Optimal Designs for Contingent Response Models with Application to Toxicity-Efficacy Studies– A Technical Report

Huwaida Rabie
Phenomics and Bioinformatics Research Centre,
School of Mathematics and Statistics, University of South Australia,
Huwaida.Rabie@unisa.edu.au
and
Nancy Flournoy
Department of Statistics
University of Missouri
flournoyn@missouri.edu

December 15, 2011

Abstract

We generalize results in the literature to obtain a general family contingent response models. These models have ternary outcomes constructed from two Bernouli outcomes, where one outcome is only observed if the other outcome is positive. This family is represented in a canonical form which yields general results for its Fisher information. D and c optimal designs found numerically for a contingent response model with expected response probabilities taken from a bivariate extreme value distribution illustrate the model and motivate limiting results. Optimal designs for even modestly complex nonlinear response models cannot be expressed in closed form; and this includes the contingent response model. Limiting D optimal designs obtained in closed form can be used to approximate exact D and c optimal designs, as they are shown to be efficient over a wide range of parameter values, or they can be used to provide starting values in numerical searches for exact optimal designs. To provide a motivating context, we describe the two binary outcomes that compose the contingent responses as toxicity and no efficacy. Efficacy or lack thereof is assumed only to be observable in the absence of toxicity, resulting in the ternary response {toxicity, efficacy without toxicity, neither efficacy nor toxicity}. The rate of toxicity, and the rate of efficacy conditional on no toxicity, are assumed to increase with dose. The results provided in this paper are useful for the construction of efficient designs under a broad class of such models.

1 INTRODUCTION

Dose response models with single binary outcomes are used extensively in science and industry. Bivariate response models are of importance in numerous clinical trials, engineering strength tests and other applications. In clinical trials, a bivariate response may be efficacy (yes/ no) and toxicity (yes/ no). A subject's response to a drug then falls in one of four categories: 1) toxicity with no efficacy; 2) toxicity with efficacy; 3) no toxicity with efficacy (success/cure); and 4) no toxicity and no efficacy. Draglin & Fedorov (2006) numerically found optimal designs for several four outcome models; Draglin et al. (2008) introduced methods for dose finding based on (numerically obtained) penalized optimal designs for the bivariate probit model.

Sometimes efficacy can only be observed contingent on no toxicity. This happens when toxicity is lethal, or so serious that the subject is then treated "off protocol". Durham et al. (1998) describe such an application in some detail. In such applications, response categories, 1 and 2 are not distinguishable, and combining them yields a ternary response. In this context, Fan & Chaloner (2001, 2004) studied Bayesian designs for the continuation ratio model [which was introduced by Cox (1972) and is described in Agresti (2002)] applied to two binary responses. Rabie & Flournoy (2004) studied the contingent response model of Li et al. (1995) (see also Durham et al. (1998)); they later recognized that the continuation ratio model applied to two binary responses is a special case of the general contingent response model we call the logistic–logistic model. The contingent response model is described in Section 2.

Denote a K point design by $\xi = \{x_i, \xi_i\}_{i=1}^K$, where $\{x_i\}$ is the set of design points and $\{\xi_i\}$ are their corresponding frequencies of use. Rabie & Flournoy (2004) obtained the following invariance property for D optimal designs, generalizing a proof for the logistic–logistic model (Fan & Chaloner, 2001, 2004): if the design $\xi_0^* = \{x_i^*, \xi_i\}_{i=1}^K$ is locally D optimal, then the design with location-scale parameters (α, β), $\xi^* = \{(x_i^* - \alpha)/\beta, \xi_i\}_{i=1}^K$, is locally D optimal. This invariance property has also been shown to hold for the examples of c optimality in this paper, but the invariance for c optimality must be proven on a model by model basis.

In this paper, we find D and c optimal designs for the contingent response model. Motivated by the limiting optimal designs of Fan & Chaloner (2001, 2004), we show, for contingent response models that are constructed from continuous distribution functions, limiting D optimal designs can be constructed from the D optimal designs derived by considering the toxicity and conditional efficacy response models separately. We illustrate this using extreme value functions.

We also conjecture that the limiting c optimal designs for the dose maximizing the probability of efficacy without toxicity consist of at least two points. We prove that the limiting c optimal design is two points for a contingent response model constructed from extreme value functions. It is important to note that the c optimal design does not coincide with the well intended practice of seeking to place subjects at the dose that maximizes the chance of efficacy without toxicity.
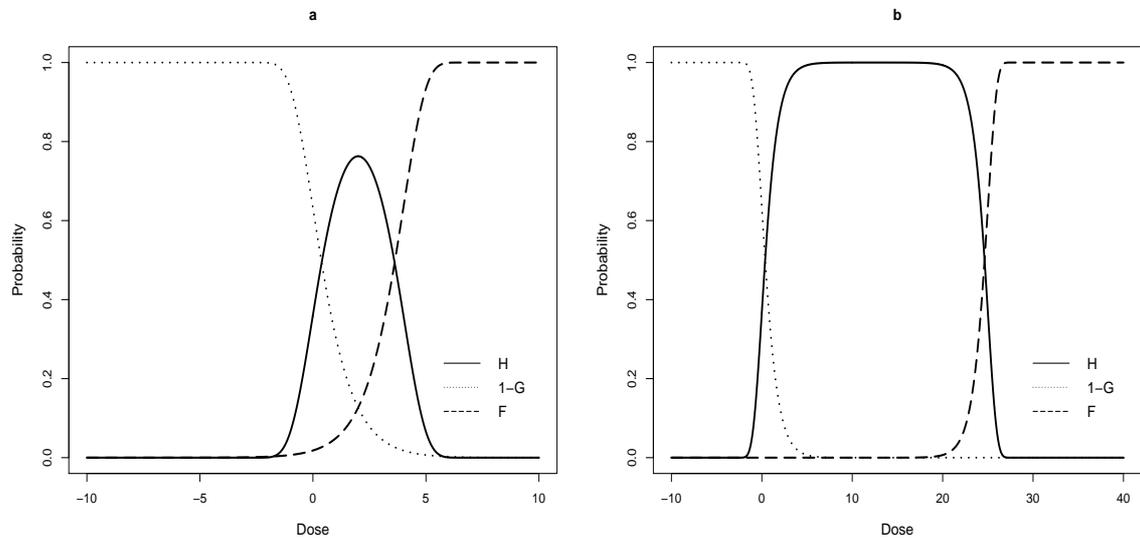
Figure 1: Examples of positive–negative extreme value models. a : $\bar{F}(x) = \exp(-\exp(x - 4)), G(x) = (\exp(-\exp(-x)))$; b : $\bar{F}(x) = \exp(-\exp(x - 25)), G(x) = (\exp(-\exp(-x)))$; $H(x) = \bar{F}(x)G(x)$.

Fisher information for the canonical contingent response model is written as a block diagonal matrix. We show that, as the two failure functions diverge, the blocks converge to the information matrices for the separate component models. We evaluate optimal designs and describe their behavior as the component models diverge. The limiting D and c optimal designs can be obtained in closed form, and they are found to be efficient for a variety of parameter values.

2 The Contingent Response Model

In the contingent response model, there are two types of failure. The failure types will have different labels in different applications, but in order to provide context, we call one failure type *toxicity* and the other *no efficacy*. We assume efficacy is *contingent* on no toxicity in that efficacy is only observed in the absence of toxicity. The probability of toxicity and the conditional probability of efficacy given no toxicity are assumed to increase with the dose.

Figure 1 shows examples of contingent response models in which the probabilities of ' toxicity' and 'efficacy given no toxicity' are modeled by the negative extreme value function with $\alpha_1 = -6$ and $-20, \beta_1 = 1$ and the positive extreme value function $G(x) = (\exp(-\exp(-x)))$.

2

To be more specific, define

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

$$Y_{2j} = \begin{cases} 1 & \text{if the } j\text{th subject has no efficacy} \\ 0 & \text{else} \end{cases}$$

for $j = 1, \ldots, N$. In the contingent response model, only three outcomes are observable for each subject, namely, success (i.e., no toxicity and efficacy) when $\{Y_{1j} = 0, Y_{2j} = 0\}$, no toxicity and no efficacy when $\{Y_{1j} = 0, Y_{2j} = 1\}$ and toxicity when $\{Y_{1j} = 1\}$.

Here when we refer generally to the contingent response model, we mean one that is a concatenation of two continuous location–scale cumulative distribution functions (cdf): a toxicity function $P\{Y_{1j} = 1 \mid x\} = F_{\alpha_1, \beta_1}, \alpha_1 \in R^1, \beta_1 > 0$ and a conditional efficacy function $P\{Y_{2j} = 0 \mid Y_{1j} = 0, x\} = G_{\alpha_2, \beta_2}, \alpha_2 \in R^1, \beta_2 > 0$, with $\bar{F} = 1 - F$ and $\bar{G} := 1 - G$. We assume F and G have the same support space and that the stimulus x has been transformed onto the same space as is $\arg F(x)$ and $\arg G(x)$. In addition, we assume that there exist link functions η_F and η_G such that

$$\eta_F(F_{\alpha_1, \beta_1}) = \alpha_1 + \beta_1 x;$$
$$\eta_G(G_{\alpha_2, \beta_2}) = \alpha_2 + \beta_2 x.$$

Selected cumulative distribution functions that are commonly used to model monotone dose-response functions are given in Table 1. The same models, or others, can be used for G. When $1 - F$ is the negative extreme value model, it is convenient to let G be the positive extreme value model which is $G(x) = exp(-exp(-\eta_G)), \beta_2 \geq 0$. Rabie & Flournoy (2004) showed that the continuation–ratio model studied by Fan & Chaloner (2001, 2004) is a contingent response function with logistic link functions for both F and G. The logistic link is popular in clinical studies; the probit is popular in acute toxicity studies; and the positive–negative extreme value link is popular in engineering. The extreme value model is attractive in that it is not symmetric and we use it for illustration.

The probability of success (that is, efficacy without toxicity) is

$$H(x) = P\{Y_{2j} = 0 \mid Y_{1j} = 0, x\} P\{Y_{1j} = 0 \mid x\} = \bar{F}(x)G(x), \tag{1}$$

and the probability of no efficacy and no toxicity is $P(Y_{1j} = 0, Y_{2j} = 1) = \bar{F}\bar{G}$. We define the optimal dose, if it exists, to be the dose that maximizes $H(x)$, and denote it by $\nu = argmax_x H(x)$. Li et al. (1995) give conditions for the existence and uniqueness of ν. For the positive–negative contingent response model, there is an explicit expression for the optimal dose which is $\nu = log(\beta_2/\beta_1) - (\alpha_1 + \alpha_2)/(\beta_1 + \beta_2)$.

We consider two sets of parameters: $\Theta = (\alpha_1, \beta_1, \alpha_2, \beta_2)$ and $\Theta = (\alpha_1, \beta, \alpha_2)$, where $\beta_1 = \beta_2 = \beta$. When $\alpha_2 = 0$ and $\beta_2 = 1$, we say that the model is in canonical form. Because of the invariance property mentioned in the introduction we can, without loss of generality, work with models in their canonical form.

3

Table 1: Selected toxicity function information.

Toxicity function	$\bar{F}(x)$	$v_F(x) = \frac{F'^2}{(1-F)F}$
Negative extreme value	$\exp(-\exp(\eta_F))$	$\dfrac{\exp(2\eta_F)\exp(-\exp(\eta_F))}{1-\exp(-\exp(\eta_F))}$
Logistic	$\dfrac{\exp(\eta_F)}{1+\exp(\eta_F)}$	$\dfrac{\exp(\eta_F)}{(1+\exp(\eta_F))^2}$
Probit	$\displaystyle\int_{\eta_F}^{\infty}\dfrac{\exp(-(\eta_F)^2/2)}{2\pi}\,dx$	$\dfrac{\exp(-(\eta_F)^2)}{2\pi\Phi(\eta_F)(1-\Phi(\eta_F))}$
Exponential	$\exp(\eta_F); x \le -\alpha_1/\beta_1$	$\dfrac{\exp(\eta_F)}{1-\exp(\eta_F)}$

3   FISHER'S INFORMATION

3.1   Preliminary Results

Let $I_F(x)$ and $I_G(x)$ be the Fisher information matrices for a single subject treated under the component response functions F and G, respectively:

$$I_F(x) = v_F(x)\begin{pmatrix} 1 & x \\ x & x^2 \end{pmatrix} \text{ and } I_G(x) = v_G(x)\begin{pmatrix} 1 & x \\ x & x^2 \end{pmatrix},$$

where $v_W(x) = W'^2(x)/\left[W(x)\bar{W}(x)\right]$, W is a continuous cumulative distribution function (e.g., F, G), $\bar{W} = 1 - W$, and W' is the derivative of W with respect to η_W. Ford et al. (1992) introduce $v_W(x)$ as weights in expressions of information for analogous models with only one outcome. Weights v_F are given in Table 1 for selected cumulative distribution functions. To simplify notation, define $v_{Wj} = v_W(z_j)$ and $v_{Wji} = v_W(z_{ji})$. We have the following proposition for these weight functions.

Proposition 3.1 *Let W be a parameter free cumulative distribution function and let $\eta = \alpha + \beta x$. Suppose $\mid \eta \mid^a v_W(\eta)$ has a limit for $a \in \{0, 1, 2\}$ that is either finite or infinite as $\eta \to -\infty$; then this limit must be zero.*

The proof is in Appendix 1.

The following lemma gives the information matrix, $I(x)$, for a single subject treated at dose x under the general contingent response model. The proof can be found in Rabie (2004)

Lemma 3.1   *(i) If $\Theta = \{\alpha_1, \beta_1, \alpha_2, \beta_2\}$, then $I(x) = diag\,(A, B)$, where $A = I_F(x)$ and $B = \bar{F}(x)I_G(x)$;*

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

$$I(x) = v_F(x)\begin{pmatrix} 1 & x & 0 \\ x & x^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \bar{F}(x)v_G(x)\begin{pmatrix} 0 & 0 & 0 \\ 0 & x^2 & x \\ 0 & x & 1 \end{pmatrix}. \qquad (2)$$

4

3.2 Parameter Set I: $\Theta = \{\alpha_1, \beta_1, \alpha_2, \beta_2\}$

It is well known that D optimal designs for commonly used sigmoidal response functions are two point designs [cf. Atkinson & Donev (1992)]. We now show that, as $\alpha_1 \to -\infty$, the information in a four point design under the four parameter contingent response model is a concatenation of information in two independent two point designs. In Section 4, we show that, as $\alpha_1 \to -\infty$, the limiting D optimal design under the canonical contingent response model is composed of the elements from the two D optimal designs under F and G separately.

Many optimal designs obtained in the paper were found to be symmetric, containing both a dose x and $-x$. Thus it is convenient to define standardized transformations:

$$z_1 = (-x - \alpha_1)/\beta_1; \quad z_{1i} = (-x_i - \alpha_1)/\beta_1;$$
$$z_2 = (x - \alpha_2)/\beta_2; \quad z_{2i} = (x_i - \alpha_2)/\beta_2.$$

Define equally weighted two-point experiments under F and G, respectively:

$$\xi_F = \begin{pmatrix} z_{11} & z_{12} \\ 0.50 & 0.50 \end{pmatrix}; \quad \xi_G = \begin{pmatrix} z_{21} & z_{22} \\ 0.50 & 0.50 \end{pmatrix}. \tag{3}$$

The information matrices for these two experiments, respectively, are

$$M_F = 0.5 \sum_{i=1}^{2} I_F(z_{1i}) = 0.5 \sum_{i=1}^{2} v_F(z_{1i}) \begin{pmatrix} 1 & z_{1i} \\ z_{1i} & z_{1i}^2 \end{pmatrix};$$

$$M_G = 0.5 \sum_{i=1}^{2} I_G(z_{2i}) = 0.5 \sum_{i=1}^{2} v_G(z_{2i}) \begin{pmatrix} 1 & z_{2i} \\ z_{2i} & z_{2i}^2 \end{pmatrix}.$$

A design that contains all the points from two separate designs, each with half of its original weight, we call a concatenation of the two separate designs. So the concatenation of ξ_F and ξ_G is

$$\xi_C = \left\{ \begin{array}{cccc} z_{11} & z_{12} & z_{21} & z_{22} \\ 0.25 & 0.25 & 0.25 & 0.25 \end{array} \right\}.$$

We now show how the limiting information matrix at ξ_C decomposes into the component information matrices at ξ_F and ξ_G. By Lemma 3.1, the information matrix for the contingent response model at ξ_C is

$$M_C = \begin{pmatrix} M_{C_A} & 0 \\ 0 & M_{C_B} \end{pmatrix}, \tag{4}$$

where

$$M_{C_A} = 0.25 \sum_{i=1}^{2} \left[I_F(z_{1i}) + I_F(z_{2i}) \right]; \tag{5}$$

$$M_{C_B} = 0.25 \sum_{i=1}^{2} \left[\bar{F}(z_{1i}) I_G(z_{1i}) + \bar{F}(z_{2i}) I_G(z_{2i}) \right]. \tag{6}$$

Lemma 3.2 $M_C \to 0.5 \begin{pmatrix} M_F & 0 \\ 0 & M_G \end{pmatrix}$ as $\alpha_1 \to -\infty$.

Proof:

For the canonical model, $z_{2i} = x_i$ does not depend on parameters and neither does $\bar{F}_{z_{1i}}$ and v_{G2i}. For $a \leq 2$ and $i = 1, 2$, it follows from Proposition 3.1 that $z_{2i}^a v_{F2i}$ and $z_{1i}^a v_{G1i}$ as functions of α_1 both go to zero as $\alpha_1 \to -\infty$. This implies a. the second term in (5) goes to zero as $\alpha_1 \to -\infty$. and hence M_{C_A} converges to $0.25 \sum_{i=1}^2 I_F(z_{1i}) = 0.5 M_F$. b. the first term in (6) together with the fact that $\bar{F}_{z_{1i}}$ does not depend on parameters goes to zero. Now $\bar{F}_{z_{2i}} \to 1$, $i = 1, 2$ when $\alpha_1 \to -\infty$, which implies that the second term in (6) converges to $0.25 \sum_{i=1}^2 I_G(z_{2i}) = 0.5 M_G$. This concludes the proof. □

It is interesting to note that as $\alpha_1 \to \infty, M_{C_A} \to 0.5 M_F$, but $M_{C_A} \to 0$.

3.3   Parameter Set II: $\Theta = \{\alpha_1, \beta, \alpha_2\}$, where $\beta_1 = \beta_2 = \beta$

Consider the special case in which $\beta_1 = \beta_2$. For example, if the toxicity rate equals the rate of failure to benefit the subject, one would have $\beta_1 = \beta_2 = 1$. The following information for the four point design $\xi_D = \left\{ \begin{array}{cccc} z_{11} & z_{12} & z_{21} & z_{22} \\ \xi_{11} & \xi_{12} & \xi_{21} & \xi_{22} \end{array} \right\}$ is used in Section 4 for deriving D optimality in this special case:

$$
\begin{aligned}
M_D &= \sum_{k=1}^2 \sum_{i=1}^2 \xi_{ki} I_{3x3}(z_{ki}) \\
&= \sum_{k=1}^2 \sum_{i=1}^2 \xi_{ki} \left[v_{Fki} \left(\begin{array}{ccc} 1 & z_{ki} & 0 \\ z_{ki} & z_{ki}^2 & 0 \\ 0 & 0 & 0 \end{array} \right) + \bar{F}(z_{ki}) v_{Gki} \left(\begin{array}{ccc} 0 & 0 & 0 \\ 0 & z_{ki}^2 & z_{ki} \\ 0 & z_{ki} & 1 \end{array} \right) \right].
\end{aligned}
\tag{7}
$$

Lemma 3.3 *Let $\tilde{M}_D = \lim_{\alpha_1 \to -\infty} M_D$ then*

$$
\tilde{M}_D = \sum_{i=1}^2 \left[\xi_{1i} v_{F1i} \left(\begin{array}{ccc} 1 & z_{1i} & 0 \\ z_{1i} & z_{1i}^2 & 0 \\ 0 & 0 & 0 \end{array} \right) + \xi_{2i} v_{G2i} \left(\begin{array}{ccc} 0 & 0 & 0 \\ 0 & z_{2i}^2 & z_{2i} \\ 0 & z_{2i} & 1 \end{array} \right) \right].
\tag{8}
$$

Proof: By Proposition 3.1, for $a \leq 2$, $z_{1i}^a v_{F2i}$ and $z_{2i}^a v_{G1i}$ both go to zero as $\alpha_1 \to -\infty$ for the canonical model. Note $\bar{F}(z_{1i})$, v_{F1i}, and v_{G2i} do not depend on parameters. This implies that the second and the third elements in (7) go to zero. Also $\bar{F}(z_{2i}) \to 1$ as $\alpha_1 \to -\infty$. This implies that M_D can be approximated by (8) when the individual response functions sufficiently diverge. □

4   LOCALLY OPTIMAL DESIGNS

4.1   D optimality

Locally D optimal designs maximize the log determinant of Fisher's information with respect to the parameters. Optimization procedures produce continuous designs where

the weight ξ_i for each point x_i is a positive number and $\sum_{i=1}^{K} \xi_i = 1$ [cf. Atkinson & Donev (1992), pp. 93]. Let n_i denote the number of subjects treated at x_i with $N = \sum_i n_i$. To implement ξ, set n_i as close as possible to $N\xi_i$.

For nonlinear response functions, optimal design points are functions of the unknown parameters rather than real numbers, but implementation can be accomplished with a sequential treatment allocation procedure that mimics numerical iterative procedures for finding optimal design points. See Wynn (1970), Fedorov (1974), Silvey (1980); Cook & Nachtsheim (1982), Atkinson & Donev (1992), and Fedorov & Hackl (1997).

The directional derivative for the D optimality criterion under a model with information matrix M is $\mathcal{D}_M = Tr[I(x)M^{*-1}]$, where $I(x)$ varies with x and M^* is the information evaluated at the optimal design, ξ^*. By the General Equivalence Theorem (Kiefer, 1974), a design is D optimal if and only if $\mathcal{D}_M \leq p$ for all x with equality holding for $x \in$ support(ξ^*), where p is the number of parameters. Optimal designs described in the example below were found numerically as described in Rabie (2004)and verified using the General Equivalence Theorem.

Example

Table 2 provides locally D optimal designs for selected canonical positive-negative extreme value model with $\Theta = \{\alpha_1, \beta_1, 0, 1\}$ (also see Table 3 in Appendix 5). We found locally D optimal designs consisting of two, three, and four points for small, moderate and large negative values of α_1 and different β_1 values, respectively. For positive values of α_1, optimal designs consist of two and three design points depending on the value of β_1.

Note as $-\alpha_1$ gets larger, optimal designs come to consist of four approximately equally weighted design points. These design points are roughly the same as the concatenated D optimal design points found for two separate experiments under the component models $F(z_1)$ and $G(z_2)$, respectively. Whereas optimal design points under the separate component models have closed form expressions, the design points for the contingent response model must be found numerically. In Section 5, we illustrate how the designs derived from the component models $F(z_1)$ and $G(z_2)$ can be used to approximate optimal designs. Table 4 (Appendix 5) shows the locally D optimal designs assuming $\Theta = \{\alpha_1, 1, 0\}$. Fan & Chaloner (2001, 2004) found similar results for the logistic–logistic model. That is, D optimal designs were found to consist of two, three, and four points depending on values of α_1.

4.2 c optimality

In the contingent response model, the dose $\nu = \arg \max H(x)$ is of great interest. The corresponding c optimality criterion minimizes the asymptotic variance of the maximum likelihood estimate of ν.

For positive–negative extreme value models and logistic–logistic models, it has been shown (Rabie & Flournoy, 2004; Fan & Chaloner, 2004) that if ξ_0^* is the locally c optimal design for the canonical $\{\alpha_1, \beta_1, 0, 1\}$ or $\{\alpha_1, 0, 1\}$ models, then designs ξ^* are locally c optimal for the general $\{\alpha_1, \beta_1, \alpha_2, \beta_2\}$ and $\{\alpha_1, \beta, \alpha_2\}$ models, respectively. We conjecture that this location-scale invariance holds for the class of contingent response models considered herein, but proofs must be done model by model.

Table 2: D-optimal designs for selected canonical positive–negative extreme value models

α_1	$\beta_1 = 0.5$		$\beta_1 = 1$		$\beta_1 = 2$	
	Design Points	Weight	Design Points	Weight	Design Points	Weight
0	-1.2752	0.4720	0.4755	0.5000	-1.2158	0.4391
	0.5985	0.3382	-1.2808	0.5000	-0.17186	0.2660
	1.948	0.1898			0.34168	0.2945
-1	-1.0982	0.4005	-1.1222	0.4087	-1.1287	0.3566
	0.8243	0.3635	0.4647	0.3447	0.08121	0.3725
	3.9569	0.2360	1.8530	0.2466	0.9055	0.2708
-3	-0.9329	0.3312	-0.9414	0.3092	-1.0136	0.2810
	1.4913	0.4200	1.2863	0.4393	0.7675	0.4573
	7.6891	0.2488	3.8609	0.2515	1.9332	0.2618
-5	-0.9101	0.2931	-0.8454	0.2717	-0.9256	0.2584
	1.6895	0.3729	2.2797	0.477	1.5614	0.4831
	7.6511	0.0948	5.8125	0.2513	2.9311	0.2585
	11.6989	0.2392				
-10	-0.9778	0.2509	-0.973	0.25	-0.8987	0.2418
	1.3465	0.2524	1.362	0.2515	1.3106	0.1511
	17.3336	0.2473	8.6396	0.2793	4.0744	0.3544
	21.955	0.2493	10.9725	0.2491	5.4483	0.2526
-15	-0.9796	0.2500	-.9795	0.2500	-0.9709	0.2476
	1.3379	0.2500	1.3381	0.2500	1.3222	0.2385
	27.3247	0.2500	13.6618	0.2500	6.7937	0.2635
	31.9592	0.2500	15.9795	0.2500	7.9835	0.2505
-20			-0.9796	0.2500	-0.9784	0.2496
			1.3378	0.2500	1.3352	0.2485
			18.6623	0.2500	9.3257	0.2519
			20.9796	0.2500	10.4889	0.2501

Tables 5- 7 (Appendix 5) show c optimal designs for selected canonical $\{\alpha_1, \beta_1, 0, 1\}$ and $\{\alpha_1, 1, 0\}$ positive-negative extreme-value models, respectively. With and without assuming equal slopes, the c optimal designs are found to consist of two design points over a wide range of parameters. These designs were obtained numerically and verified using the General Equivalence Theorem (Kiefer, 1974). Singular optimal designs occur. In such cases, Silvey's Theorem (Silvey, 1980) was used to verify optimality. For large positive values of α_1 the maximum probability of success is negligible for all doses which leads to numerical instabilities when trying to find the optimal design.

5 LIMITING LOCALLY OPTIMAL DESIGNS

5.1 Limiting D optimality when $\Theta = \{\alpha_1, \beta_1, \alpha_2, \beta_2\}$

Let M_F^* and M_G^* denote the information matrices evaluated at the optimal designs ξ_F^* and ξ_G^* for the separate response functions F and G, respectively. Let $\xi_C^\diamond$ denote the concatenation of ξ_F^* and ξ_G^*, whereas ξ_C^* denotes the D optimal design for contingent response model. Lemma 3.2 and Proposition 3.1 lead to

Theorem 5.1 $\xi_C^\diamond \longrightarrow \xi_C^*$ *as* $\alpha_1 \to -\infty$.

By the General Equivalence Theorem, Theorem 5.1 is true iff $\lim_{\alpha_1 \to -\infty} \mathcal{D}_{M_C} \leq 4$ for all $x \in \text{support}(\xi_C)$ with equality holding for $x \in \xi_C^\diamond$. To provide insight into the proof of Theorem 5.1, Figure 2 shows directional derivatives for some positive-negative extreme value models. Figure 2(a) displays directional derivatives for three component models: $\mathcal{D}_{M_G} = Tr\left[I_G^{-1} M_G^{*-1}\right]$, where G is the negative canonical extreme value model and $\mathcal{D}_{M_F} = Tr\left[I_F^{-1} M_F^{*-1}\right]$, where F with $\beta_1 = 2$ and $\alpha_1 = -10, -15, -25$. The information matrices, M_F^* and M_G^*, are evaluated at the optimal designs, ξ_F^* and ξ_G^*, respectively, so each curve reaches 2, the number of model parameters, at the optimal design points for the respective model. Compare these with Figure 2(b), which shows $2\left(\mathcal{D}_{M_G} + \mathcal{D}_{M_F}\right)$ for $\beta_1 = 2$ and $\alpha_1 = -10, -15, -25$. The height of $2\left(\mathcal{D}_{M_G} + \mathcal{D}_{M_F}\right)$ at its modes is not equal to 4 when $\alpha_1 = -10$; when $\alpha_1 = -15$, the height at the modes is closer to 4; but when $\alpha_1 = -25$, they equal 4 indicating that the optimal design ξ_C^* has been found. Note that the modes shift with α_1 until they are located at the same points as are the modes shown in Figure 2(a) of $\mathcal{D}_{M_F}$ with $\alpha_1 = -25$ and canonical $\mathcal{D}_{M_G}$.

Proof

The proof amounts to showing that $\lim_{\alpha_1 \to -\infty} \mathcal{D}_{M_C} - 4$ is nonpositive and attains its supremum at all $x \in \text{support}\left(\xi_C^\diamond\right)$. By Lemma 3.2,

$$M_C^{-1} = \begin{pmatrix} M_{C_A}^{-1} & 0 \\ 0 & M_{C_B}^{-1} \end{pmatrix} \xrightarrow[\alpha_1 \to -\infty]{} 2 \begin{pmatrix} M_F^{-1} & 0 \\ 0 & M_G^{-1} \end{pmatrix}.$$

Hence

$$\mathcal{D}_{M_C}(x) - 4 \xrightarrow[\alpha_1 \to -\infty]{} 2\left(\mathcal{D}_{M_F}(x) + \mathcal{D}_{M_G}(x)\right) - 4.$$

Consider $\mathcal{D}_{M_F} = Tr\left[I_F(x) M_F^{*-1}\right]$. By Proposition 3.1, each element of $I_F(x)$ goes to zero as $\eta_F(x) \to -\infty$, while M_F^{*-1} is fixed. Therefore, $\mathcal{D}_{M_F} \to 0$ as $\eta_F(x) \to \pm\infty$, which for fixed x occurs if $\alpha_1 \to -\infty$, and vise versa. Thus, given $\epsilon_F > 0$, there exists points, say x_F^1 and x_F^2, such that $\mathcal{D}_{M_F}(x) < \epsilon_F$ for all $x \leq x_F^1$ and for all $x \geq x_F^2$. Analogously, $\mathcal{D}_{M_G} \to 0$ as $\eta_G(x) \to \pm\infty$, and given $\epsilon_G > 0$, there exists points, say x_G^1 and x_G^2, such that $\mathcal{D}_{M_G}(x) < \epsilon_G$ for all $x \leq x_G^1$ and for all $x \geq x_G^2$. With the model in canonical form, x_G^1 and x_G^2 are fixed, while x_F^1 and x_F^2 shift with the location parameter α_1. Thus, letting ϵ_F and ϵ_G go to zero, the following scenarios develop as α_1 becomes increasingly negative and F shifts toward infinity:

9

1. Case 1a. $\alpha_1 << 0$ and $x_G^1 < x_F^1 < x_G^2 < x_F^2$:

$$\mathcal{D}_{M_C} = \begin{cases} 0 & \text{if } x < x_G^1; \\ \mathcal{D}_{M_G} & \text{if } x_G^1 \leq x < x_F^1 \ ; \\ \mathcal{D}_{M_C} & \text{if } x_F^1 \leq x < x_G^2 \ ; \\ \mathcal{D}_{M_F} & \text{if } x_G^2 \leq x < x_F^2; \\ 0 & \text{if } x \geq x_F^2. \end{cases} \tag{9}$$

2. Case 1b. $\alpha_1 \to -\infty$, so $x_G^1 < x_G^2 < x_F^1 < x_F^2$: (9) goes to

$$\mathcal{D}_{M_C} = \begin{cases} 0 & \text{if } x < x_G^1; \\ \mathcal{D}_{M_G} & \text{if } x_G^1 \leq x < x_G^2 \ ; \\ 0 & \text{if } x_G^2 \leq x < x_F^1 \ ; \\ \mathcal{D}_{M_F} & \text{if } x_F^1 \leq x < x_F^2; \\ 0 & \text{if } x \geq x_F^1. \end{cases} \tag{10}$$

Now for any point $x^* \in \left(x_G^2, x_F^1 \right)$, the directional derivative for design $\xi_C^\diamond$ is

$$\mathcal{D}_{M_C} - 4 \xrightarrow[\alpha_1 \to -\infty]{} 2\left(\mathcal{D}_{M_F} + \mathcal{D}_{M_G} - 2 \right) = \begin{cases} 2(\mathcal{D}_{M_G} - 2) & x \leq x^* \\ 2(\mathcal{D}_{M_F} - 2) & x > x^*. \end{cases}$$

Now $(\mathcal{D}_{M_G} - 2)$ and $(\mathcal{D}_{M_F} - 2)$ are the directional derivatives for G and F, respectively. By the General equivalence theorem each is non-positive and attains its maximum at ξ_F^* and ξ_G^*. Thus $\mathcal{D}_{M_C}(x) - 4$ is non-positive and attains its maximum at $\xi_C^\diamond$, therefore ξ_C^* is the D-optimal design for the contingent response model. □

Corollary 5.1 *(i) For the canonical positive–negative extreme value model with $\Theta = \{\alpha_1, \beta_1, \alpha_2, \beta_2\}$, the locally limiting D optimal design is given by*

$$\xi_C^* = \left\{ \begin{array}{cccc} -0.9796 & 1.3377 & \frac{0.9796 - \alpha_1}{\beta_1} & \frac{-1.3377 - \alpha_1}{\beta_1} \\ 0.25 & 0.25 & 0.25 & 0.25 \end{array} \right\}. \tag{11}$$

(ii) For the canonical logistic–logistic model with $\Theta = \{\alpha_1, \beta_1, \alpha_2, \beta_2\}$, the locally limiting D optimal design is given by

$$\xi_C^* = \left\{ \begin{array}{cccc} -1.543 & 1.543 & \frac{1.543 - \alpha_1}{\beta_1} & \frac{-1.543 - \alpha_1}{\beta_1} \\ 0.25 & 0.25 & 0.25 & 0.25 \end{array} \right\}. \tag{12}$$
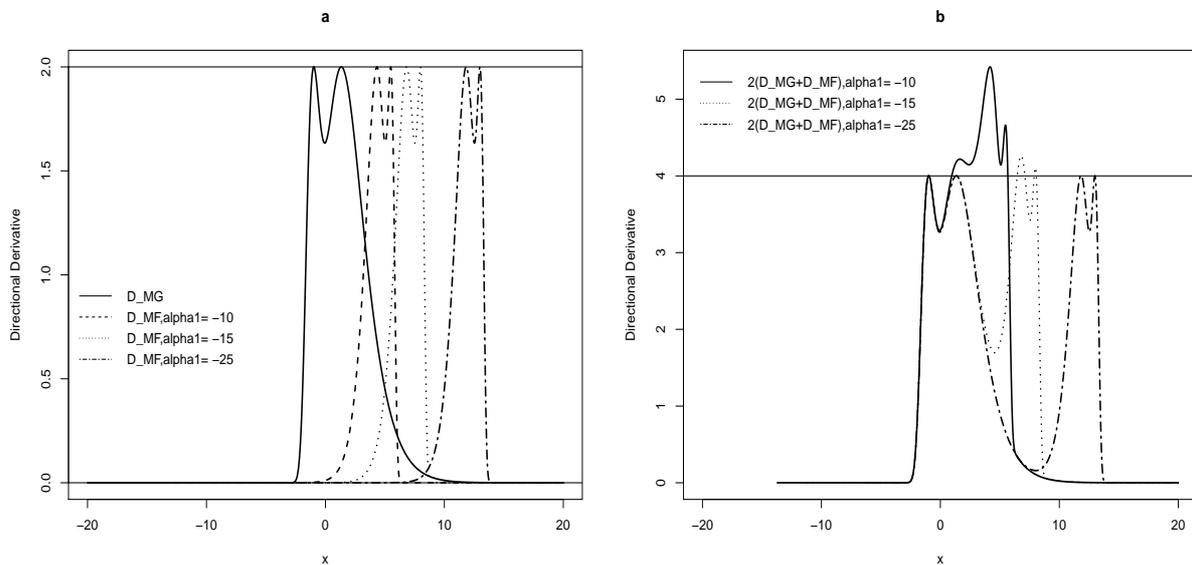
Figure 2: (a): The directional derivatives for the canonical extreme value for G, $\mathcal{D}_{M_G}$, and F, $\mathcal{D}_{M_F}$, for several values of α_1. b: $2(\mathcal{D}_{M_F} + \mathcal{D}_{M_G})$.

Proof

(i) The optimal design points for a single positive extreme value function with scale parameter $\beta_2 = 1$ and location parameter $\alpha_2 = 0$ are $(-0.9796, 1.3377)$ with equal weights. The optimal design points for the negative extreme value function, when $\beta_1 = 1, \alpha_1 = 0$, are $(0.9796, -1.3377)$. For more detail see Ford et al. (1992). By the invariance property of the D optimal designs for linearly transformed design spaces, the optimal design points for the negative extreme value function with scale parameter β_1 and location parameter α_1, are $(0.9796 - \alpha_1)/\beta_1$ and $(-1.3377 - \alpha_1)/\beta_1$ with equal points. Hence by Lemma 3.2, ξ_C^* is the limiting optimal design for the canonical model.

(ii) The optimal design points for a single logistic response with equal weights are $(-1.543, 1.543)$ when $\alpha_1 = 0$ and $\beta_1 = 1$ (White , 1975). The result follows from the same argument as in (i).

□

Fan & Chaloner (2001) obtained the result for logistic–logistic model using a different method of proof.

5.2 Limiting D optimality when $\Theta = \{\alpha_1, \beta, \alpha_2\}$, $\beta_1 = \beta_2$.

In this section, we derive closed form for the limiting D optimal designs based on the results given in Table 4 (Appendix 5) for the canonical positive–negative extreme value

model.

Theorem 5.2 *For the canonical positive–negative extreme value model with* $\Theta = \{\alpha_1, 1, 0\}$, *the limiting locally D optimal design is given by*

$$\xi_D^* = \left\{ \begin{array}{cccc} -0.8537 & 1.0773 & (0.8537 - \alpha_1) & (-1.0773 - \alpha_1) \\ 0.2900 & 0.2100 & 0.2900 & 0.2100 \end{array} \right\}.$$

The detailed proof in Appendix 2 amounts to maximizing the determinant of $\tilde{M}_D$ in Lemma 3.3 and then showing that the directional derivative $\mathcal{D}_{\tilde{M}_D} - 3$ goes to zero as $\alpha_1 \to -\infty$. Although ξ_D^* consists of four design points, these points are not the points obtained from concatenating the optimal designs points of the separate positive and negative extreme value models. An analogous result for the logistic–logistic model was found by Fan & Chaloner (2004). They found the limiting optimal design consists of four equally weighted points: $\{1.223 - \alpha_1, -1.223 - \alpha_1, -1.223, 1.223\}$.

5.3 Limiting c optimality when $\Theta = \{\alpha_1, \beta, \alpha_2\}$

We conjecture from Tables 5 – 7 in Appendix 5 that locally c optimal designs for the canonical model for estimating the dose with maximum success probability consist of two design points with weights depending on β and α_1. Thus we conjecture that the limiting c optimal design will also have two points. In particular, when $\beta_2 = \beta_1$ for the canonical model, we conjecture that these points are $z_1^* = \arg\max_x v_F$ and $z_2^* = \arg\max_x v_G$.

We have this result for the positive–negative extreme value model:

Theorem 5.3 *For the canonical positive–negative extreme value model with* $\{\Theta\} = \{\alpha_1, 1, 0\}$, *the locally limiting c optimal design is given by*

$$\xi_E^* = \left\{ \begin{array}{cc} \arg\max_x v_F & \arg\max_x v_G \\ 0.5000 & 0.5000 \end{array} \right\},$$

where $z_2^* = \arg\max_x v_G = -0.466$ *and* $z_1^* = \arg\max_x v_F = 0.466 - \alpha_1$.

The proof is in Appendix 3. The analogous locally c optimal design for the logistic–logistic model is two equally weighted points at $\arg\max_x v_F = -\alpha_1$ and $\arg\max_x v_G = 0$ [Fan & Chaloner (2004)].

6 EFFICIENCY OF THE LIMITING OPTIMAL DESIGNS

Now we illustrate the efficiency of limiting D and c optimal designs for the canonical positive-negative extreme value model. Efficiency is defined as the sample size n with the optimal design (ξ^*) that produces the same criterion value as does the sample size from the limiting design ξ_L^*. For D and c optimality, respectively, the efficiency is $n = [\det[M_L]/\det[M_L^*]]^{1/p}$ and $n = (\dot{\nu} \det[M_L^*]^{-1} \nu)/(\dot{\nu} M_L^{-1} \nu)$, where $\dot{\nu}$ is the
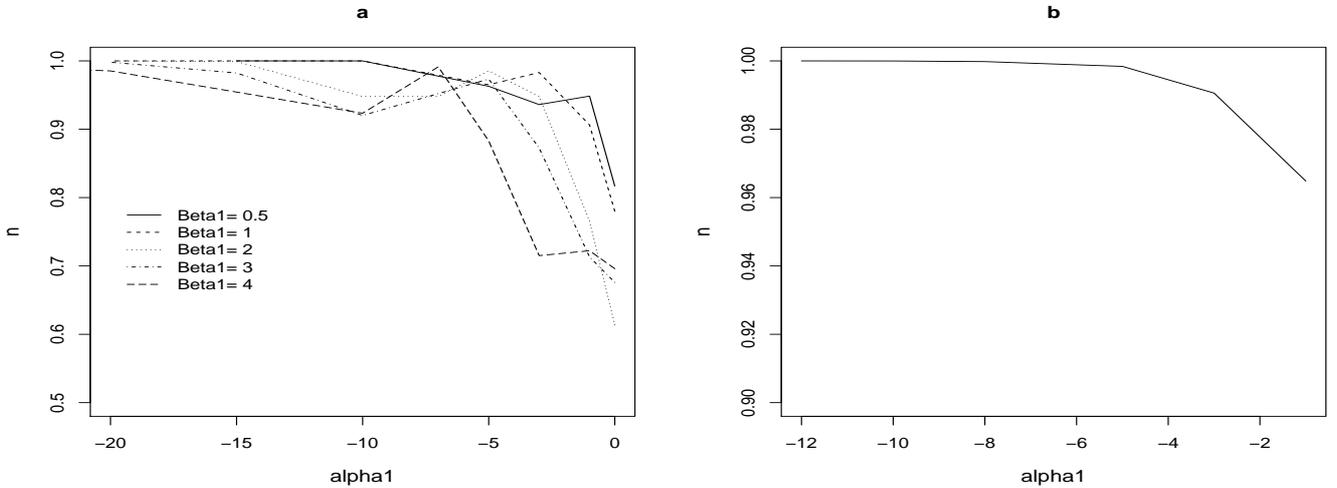
Figure 3: a: Efficiency of limiting D optimal for $\beta_1 = 0.5, 1, 2, 3, 4$ and b: c optimal designs when $\beta_1 = \beta_2 = 1$, for several values of α_1 for canonical positive-negative extreme value models.

derivative of ν and p is the number of parameters. Figure 3a displays the reduced sample size n required as function of α_1 when $\Theta = \{\alpha_1, \beta_1, \alpha_2, \beta_2\}$ for different values of β_1/β_2, which is equal to β_1 for the canonical extreme value model. The efficiencies for large negative values of α_1 approach 1.00 with both large and small values of β_1/β_2. For small to moderate values of $|\alpha_1|$, the efficiencies vary with the value of β_1/β_2, and they range from $.71 - .91$. It can be seen that the efficiencies are higher for $\beta_1/\beta_2 = 0.5$ and $\beta_1/\beta_2 = 1$ than larger values of β_1/β_2, but they are still reasonable. Figure 3b and Figure 1 in Appendix 6 show the efficiencies versus α_1 for limiting D and c optimal designs when $\Theta = \{\alpha_1, \beta, \alpha_2\}$, respectively. For D optimal designs, these efficiencies are all higher than .92 and they approach .96 for large values of $|\alpha_1|$ examined here. For the c optimal designs, the efficiencies of limiting optimal designs are higher than 0.965 over the range we examined.

7 CONCLUDING REMARKS

Generalizing models in the literature, we presented a family of contingent response models in which one binary response can only be observed contingent on another. Writing these models in a canonical form simplified characterization of their Fisher information. We have shown that when the two response functions diverge, the information matrix for contingent response model converges to a block diagonal matrix in which each block contains the information matrix of a single response treated separately. This result was used to show that limiting D optimal designs consist of the concatenation of the optimal designs of each single response. Thus limiting D

13

optimal designs can be surmised from the knowledge of individual optimal designs found separately under single response models, F and G. In spite of recent progress in finding optimal designs analytically for some classes of three and four parameter models (see Li & Majumdar (2008) and Yang (2010)), it is not clear these results extend to the contingent response family. However, optimal designs for many common two–parameter single response models are available in the literature, and we show how they can be used directly to construct efficient approximate optimal designs under three and four parameter contingent response models, avoiding the challenge of numerical searches. Alternatively, the limiting designs can serve as starting values in numerical searches for exact optimal designs or as initial designs in Bayesian or frequentist sequential procedures. The c optimal design for the optimal dose (i.e., the dose that maximizes the probability of efficacy without toxicity) was found to be a two point design for the positive–negative extreme value model, as is also the case for the logistic–logistic model (Fan & Chaloner (2001, 2004)). We conjecture the theorem holds generally for other models in this family, but the proof must be done case by case. This is an important result. Minimizing the variance of a percentile in toxicity studies often results in one point optimal designs at the target dose. In such cases, the optimal design is consistent with the popular use of procedures that seek to treat subjects at the target dose. This strategy has been transferred to the toxicity–efficacy setting, with procedures proposed that seek to place subjects at the optimal dose. However, in the toxicity–efficacy setting, the objective of treating subjects at the optimal dose and the objective of estimating that dose efficiently conflict. This conflict needs to be more widely recognized and trade–offs in the objective functions consciously considered in the design process. These findings extend the results of Fan & Chaloner (2001, 2004) and Rabie & Flournoy (2004) and support the use of limiting optimal designs with contingent response models other than the logistic–logistic and the positive-negative extreme value models, such as the probit–probit and the logistic–exponential models.

ACKNOWLEDGMENT

Bibliography

AGRESTI, A. (2002). Categorical Data Analysis. Wiley.

ATKINSON, A.C. & DONEV, A.N. (1992). Optimum Experimental Designs. Oxford University Press, New York.

COOK, R.D. & NACHTSHEIM, C.J. (1982). Model Robust, Linear-Optimal Designs. *Technometrics* **24**, 49-54.

COX, D.R. (1972). Regression Models & Life-Tables. *Journal of the Royal statistical Society: Series B* **34**, 187–220.

DRAGALIN, V. & FEDOROV, V. (2006). Adaptive Designs for Dose-finding based on Efficacy-Toxicity Response. *Journal of Statistical Planning* & Inference **136**, 1800–1823.

DRAGLIN, V., FEDOROV, V. & WU, Y. (2008). Adaptive Designs for Selecting Drug Combinations Based on Efficacy-Toxicity Response. *Journal of Statistical Planning* & Inference 1**38**, 352–373.

DURHAM, S.D, FLOURNOY, N. & LI, W. (1998). A Sequential Design for Maximizing the Probability of a Favourable Response. *Canandian Journal of Statistics* **26**, 479–495.

FAN, S.K. & CHALONER, K. (2001). Optimal Designs for a Continuation–Ratio Model. *In mODa6-Advances in Model-Oriented Design & Analysis*. eds. Atkinson, A.C., Hackl, P. & Muller. W.G. Physica–Verlag, 77–86.

FAN, S.K. & CHALONER, K. (2004). Optimal Designs & Limiting Optimal Designs for a Trinomial Response. *Journal of Statistical Planning* & Inference **126**, 347-60.

FEDOROV, V.V. (1974). Theory of Optimal Experiments, translated & edited by Studden, W.J. & Klimko. E.M. Academic Press New York & London.

FEDOROV, V.V. & HACKL, P. (1997). Model-Oriented Design of Experiments. *Lecture Notes in Statistics* 125, Springer–Verlag: New York.

FORD, I., TORSNEY, B. & WU, C.F.J. (1992). The Use of a Canonical Form in the Construction of Locally Optimal Designs for Non-linear Problems. *Journal of the Royal Statistical Society Series B* **54**, 569–583.

GILL, P.E., MURRAY, W., SAUNDERS, M.A. & WRIGHT,M.H. (1998). User's Guide For NPSOL 5:0 A Fortran Package For Nonlinear Programming. *Technical Report SOL* 86–1. Department of EESOR Stanford University.

Heise, M.A. & Myers, R.H. (1996). Optimal Designs for Bivariate Logistic Regression. *Biometrics* **52**, 613-624.

KIEFER, J.(1974). General Equivalence Theory for Optimum Designs (approximate theory). *Annals of Statistics* **2**, 849-879.

LI, Z., DURHAM, S.D. & FLOURNOY,N. (1995). An Adaptive Design for Maximization of a Contingent Binary Response. *IMS Lecture Notes-Monograph Series* **25**, 179–196. Hayward, CA: Institute of Mathematical Statistics.

LI, G. & MAJUMDAR, D. (2008). D-optimal designs for logistic models with three & four parameters. Journal of Statistical Planning & Inference 138, 1950-1959.

LISKI, E.P., MANDAL, N.K., SHAH, K.R. & SINIHA, B.K. (2002). Topics in Optimal Design. New York: Springer–Verlag.

PUKELSHEIM, F. (2008). Optimal Design of Experiments. Society for Industrial & Applied Mathematics (SIAM), Philadelphia, PA.

RABIE, H.S. (2004). Optimal Designs with Opposing Failure in the Contingent Response Model. Ph.D Thesis, Department of Statistics, University of Missouri-Columbia.

RABIE, H.S. & FLOURNOY, N. (2004). Optimal Designs for Contingent Response Models. In mODa7-Advances in Model-Oriented Design & Analysis, Di Bucchianico, A., Lauter, H. Wynn, H.P. (Eds.), Physica–Verlag, 133–142.

SILVEY, S.D. (1980). Optimal Design. New York: Chapman & Hall.

WHITE, L. (1975). Ph.D Thesis. Imperial College, London.

WHITTLE, P. (1973). Some General Points in the Theory of Optimal Experimental Design. Journal of the Royal Statistal Society Series B **35**, 123–130.

WYNN, H.P. (1970). The Sequential Generation of D optimal Experimental Designs. *Annals Mathematical Statistics* **41**, 1655–1664.

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

Appendix 1

PROOF OF PROPOSITION 3.1

1. Case $I : lim|x|^a v_F(x) \to 0$ as $x \to \infty$.
   Let $h(x) = \bar{F}(x)$. Then the following are true: $h(x) \geq 0, h'(x) \leq 0$ and $h(x) \to 0$ as $x \to \infty$. Since $F(x) \to 1$ as $x \to \infty$, we only need to show

$$\frac{|x|^a (h'(x))^2}{h(x)} \to 0 \ as \ x \to \infty. \tag{13}$$

   Suppose that the limiting value of (13) is either finite and non zero or infinite. Then there exists $K > 0$ and $k > 0$ such that

$$\frac{x^{a/2}|h'(x)|}{\sqrt{h(x)}} > k, \forall \ x > K.$$

   Now suppose $h'(x) \leq 0$; then for any $y > K$, we have

$$\frac{-h'(x)}{\sqrt{h(x)}} > kx^{-a/2} \Rightarrow -\int_K^y \frac{h'(x)}{\sqrt{h(x)}}dx > \int_K^y kx^{-a/2}dx$$

   The left hand integral is evaluated to $2\sqrt{h(K)} - \sqrt{h(y)}$, which has a limiting value of $2\sqrt{h(K)}$,
   as $y \to \infty$ since $h(y) \to 0$ as $y \to \infty$. However, the right hand integral becomes unbounded as $y \to \infty$ since the exponent $-a/2 \geq -1$ which contradicts the assumption. Therefore, we conclude that if (13) has a limiting value it must be zero.

2. Case $II : lim|x|^a v_F(x) \to 0$ as $x \to -\infty$.
   The proof is exactly as in Case I with $h(x) = F(x)$.

$\square$

Appendix 2

PROOF OF THEOREM 5.2

Maximizing the determinant of $\tilde{M}_D$ for the canonical positive–negative extreme value model using the NPSOL algorithm of Gill et al. (1998) yields $z_{11} = -0.8536657$, $z_{12} = 1.077288$, $z_{21} = -(z_{11} + \alpha_1)$,and $z_{22} = -(z_{12} + \alpha_1)$ with weights $\xi_{11} = \xi_{21} = 0.2895051$ and $\xi_{12} = \xi_{22} = 0.2104949$. To confirm that this solution is globally optimal, we appeal to the General Equivalence Theorem. That is, we show that $lim_{\alpha_1 \to -\infty} \mathcal{D}_{M_D} - 3$ is non-positive and its maximum is zero for $x \in \xi_D$.

$$\begin{aligned} \mathcal{D}_{\tilde{M}_D} &= Tr\left(\mathbf{I}(x)M_D^{-1}\right) \\ &= Tr\left(\mathbf{I}(x)\tilde{M}_D^{-1}\right) + Tr\left(\mathbf{I}(x)\left(M_D^{-1} - \tilde{M}_D^{-1}\right)\right). \end{aligned}$$

17

Because each element of $I(x)$ is a bounded function, $Tr[I(x)(M_D^{-1} - \tilde{M}_D^{-1})] \to 0$ as $\alpha_1 \to -\infty$. So $lim_{\alpha_1 \to -\infty} \mathcal{D}_{M_D} = lim_{\alpha_1 \to -\infty} \mathcal{D}_{\tilde{M}_D}$.

We now show $\mathcal{D}_{\tilde{M}_D} - 3 \leq 0$. Denote $v_F(x)$ and $\bar{F}(x)v_G(x)$ by v_x and w_x, respectively. For the canonical positive–negative extreme value model,

$$v_G(x) = \frac{\exp(-2x)\exp(-\exp(-x))}{1 - \exp(-\exp(-x))}.$$

Maple software provides the following equality:

$$
\begin{aligned}
\mathcal{D}_{\tilde{M}_D} =\ & v_x[-2.091898121\alpha_1 + 3.023967019\alpha_1^2 + 4.732702573] \\
& + 2v_x x[-1.045949058 + 3.023967013\alpha_1] + 3.023967012 v_x x^2 \\
& + 3.023967012 w_x x^2 + 2.091898118 w_x x + 4.732702591 w_x.
\end{aligned}
$$

Recall that that $w_x \leq v_G(x)$ for all x.

1. Case 1: $x < -\alpha_1/2$.

   Assuming different large negative values of α_1, for each sequence such that $x < -\alpha_1/2$ the following statements were verified by S-plus: $v_x < v_{-\alpha_1/2}$, $-(x + \alpha_1)v_x < (-\alpha_1/2)v_{-\alpha_1/2}$ and $(x + \alpha_1)^2 v_x < (\alpha_1^2/4)v_{-\alpha_1/2}$. Assuming these inequalities hold and keeping four significant digits, we rearrange terms to obtain

   $$
   \begin{aligned}
   \mathcal{D}_{\tilde{M}_D} =\ & 3.0240(x + \alpha_1)^2 v_x - 2.0919(x + \alpha_1)v_x + 4.7327 v_x \\
   & + (3.0240 x^2 w_x + 2.0919 x w_x + 4.7327 w_x) \\
   <\ & 3.0240(\alpha_1^2/4)v_{-\alpha_1/2} + 2.0919(-\alpha_1/2)v_{-\alpha_1/2} + 4.7327 v_{-\alpha_1/2} \\
   & + (3.0240 x^2 v_G(x) + 2.0919 x v_G(x) + 4.7327 v_G(x)) \\
   \leq\ & 3.0240(\alpha_1^2/4)v_{-\alpha_1/2} + 2.0919(-\alpha_1/2)v_{-\alpha_1/2} + 4.7327 v_{-\alpha_1/2} \\
   & + \max(3.0240 x^2 v_G(x) + 2.0919 x v_G(x) + 4.7327 v_G(x)) \\
   =\ & 3.0240(\alpha_1^2/4)v_{-\alpha_1/2} + 2.0919(-\alpha_1/2)v_{-\alpha_1/2} \\
   & + 4.7327 v_{-\alpha_1/2} + 3.
   \end{aligned}
   $$

2. Case 2: $x \geq (-\alpha_1/2)$.

   For each sequence of x values such that $x > -\alpha_1/2$ with different series of large negative values of α_1 the following statements hold. They were verified by S-plus: $v_G(x) < v_G(-\alpha_1/2)$, $x v_G(x) < (-\alpha_1/2)v_G(-\alpha_1/2)$ and $x^2 v_G(x) < (\alpha_1^2/4)v_G(-\alpha_1/2)$. Assuming these inequalities hold and keeping

four significant digits, we rearrange $\mathcal{D}_{\tilde{M}_D}$ as

$$
\begin{aligned}
\mathcal{D}_{\tilde{M}_D} \;=\; & 3.0240(x+\alpha_1)^2 v_x - 2.0919(x+\alpha_1)v_x + 4.7327 v_x \\
& + (3.0240 x^2 w_x + 2.0919 x w_x + 4.7327 w_x) \\
\leq \; & (3.0240(x+\alpha_1)^2 v_x - 2.0919(x+\alpha_1)v_x + 4.7327 v_x) \\
& + (3.0240 x^2 v_G(x) + 2.0919 x v_G(x) + 4.7327 v_G(x)) \\
< \; & (3.0240(x+\alpha_1)^2 v_x - 2.0919(x+\alpha_1)v_x + 4.7327 v_x) \\
& + 3.0240(\alpha_1^2/4) v_G(-\alpha_1/2) + 2.0919(-\alpha_1/2) v_G(-\alpha_1/2) \\
& + 4.7327 v_G(-\alpha_1/2) \\
\leq \; & \max(3.0240(x+\alpha_1)^2 v_x - 2.0919(x+\alpha_1)v_x \\
& + 4.7327 v_x) + 3.0240(\alpha_1^2/4) v_G(x)|(-\alpha_1/2) \\
& + 2.0919(-\alpha_1/2) v_G(-\alpha_1/2) + 4.7327 v_G(-\alpha_1/2) \\
< \; & 3 + 3.0240(\alpha_1^2/4) v_G(-\alpha_1/2) + 2.0919(-\alpha_1/2) v_G(-\alpha_1/2) \\
& + 4.7327 v_G(-\alpha_1/2.)
\end{aligned}
$$

Note that $v(-\alpha_1/2) = v_G(-\alpha_1/2) = \exp(u)\exp(-\exp(\alpha_1/2))/(1-\exp(-\exp(\alpha_1/2))) \to 0$ as $\alpha_1 \to -\infty$, which implies that $\mathcal{D}_{\tilde{M}_D} \to 3$.

Therefore, the directional derivative $\mathcal{D}_{\tilde{M}_D}$ is non-positive and the maximum is zero as $\alpha_1 \to -\infty$ and the proof of Theorem 5.2 is complete. $\qquad\square$

Appendix 3

PROOF OF THEOREM 5.3

The proof for the positive-negative extreme model with $\Theta = \{\alpha_1, \beta, \alpha_2\}$ involves showing that the limiting directional derivative is non-positive and has a maximum of zero at ξ_{E^*}. Denote this model by 3PNEM. The limiting directional derivative is given by

$$
\lim_{\alpha_1 \to -\infty} F_D = \lim_{\alpha_1 \to -\infty} \dot{\nu}^T M_{E^*}^{-1} I(x) \, M_{E^*}^{-1} \dot{\nu} - \dot{\nu}^T M_{E^*}^{-1} \dot{\nu}, \tag{14}
$$

where $I(x)$ is given in Lemma 3.1; M_{E^*} is Fisher's information given design ξ_{E^*}; the optimal dose is $\nu = \arg\max_x H(x) = -(\alpha_1 + \alpha_2)/2\beta$ for the PNEM with gradient $\dot{\nu} = \partial \nu / \partial(\alpha_1, \beta, \alpha_2) = \begin{pmatrix} -1/2\beta & (\alpha_1+\alpha_2)/2\beta^2 & -1/2\beta \end{pmatrix}^T$, which reduces to $\dot{\nu} = 1/2 \begin{pmatrix} -1 & \alpha_1 & -1 \end{pmatrix}^T$ for the canonical 3PNEM.

Using Maple software, we establish (with details below) that (14) simplifies to

$$
\lim_{\alpha_1 \to -\infty} F_D = v_G(x) v_{F1}^{-2} - v_{F1}^{-1}. \tag{15}
$$

Evaluated at $z_2^* = \arg\max_x[v_G(x)]$, the limit (15) attains its maximum of zero; this can be seen by noting that $z_1^* = \arg\max_x[v_F(x)] = -z_2^* - \alpha_1$ and $v_{G2} = v_{F1}$, where

v_{G2} is $v_G(x)$ evaluated at z_2^*. Thus the directional derivative (15) is non-positive. We then show that $v_G(x)$ has a unique maximum at $z_2^* = -0.466$ We conclude that z_2^* and z_1^* are the optimal design points and the proof of Theorem 5.3 is complete. $\qquad \square$

Appendix 4

Preliminary Results for the canonical 3PNEM

To simplify notation, let $\bar{F}_i = \bar{F}(z_i)$ and $w_i = \bar{F}_i v_{Gi}, i = 1, 2$. Note

$$\bar{F}_1 = G_2 = \exp\left(-e^{-z_2}\right);$$
$$v_{F1} = v_{G2} = \frac{\exp\left(-e^{-z_2}\right) e^{-2z_2}}{1 - \exp\left(-e^{-z_2}\right)};$$

The following limits, taken as $\alpha_1 \to -\infty$, can be easily verified:

$$G_1 = \bar{F}_2 = \exp\left(-e^{\alpha_1 + z_2}\right) \to 1;$$
$$v_{G1} = v_{F2} = \frac{\exp\left(-e^{\alpha_1 + z_2}\right) e^{-2(\alpha_1 + z_2)}}{1 - \exp\left(-e^{\alpha_1 + z_2}\right)} \to 0;$$
$$w_1 = \bar{F}_1 v_{F2} \to 0;$$
$$w_2 = \bar{F}_2 v_{G2} \to v_{F1}.$$

Consider $\bar{F}$ and ν_F as an exponential functions of α_1, then $\bar{F}$ goes to 1 and by Proposition 3.1 ν_F goes to zero. Hence the information matrix
$$I(x) = \begin{pmatrix} \nu_F & x\nu_F & 0 \\ x\nu_F & x^2(\nu_F + \bar{F}\nu_G) & x\bar{F}\nu_G \\ 0 & x\bar{F}\nu_G & \bar{F}\nu_G \end{pmatrix}$$
goes to $I(x) = v_G \begin{pmatrix} 0 & 0 & 0 \\ 0 & x^2 & x \\ 0 & x & 1 \end{pmatrix}$.

Using Lemma (3.1), the information in the two equally weighted points $\{z_1, z_2$ under the canonical 3PNEM can be written as

$$M_E = \frac{1}{2} \begin{pmatrix} v_{F1} + v_{F2} & -v_{F1}(z_2 + \alpha_1) + v_{F2}z_2 & 0 \\ -v_{F1}(z_2 + \alpha_1) & (v_{F1} + w_1)(z_2 + \alpha_1)^2 & w_2 z_2 \\ +v_{F2}z_2 & +(w_2 + v_{F2})z_2^2 & -w_1(z_2 + \alpha_1) \\ 0 & w_2 z_2 - w_1(z_2 + \alpha_1) & w_1 + w_2 \end{pmatrix}.$$

Maple software establishes the following propositions.

1. Proposition. $M_E^{-1} = \frac{1}{\det(M_E)}\left(m^{ij}\right)$, $i,j = 1,2,3$, where det denotes the determinant and

$$4\det(M_E) = (v_{F1}v_{F2}(w_1 + w_2) + (v_{F2} + v_{F1})w_1w_2)(z_2 - z_1)^2;$$
$$2m^{11} = (v_{F2}(w_1 + w_2) + w_2w_1)z_2^2$$
$$+ (v_{F1}(w_1 + w_2) + w_1w_2)(z_2 + \alpha_1)^2 + 2w_1w_2z_2(z_2 + \alpha_1);$$
$$2m^{12} = 2m^{21} = -v_{F2}(w_1 + w_2)z_2 + v_{F1}(w_1 + w_2)(z_2 + \alpha_1);$$
$$2m^{13} = 2m^{31} = v_{F2}w_2z_2^2 + v_{F1}w_2(z_2 + \alpha_1)^2$$
$$- (v_{F1}w_2 + v_{F2}w_1)z_2(z_2 + \alpha_1);$$
$$2m^{22} = v_{F2}(w_1 + w_2) + v_{F1}(w_1 + w_2);$$
$$2m^{23} = 2m^{32} = -(v_{F1} + v_{F2})w_2z_2 + (v_{F1} + v_{F2})w_1(z_2 + \alpha_1);$$
$$2m^{33} = (v_{F1}w_2 + v_{F1}v_{F2} + v_{F2}w_2)z_2^2$$
$$+ (v_{F1}v_{F2} + v_{F2}w_1 + v_{F1}w_1)(z_2 + \alpha_1)^2 + 2v_{F1}v_{F2}z_2(z_2 + \alpha_1).$$

2. Proposition. The row vector $Y = \dot{\nu}^T M_E^{-1}$ has elements $\{y_1, y_2, y_3\}$ determined by

$$-4\det(M_E)y_1 = 2(v_{F1}w_1 + v_{F2}w_2 + 2w_2w_1)z_2^2$$
$$+ (v_{F2}w_2 + 3v_{F1}w_1 + 4w_1w_2)\alpha_1z_2$$
$$+ (w_1w_2 + v_{F1}w_1)\alpha_1^2;$$
$$4\det(M_E)y_2 = 2(v_{F2}w_2 - v_{F1}w_1)z_2 + (v_{F2}w_2 - v_{F1}w_1)\alpha_1;$$
$$-4\det(M_E)y_3 = 2(2v_{F1}v_{F2} + v_{F1}w_1 + v_{F2}w_2)z_2^2$$
$$+ (4v_{F1}v_{F2} + 3v_{F1}w_1 + v_{F2}w_2)\alpha_1z_2$$
$$+ (v_{F1}v_{F2} + v_{F1}w_1)\alpha_1^2.$$

3. Combining results from Proposition 1 and 2 we have

$$8\,\dot{\nu}^T M_E^{-1}\,\dot{\nu} = \frac{(v_{F1}v_{F2} + v_{F1}w_2 + v_{F2}w_1 + w_2w_1)(z_2 - z_1)^2}{\det(M_E)};$$
$$\lim_{\alpha_1 \to -\infty} Y = \left(-\frac{2}{v_{F1}(1 + \exp(\exp(-z_2)))} \quad 0 \quad -\frac{1}{v_{F1.}} \right),$$

and finally (15) follows from

$$\lim_{\alpha_1 \to -\infty} \dot{\nu}^T M_E^{-1}\,\dot{\nu} = v_{F1}^{-1};$$
$$\lim_{\alpha_1 \to -\infty} \dot{\nu}^T M_E^{-1}I(x)M_E^{-1}\,\dot{\nu} = \lim_{\alpha_1 \to -\infty} Y^T I(x)Y = v_G(x)v_{F1}^{-2}.$$

Proof of $\arg\max_x v_G(x) = -0.466$
$\arg\max_x v_G(x) = -0.466$ We show that the maximum of $v_G(x)$ is unique and equals to -0.466. Define $u(x) = 1/v_G(x)$ which equals $e^{2x + e^{-x}} - e^{2x}$ for the positive-negative extreme value model. Now $\dot{u}(x) = e^{2x}h(x)$, where $h(x) = (2 - e^{-x})e^{-2x}$

21

-2. Note that $h(x)$ and $\dot{u}(x)$ have the same sign and $v_G(x)$ and $\dot{u}(x)$ have opposite sign. There is only one solution to $h(x) = 0$; it is between $x = -0.46601083$ and $x = -0.46601084$. We show that $\dot{u}(x)$ has a minimum at $x = 0.466$, that is, we need only to show that

$$\dot{u}(x) \le 0 \text{ when } x \le -0.466$$
$$\dot{u}(x) \ge 0 \text{ when } x \ge -0.466.$$

which means $v_G(x)$ has a maximum at $x = 0.466$. Because

$$\dot{h}(x) = e^{-x+e^{-x}}\left(e^{-x}-1\right)\begin{cases} \ge 0 & \text{when } x < 0 \\ \le 0 & \text{when } x > 0, \end{cases}$$

$h(x)$ is increasing when $x < 0$ and decreasing when $x > 0$. Because $h(-0.466) = 0$, $h(0) > 0$, $\lim_{x\to\infty} h(x) = 0$ and $h(x)$ is a decreasing function when $x \ge 0$, implies that $h(x) > 0$ when $x > 0$. Also, because $h(0) > 0$, $\lim_{x\to-\infty} h(x) = -\infty$ and $h(x)$ is a increasing function when $x \le 0$, it follows that $h(x) < 0$ when $x \le -.466$ and $h(x) > 0$ when $x \ge -.466$. Because $h(x)$ and $\dot{u}(x)$ have the same sign, it follows that $\dot{u}(x) \le 0$ when $x < -.466$ and $\dot{u}(x) \ge 0$ when $x > -.466$. $\qquad\square$

Appendix 5

TABLES

Table 3: Table D-optimal design points for the canonical positive–negative extreme value models with unequal slopes at selected values of α_1, and $\beta_1 = 3$ and 4.

α_1	$\beta_1 = 3$		$\beta_1 = 4$	
	Design Points	Weight	Design Points	Weight
3	-1.5878	0.5000	-1.3721	0.5000
	-0.8112	0.5000	-0.6477	0.5000
1	-1.2783	0.4145	-1.2976	0.308
	-0.5062	0.26675	-0.4986	0.3971
	-0.1083	0.3188	-0.0501	0.2949
0	-1.2187	0.3374	-1.2558	0.2796
	-0.2938	0.3749	-0.2914	0.4365
	0.2638	0.2878	0.2099	0.2839
-1	-1.1701	0.2956	-1.2133	0.2649
	-0.0434	0.4275	-0.07508	0.4578
	0.6136	0.2769	0.4649	0.2773
-3	-1.081	0.2621	-1.1278	0.2534
	0.5085	0.4702	0.3732	0.4768
	1.2903	0.2676	0.9683	0.2698
-5	-1.0481	0.2505	-0.9980	0.2530
	0.8347	0.4838	1.0965	0.4838
	1.4682	0.2657	1.9570	0.2632
-10	-0.8393	0.25	-0.8149	0.2499
	2.6464	0.4912	2.02326	0.489
	3.618	0.2588	2.7155	0.2611
-20	-0.9663	0.2464	-0.9337	0.2413
	1.3088	0.231	1.2397	0.1814
	6.1891	0.2716	4.6004	0.3235
	6.9873	0.251	5.2322	0.2538

Table 4: D-optimal design points for canonical positive–negative extreme value models with equal slopes at selected values of α_1; $\theta = (0, 1, \alpha_1)$.

α_1	Design Points	Weight	α_1	Design Points	Weight
5	-6.3369	0.3316	3	-4.1760	0.3333
	-4.0204	0.3305		-1.7889	0.6667
	-2.5862	0.3379			
2	-3.5131	0.3334	1	-1.3120	0.6164
	-1.2554	0.6667		0.0031	0.3836
-1	-0.5911	0.6496			
	1.8519	0.3504			
-2	-0.6450	0.4091	-5	-0.6986	0.3367
	0.5111	0.2675		2.101	0.3407
	2.7947	0.3233		5.6449	0.3226
-8	-0.8299	0.2949	-10	-0.8462	0.2897
	1.1613	0.2140		1.0914	0.2084
	6.8574	0.2001		8.9041	0.2113
	8.8203	0.2910		10.8483	0.2906
-12	-0.8522	0.2595	-20	-0.8537	0.2895
	1.080	0.2101		1.0773	0.2105
	10.9189	0.2107		18.9227	0.2105
	12.8526	0.2897		20.8537	0.2895
-25	-0.8537	0.2895	-30	-0.8537	0.2895
	1.0773	0.2105		1.0773	0.2105
	23.9227	0.2105		28.9223	0.2105
	25.8537	0.2895		30.8537	0.2895

Table 5: c-optimal design points for canonical positive–negative extreme value models with different slopes at selected values of α_1 and $\beta_1 = 0.5, 1, 4$.

α_1	$\beta_1 = 0.5$		$\beta_1 = 1$		$\beta_1 = 4$	
	Design Points	Weight	Design Points	Weight	Design Points	Weight
1			-1.4536	0.6030	-1.3700	0.6223
			0.1765	0.3970	-0.1773	0.3777
0	-0.5830	0.5265	-1.0323	0.5435	-1.2425	0.5736
	2.3137	0.4735	1.0106	0.4565	0.0110	0.4264
-1	0.1037	0.6005	-0.5643	0.5437	-1.1319	0.5185
	3.8163	0.3995	1.7731	0.4563	0.1980	0.4815
-3	1.4519	0.7756	0.3817	0.5890	-0.9528	0.4088
	6.4685	0.2244	2.9918	0.4110	0.5817	0.5912

Table 6: Continued c-optimal design points for canonical positive–negative extreme value models with different slopes at values of α_1 when $\beta_1 = 2$ and 3.

α_1	$\beta_1 = 2$		$\beta_1 = 3$	
	Design Points	Weight	Design Points	Weight
1	-1.4635	0.6294	-1.4163	0.6336
	-0.05416	0.3706	-0.1217	0.3664
0	-1.2133	0.5842	-1.2376	0.5923
	0.3745	0.4158	0.1390	0.4077
-1	-0.9531	0.5489	-1.0712	0.5415
	0.7419	0.4512	0.3743	0.4585
-3	-0.5054	0.4444	-0.8121	0.4184
	1.3595	0.5556	0.8361	0.5816
-5	-0.1536	0.3206	-0.6112	0.3081
	1.9746	0.6794	1.3181	0.6919
-7	0.2305	0.2370	-0.4127	0.2265
	2.6208	0.7631	1.8120	0.7735
-9	0.7077	0.1910	-0.1820	0.1704
	3.2801	0.8090	2.3099	0.8296

Table 7: c-optimal design points for positive–negative extreme value models with equal slopes at selected values of α_1

α_1	Design Points	Weight	α_1	Design Points	Weight
3	-4.2617	0.1729	-1	-0.1399	0.6299
	-1.6981	0.9827		1.5914	0.3702
-3	-0.3822	0.5162	-5	-0.4489	0.5024
	3.514	0.4838		5.4782	0.4977
-8	-0.4647	0.5001	-10	-0.4659	0.5001
	8.4670	0.4999		10.4663	0.5000
-12	-0.4600	0.5000	-15	-0.4600	0.50000
	12.4600	0.5000		15.46	0.5000

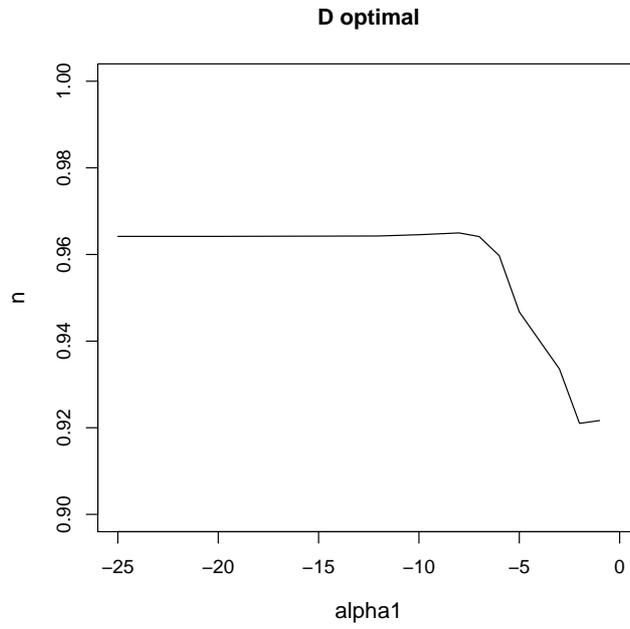# Appendix 6

FIGURES

D optimal

Figure 4: Efficiency plot for limiting D optimal designs for the canonical extreme value model when $\beta_1 = \beta_2 = 1$.