m)
    {B_matrix(nrow=m,ncol=n)
    for(i  in 1:m )
    for (j  in 1:n)
    B[i,j]<-A[j,i]
    return(B)}
```

Define Fisher information for a single point.

```
MF1<-function(y,d1,b1,b,d) {
   f<-exp(-exp(b1*y+d1))
   g<-exp(-exp(-b*y-d))
   s11<-exp(2*(b1*y+d1))*f/(1-f)
   s12<-y*exp(2*(b1*y+d1))*f/(1-f)
   s22<-(y^2)*exp(2*(b1*y+d1))*f/(1-f)
   s33<-((exp(-2*(b*y+d)))*(f)*(g))/(1-g)
   s34<-((y)*(exp(-2*(b*y+d)))*(f)*(g))/(1-g)
   s44<-((y^2)*(exp(-2*(b*y+d)))*(f)*(g))/(1-g)
   A<-matrix(ncol=4,nrow=4) A[1,]<-c(s11,s12,0,0)
   A[2,]<-c(s12,s22,0,0)
   A[3,]<-c(0,0,s33,s34)
   A[4,]<-c(0,0,s34,s44)
   return(A) }
```

The directional derivative for a seq of x values.

```
dir<-function(B,d1,b1,b,d) {
   x<- seq(-8,8,.01)
   dd<- NULL
   gc_matrix(nrow=4,ncol=1)
```

```
    gc[1]<- -1/(b1+1)
    gc[2]<-  -1/(b1*(b1+1))- ((-log(b1)-d1)/(b1+1)^2)
    gc[3]<- -1/(b1+1)
    gc[4]<- (1/(b1+1))- ((-log(b1)-d1)/(b1+1)^2)
    gct<-tran(gc,4,1)
        for(i  in 1:length(x))}
 Ix<-MF1(x[i],d1,b1,b,d)
 dd[i]<-gct*B*Ix*B*gc - gct*B*gc}
 return(dd)}


y<-dir(BB,d1,b1,b,d)
x<-seq(-8,8,.01)

plot(x,y,type="l",ylab="Directional Derivative")
 abline(h=0)
```

B.3 Up–and–Down Procedure

```
 A program to return an estimate of the optimal designs using up
and down procedure.
The optimal designs for the contingent
response model when mu =-3, r=2.
 1) x_1^*=-0.05054 , x_2^*=1.3595, with w_1=0.4444  and  w_2=0.5556

Define the negative-positive extreme value function

 ffu<-function(x,u,r) {
    f<-1- exp(-exp(r*x+u))
    return(f)}

 b3<-ffu(-.5054,-3,2)
 ggu<-function(x,u,r){
    g<-(1-exp(-exp(-x)))*(exp(-exp(r*x+u)))
    return(g) }

 b1<-ggu(-.5054,-3,2)
 r<- 2
 u<--3
```

```
 rp<- NULL
```

Define

```
 Pi_i/Pi_1 p<-b3*ggu(x[1],mu,r )/(b1*ffu(x[2],mu,r) )
 rp<-c(p,rp)for(i in 2: (length(x)-1)  )  {
 w_rp[i-1]*b3*ggu(x[i],mu,r)/(b1*ffu(x[i+1],mu,r)) rp_c(rp,w)}
 sum1<-sum(rp) p1<-1/(1+sum1)
 pp<-NULL
 pp_<-c(p1,pp)
 for(i in 1: length(rp)  ) {
 pps<-p1*rp[i]
 pp<-c(pp,pps)}
```

Repeat for the second optimal point.

```
 b3<-ffu(1.3595,-3,2)
 b1<-ggu(1.3595,-3,2)
 x<-seq(-3,6,.25)
 rp2<-NULL
 p2<-b3*ggu(x[1],mu,r)/(b1*ffu(x[2],mu,r))
 rp2<-c(p2,rp2)
 for(i in 2 :(length(x)-1))     {
 w1<-rp2[i-1]*b3*ggu(x[i],mu,r)/(b1*ffu(x[i+1],mu,r))
 rp2<-c(rp2,w1)}
 sum2<-sum(rp2)
 p11<-1/(1+sum2)
 pp1<-NULL
 pp1<-c(p11,pp1)
 for(i in 1: length(rp2)  ) {
 pps<-p11*rp2[i]
 pp1<-c(pp1,pps) }
```

Efficiency of the designs

Define the components functions of Fisher's information.

```
ff<-function(x,b1,d1) {
  f<- exp(-exp(b1*x+d1))
  return(f) }
gg<-function(x){
```

```
  g<- exp(-exp(-x))
  return(g) }
```

Define Fisher's information for a vector of x^* and w^*.

```
 FM<-function(a,y,b,d,b1,d1){
   s11<-NULL
   s12<-NULL
   s22<-NULL
   s33<-NULL
   s34<-NULL
   s44<-NULL
   k<-NULL
   h<-NULL
   for(i in 1:length(y)){
   k[i]<-ff(y[i],b1,d1)
   h[i]<-gg(y[i])
```

Define the components of Fishers M

```
   s1<-NULL
   s1[i]<-((a[i])*(exp(2*(b1*y[i]+d1)))*(k[i]))/(1-
        k[i])
   s11<-c(s11,s1[i])
   s11<-sum(s11)
   s13<-NULL
   s13[i]<-((a[i])*(y[i])*(exp(2*(b1*y[i]+d1)))*(k[i]))/(1-
         k[i])
   s12<-c(s12,s13[i])
   s12<-sum(s12)
   s2<-NULL
   s2[i]<-((a[i])*(y[i]^2)*(exp(2*(b1*y[i]+d1)))*(k[i]))/(1-
        k[i])
   s22<-c(s22,s2[i])
   s22<-sum(s22)
   s3<-NULL
   s3[i]<-((a[i])*(exp(-2*(b*y[i]+d)))*(k[i])*(h[i]))/(1-
        h[i])
   s33<-c(s33,s3[i])
   s33<-sum(s33)
   s31<-NULL
   s31[i]<-((a[i])*(y[i])*(exp(-2*(b*y[i]+d)))*(k[i])*(h[i]))/(1-
         h[i])
```

```
   s34<-c(s34,s31[i])
   s34<-sum(s34)
   s4<-NULL
   s4[i]<-((a[i])*(y[i]^2)*(exp(-2*(b*y[i]+d)))*(k[i])*(h[i]))/(1-
        h[i])
   s44<-c(s44,s4[i])
   s44<-sum(s44)}
   A<-matrix(ncol=4,nrow=4)
   A[1,]<-c(s11,s12,0,0)
   A[2,]<-c(s12,s22,0,0)
   A[3,]<-c(0,0,s33,s34)
   A[4,]<-c(0,0,s34,s44)
   return(A) }
```

The design parameters.

```
   b<-1
   d<-0
   b11<-2
   d1<--3
```

Return the transpose of a matrix.

```
 tran<-function(A,n,m){
     B_matrix(nrow=m,ncol=n)
     for(i in 1:m )
     for (j in 1:n)
     B[i,j]<-A[j,i]
     return(B) }
```

Define the gradient vector.

```
gc_matrix(nrow=4,ncol=1)
gc[1]<- -1/(b11+1)
gc[2]<-  -1/(b11*(b11+1))- ((-log(b11)-d1)/(b11+1)^2)
gc[3]<- -1/(b11+1)
gc[4]<- (1/(b11+1))- ((-log(b11)-d1)/(b11+1)^2)
gct<-tran(gc,4,1)
```

Find the variance of g(Theta) when (mu=-3,r=2) for the c-optimal design

```
a <-c(.4444,.5556)
y<-c(-.5054,1.3595)
```

```
BB<-FM(a,y,b,d,b11,d1)
BBo<-solve(BB)
varo<-gct*BBo*gc
```

```
 Find the variance of Up and  Down design.
 Sum  pi I_x  over the grid points and the pi's are (pp) and (pp1)

  a1<- a[1]*pp
  a2<-a[2]*pp1
  yy<-x
  aaa_a1+a2

 Evaluate Fisher's information at the up-and-down design and
 find the inverse.

 BB2<-FM(aaa,yy,b,d,b11,d1)
 BBd<-solve(BB2)
 varupx<-gct*BBd*gc
```

APPENDIX C

FORTRAN CODE

Optimal designs were found using NPSOL which is a fortran program to minimize a multivariate function. For details about the procedure see [19]. The program was run for each (mu, r) or \mu) model assuming the optimal design consist of 2, 3, 4,5,... until the candidate design is verified by the General Equivalence Theorem. The following is a Fortran program for minimizing log-determinant of Fisher's information assuming optimal designs consist of two design points .

```
      program det2
      implicit double precision   (a-h, o-z)
%*     ================================================================
*     Set the declared array dimensions.
*     ldA    = the declared leading dimension of  A.
*     ldcJ   = the declared leading dimension of  cJac.
*     ldR    = the declared leading dimension of  R.
*     maxn   = maximum no. of variables allowed for.
*     maxbnd = maximum no. of variables + linear \& nonlinear
constrnts.
*     liwork = the length of the integer work array.
*     lwork  = the length of the double precision work array.
%*     ================================================================
    parameter         (ldA = 1, ldcJ= 1, ldR= 4 , maxn=4,
                    liwork = 30, lwork= 150, maxbnd=
     maxn+ldA+ldcJ)
      integer           istate(maxbnd)
    integer           iwork(liwork)
    double precision  A(ldA,maxn)
```

```
      double precision  bl(maxbnd), bu(maxbnd)
      double precision  c(ldcJ), cJac(ldcJ,maxn), clamda(maxbnd)
      double precision  objgrd(maxn), R(ldR, maxn), x(maxn)
      double precision  work(lwork)
      external          funobj, funcon
      double precision  bigbnd
      character*20      lFile
      logical           byname, byunit
      parameter         (zero=0.0d+0, one=1.0d+0)
%*     ------------------------------------------------------------------
*     Assign file numbers and open files by various means.
*     (Some systems don't need explicit open statements.)
*     iOptns = unit number for the Options file.
*     iPrint = unit number for the Print file.
*     iSumm  = unit number for the Summary file.
%*     ------------------------------------------------------------------
        iOptns =4
      iPrint = 10
      iSumm=6
      byname = .true.
      byunit = .false.
      if ( byname ) then
           lFile = 'det2.opt'
           open( iOptns, file=lFile, status='OLD',     err=800 )
           lFile = 'det2.out'
           open( iPrint, file=lFile, status='UNKNOWN', err=800 )
        else if ( byunit ) then
           lUnit = iOptns
           open( lUnit, status='OLD',     err=900 )
           lUnit = iPrint
           open( lUnit, status='UNKNOWN', err=900 )
        end if
%*     =============================================================
*     Set the actual problem dimensions.
*     n      = the number of variables.
*     nclin  = the number of general linear constraints (may be
0).
*     ncnln  = the number of nonlinear constraints (may be 0).
%*     =============================================================
        n      = 4
        nclin  = 1
        ncnln  = 0
```

```
      nbnd   = n + nclin + ncnln
*Assign the data arrays.
*     A      = the linear constraint matrix.
*     bl     = the lower bounds on  x,  a'x  and  c(x).
*     bu     = the upper bounds on  x,  a'x  and  c(x).
*     bounds  .ge.    bigbnd  will be treated as plus  infinity.
*     bounds  .le.  - bigbnd  will be treated as minus infinity.
*     x      = the initial estimate of the solution.
%*     ------------------------------------------------------------------
       bigbnd = 1.0d+21
      A(1,1)= zero
    A(1,2)=zero
    A(1,3)= one
      A(1,4) = one
*     set the bounds
      bl(1) = -bigbnd
      bl(2) = -bigbnd
    bl(3) = zero
    bl(4) = zero
    bl(5) = one
    bu(1) = bigbnd
    bu(2) = bigbnd
      bu(3) = one
    bu(4) = one
    bu(5)= one
   Set the initial estimate of  X.(to be checked later)
      x(1)   = -1.25
    x(2)   =  1.05
      x(3)   = .5
      x(4)   = .5
%*     ------------------------------------------------------------------
   Set a few options in-line.
*     The Print file   will be on unit iPrint.
*     The Summary file will be on the default unit 6
*     (typically the screen).
*     ------------------------------------------------------------------
      call npopti( 'Print file          =', iPrint )
      call npoptr( 'Infinite Bound size =', bigbnd )
*     Read the Options file.
      call npfile( iOptns, inform )
      if (inform .ne. 0) then
         write(iPrint, 3000) inform
```

```
        stop
      end if
%-------------------------------------------------------------------
   Solve the problem.
  %   ------------------------------------------------------------------
      call npsol ( n, nclin, ncnln, ldA, ldcJ, ldR,
                  A, bl, bu,
                    funcon, funobj,
                    inform, iter, istate,
                   c, cJac, clamda, objf, objgrd, R, x,
                iwork, liwork, work, lwork )
      if (inform .gt. 0) go to 999
      do 100, j = 1, n
          x(j)   = x(j) + 0.1
  100 continue
*     Set some new options in-line,
*     but stop listing them on the Print file.
      call npoptn( 'Nolist'                          )
      call npoptn( 'Derivative level              0' )
      call npoptn( 'Verify                       No' )
      call npoptn( 'Warm Start'                      )
      call npopti( 'Major iterations    ',      20  )
      call npopti( 'Major print level   ',      10  )

*     Error conditions.
  800 write(iSumm , 4000) 'Error while opening file', lfile
      stop
  900 write(iSumm , 4010) 'Error while opening unit', lunit
      stop
  999 write(iPrint, 3010) inform
      stop
 3000 format(/ ' npfile terminated with  inform =', i3)
 3010 format(/ ' npsol  terminated with  inform =', i3)
 4000 format(/  a, 2x, a  )
 4010 format(/  a, 2x, i6 )
*     end of the example program for NPSOL
      end
%*+++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

      subroutine funobj (mode, n, x,objf, objgrd, nstate)
    implicit double precision(a-h, o-z)
    double precision  x(n), objgrd(n)
```

```
   double precision , parameter::a1=0d0, d1=.5d0,
   a2=0d0,d2=1d0
F1  =1- exp(-exp(a1 + d1*x(1)))
Fb1 =1-F1
G1  =exp(-exp(-(a2+ d2*x(1))))
Gb1 =1-G1
  F2 =1- exp(-exp(a1 + d1*x(2)))
Fb2 =1-F2
G2  =exp(-exp(-(a2+ d2*x(2))))
Gb2 =1-G2
s11 = x(3)*exp(2*(a1+d1*x(1)))*(Fb1/F1) +
      x(4)*exp(2*(a1+d1*x(2)))*(Fb2/F2)
  s12 = x(3)*exp(2*(a1+d1*x(1)))*x(1)*(Fb1/F1) +
  x(4)*exp(2*(a1+d1*x(2)))*x(2)*(Fb2/F2)
s22 = x(3)*exp(2*(a1+d1*x(1)))*x(1)**2*(Fb1/F1) +
     x(4)*exp(2*(a1+d1*x(2)))*x(2)**2*(Fb2/F2)
s33 = x(3)*exp(-2*(a2+d2*x(1)))*Fb1*(G1/Gb1) +
     x(4)*exp(-2*(a2+d2*x(2)))*Fb2*(G2/Gb2)
s34 = x(3)*exp(-2*(a2+d2*x(1)))*x(1)*Fb1*(G1/Gb1) +
     x(4)*exp(-2*(a2+d2*x(2)))*x(2)*Fb2*(G2/Gb2)
s44 = x(3)*exp(-2*(a2+d2*x(1)))*x(1)**2*Fb1*(G1/Gb1) +
     x(4)*exp(-2*(a2+d2*x(2)))*x(2)**2*Fb2*(G2/Gb2)
dt1 = s11*s22 - s12**2
dt2 = s33*s44 - s34**2
objf= -log(dt1*dt2)

END SUBROUTINE  funobj
  subroutine funcon (mode, ncnln ,n, ldcJ, needc,x,c, cJac,
  nstate)
implicit double precision(a-h, o-z)
integer  needc (*)
double precision   x(n), c(*), cJac(ldcJ,*)
  END SUBROUTINE  funcon
```

BIBLIOGRAPHY

[1] Agresti, A.: *Categorical Data Analysis.* Wiley, New York (1990).

[2] Atkinson, A.C., Donev, A.N.: *Optimum Experimental Designs.* Oxford University Press, New York (1992).

[3] Atkinson, A.C., Haines, L.M.: *Designs for Nonlinear and Generalized Linear Models.* S. Ghosh and C.R. Rao, eds, Handbook of Statistics, 13 (1996).

[4] Chaloner, K., Verdinelli, I.: *Bayesian Experimental Design: A Review.* Statistical Science, 10, 273-304 (1995).

[5] Chernoff, H.: *Locally Optimal Designs for Estimating Parameters.* Annals of Mathematical Statistics, 24, 586-602 (1953).

[6] Derman, C.: *Non-Parametric Up-and-Down Experimentation.* Annals of Mathematical Statistics, 28, 795-799 (1957).

[7] Durham, S.D., Flournoy, N.: *Random Walks for Quantile Estimation*. In: Gupta and Berger (Eds.), Statistical Decision Theory and Related Topics, Springer-Verlag, 467-476 (1994).

[8] Durham, S.D., Flournoy, N.: *Up–and–Down Designs I: Stationary Treatment Distributions.* In: Adaptive Designs, IMS Lecture Notes-Monograph Series, 25, 139-157 (1995).

[9] Durham, S.D., Flournoy, N., Rosenberger, W. F.: *A Random Walk Rule for Phase I Clinical Trials.* Biometrics, 53, 745-760 (1997).

[10] Elfving, G: *Optimum Allocation in Linear Regression Theory.* Annals of Mathematical Statistics, 23, 255-262 (1952).

[11] Fan, S.K., Chaloner,K.: *Optimal Designs for a Continuation–Ratio Model*. In mODa6-Advances in Model-Oriented Design and Analysis, eds. Atkinson, A.C., Hackl, P., and Muller, W.G. Physica-Verlag, 77–86 (2001).

[12] Fan, S. K., Chaloner, K.: *Optimal Designs and Limiting Optimal Designs for a Trinomial Response.* Journal of Statistical Planning and Inference. In Press.

[13] Fan, S. K.: *Multivariate Optimal Designs.* Ph.D Thesis, School of Statistics, University of Minnesota (1999).

[14] Fedorov, V.V.: *Theory of Optimal Experiments.* translated and edited by Studden, W.J. and Klimko, E.M. Academic Press, New York (1974).

[15] Flournoy, N.: *Proposal Made to The Seattle Bone Marrow Transplant Team During The Preparation of the Adult Leukemia Center grant renewal.* (1985).

[16] Flournoy, N.: *A Biased Coin Urn Design for Estimating a Target Quantile from a Sequence of Binary Responses: asymptotic results.* Proceedings of the 20th International Conference on Information Technology Interfaces 13-18. Kalpic', D. and Dobric', V.H. eds. SRCE University Computing Centre University of Zagreb, Croatia (1998).

[17] Flournoy, N., Durham, S.D., Rosenberger, W.F.: *Toxicity in Sequential Dose-Response Experiments.* Sequential Analysis, 14, 217-228 (1985).

[18] Ford, I., Torsney, B., Wu, C.F.J.: *The Use of a Canonical Form in the Construction of Locally Optimal Designs for Non-Linear Problems.* Journal of the Royal Statistical Society Series B, 54, 569–583 (1992).

[19] Gill, P.E., Murray, W., Saunders, M.A., Wright, M.H.: *User's Guide For NPSOL 5:0 A Fortran Package For Nonlinear Programming.* Technical Report SOL 86–1. Department of EESOR Stanford University (1998).

[20] Giovagnoli, A., Pintacuda, N.: *Markovian Experiments that Approximate Optimal Designs for Quantal Response Curve.* Metron, LVI, 1-2 (1998a).

[21] Giovagnoli, A., Pintacuda, N.: *Markovian eExperiments that Approximate Optimal Designs for Quantal Response Curve.* Metron, 77-94 (1998b).

[22] Gooley, T.A., Martin, P.J., Fisher, L.D., Pettinger, M.: *Simulation as a Design Tool for Phase I/II Clinical Trials: An example from bone marrow transplantation.* Controlled Clinical Trials, 15, 450-460 (1994).

[23] Haines, L.M., Perevoskaya, I., Rosenberger, W.F.: *Bayesian Optimal Designs for Phase I Clinical Trials.* Biometrics, 59, 591-600 (2003).

[24] Hardwick, J., Meyer, M.C., Stout, Q.F.: *Directed Walk Designs for Dose Response Problems with Competing Failure Modes.* Biometrics, 59, 229-236 (2003).

[25] Hayes, G.J., Edie, D.D., Durham, S.D.: *Statistical Analysis of Recoil Test Data for Carbon Fibers.* Twentieth Biennial Conference on Carbon, 326–327 (1992).

[26] Karlin, S., Taylor, H.M.: *A First Course in Stochastic Processes.* New York, Academic Press (1975).

[27] Kiefer, J.: *General Equivalence Theory for Optimum Designs (Approximate Theory).* Annals of Statistics 2, 849–879 (1974).

[28] Kpamegan, E.E., Flournoy, N.: *An Optimizing Up–and–Down Design.* In Optimal Design 2000: Atkinson, A. et al. (eds.), Khwer Academic Publishers, 211-224 (2000).

[29] Li, Zhengqing, Durham, S.D., Flournoy, N.: *An Adaptive Design for Maximization of a Contingent Binary Response.* IMS Lecture Notes-Monograph Series, 25, 179–196 (1995).

[30] McLeish, D.L., Tosh, D.: *Sequential Designs in Bioassay.* Biometrics, 46, 103-116 (1990).

[31] Melas, V.B.: *Analytical Theory of E-Optimal Designs for Polynomial Regression on a Segment.* mODa5-Advance in Model-Oriented Data Analysis and Experimental Design, eds. Atkinson, A.C., Pronzato, L., and Wynn, H.P. Physica-Verlag, 51-58 (1998).

[32] Mungo, M., Zhu, W., Rosenberger, W.F.: *Adaptive Urn Designs for Estimating Several Percentiles of a Dose-Response Curve.* Statistics in Medicine, 23(13), 2137-2150 (2004).

[33] Pazman, Andrej: *Foundation of Optimum Experimental Designs.* Kluwer Academic Publishers (1986).

[34] Rosenberger, W.F., Flournoy, N., Durham, S.D.: *Asymptotic Normality of Maximum Likelihood Estimators from Multiparameter Response–Driven Designs.* Journal of Statistical Planning and Inference, 60, 69-76 (1997).

[35] Silvapulle, M.J.: *On the Existence of Maximum Likelihood Estimation for the Binomial Response Models.* Journal of the Royal Statistical Society, Series B, 43, 310-313 (1981).

[36] Silvey, S.D.: *Optimal Design Measures with Singular Information Matrices.* Biometrika, 56, 553–559 (1978).

[37] Silvey, S.D.: *Optimal Design.* Chapman and Hall, New York (1980).

[38] Sitter, R.R., Wu, C.F.J.: *Optimal Designs for Binary Response Experiments: Filler, D, and A Criteria.* The Scandinavian Journal of Statistics, 20, 329-341 (1993).

[39] White, L.: Ph.D Thesis. Imperial College, London (1975).

[40] Wynn, H.P.: *The sequential Generation of D-optimal Experimental Design.* Annals of Mathematical Statistics, 41,1655-1664 (1970).

[41] Wynn, H.P.: *Results in the Theory and Construction of D-optimum Designs.* Journal of the Royal Statistical Society, B, 34, 133-147 (1970).

[42] He, Z., Studden, W.J., Dongchu, S.: *Optimal Designs for Rational Models.* Annals of Statistics, 24, 2128-2147 (1996).

[43] Zocchi, Silvo S., Atkinson, A.C., Doney, A.N.: *Optimum Experimental Designs for Multinomial Logistic Models.* Biometrics, 55, 437–444 (1999).

VITA

Huwaida Rabie was born in 1958. She graduated from Teacher's Training College in 1978 and got a Diploma in Mathematics. She got her Bachelor Degree in Mathematics in 1984 from Bethlehem University in Palestine. She worked as a high school teacher for several years before coming to U.S.A. in 1998. Her dream in pursuing her studies became true when she got her Master Degree in Statistic in 2000 at American University in Washington D.C. She got a scholarship to pursue her Ph.D in statistics. She completed her course work at American University before she moved to the University of Missouri-Columbia with her advisor Dr. Nancy Flournoy.

Huwaida's dissertation is in Design of Experiments related to phase II clinical trials. She will join Georgetown University as a Postdoctoral fellow in the Cancer Research Center.