UNIVERSITY OF
KWAZULU-NATAL

D-optimal Designs for Drug Synergy

Muregancuro Gaëtan KABERA

Submitted in fulfillment of the academic
requirements for the degree of

DOCTOR OF PHILOSOPHY
in
Statistics

in the
School of Statistics and Actuarial Science
University of KwaZulu-Natal
Pietermariztburg
2009

Dedication

This thesis is dedicated to my relatives and many generous people who offered me unconditional love and support throughout the course of day-to-day life.

Declaration

The research work described in this thesis was carried out in the School of Statistics and Actuarial Science, University of KwaZulu-Natal Pietermaritzburg, under the supervision of Professor Dr Ndlovu and Professor Linda Haines.

This thesis presents original work by the author and has not been submitted in any form for any degree or diploma to any University. Where use has been made of the work of others it is duly acknowledged in the text.

November, 2009.

Student: ____________________ ____________________

Muregancuro Gaëtan Kabera Date

Supervisor: ____________________ ____________________

Dr Principal Ndlovu Date

Co-supervisor: ____________________ ____________________

Prof Linda Haines Date

Acknowledgements

In the first place I would like to express my deepest gratitude and appreciation to my supervisors, Dr Principal Ndlovu and Prof Linda Haines, for their guidance, patience and encouragement in diverse situations during the writing of this thesis. In particular, I owe my sincere gratitude to Prof Linda Haines for raising initial funds for this study.

I am very grateful to Prof Delia North, Prof Temesgen Zewotir, Dr Peter Njuho, Prof Carl Lombard and Ms Cathy Connolly who supplied moral and financial supports after expiration of funds granted by Prof Linda Haines. Special thanks to the National Research Foundation, the University of KwaZulu-Natal, and the Medical Research Council for providing grants.

I am thankful to Prof Timothy O'Brien for his fruitful advice and for providing important documentation related to this research work. Also thanks to Dr Samuel Manda and my wife Pénine Uwimbabazi who accepted to read the manuscript of this research work.

I owe much to staff members of the old School of Mathematics, Statistics and Information Technology of the University of KwaZulu-Natal Pietermaritzburg for their assistance during this period of hard work. Thank you Prof Johan Swart, Prof John van den Berg, Dr Henry Mwambi, Dr Petrovious Horton, Dr Shaun Ramrop, Mr Rob Dempster, Bev Bonhomme, Faith Nzimande, Jenny Wilson, Sebe Ndlovu and Shaun Moodley for your supports at various levels.

I am grateful to you my colleagues at MRC, especially at the Biostatistics Unit, as well at KIST and KIE, for your support in diverse situations. Special thanks to you Aloys, Alphonse, Anecto, Ayoub, Bernardo, Denis, Dikokole, Eugène, Faizi, Laurence, Makhala, Médard, Mohammed, Pedro, Saido, Sam, Shampiyona, Thekiso, Yongama, Zodwa, and fellow Rwandese and Eastern African Student Community on campus for your generosity and encouragements.

I am very proud of my country without which dignity and access to all rights are impossible.

Finally, I am thankful to my sisters and brothers, my wife Pénine and my sons Gustave, Gislain, Octave and late Jordan for love and tolerance to my lengthy absences from home.

Abstract

This thesis is focused on the construction of optimal designs for detecting drug interaction using the two-variable binary logistic model. Two specific models are considered: (1) the binary two-variable logistic model without interaction, and (2) the binary two-variable logistic model with interaction. The two explanatory variables are assumed to be doses of two drugs that may or may not interact when jointly administered to subjects. The main objective of the thesis is to algebraically construct the optimal designs. However, numerical computations are used for constructing optimal designs in cumbersome cases. The problem of constructing optimal designs is to allocate weights to specific points of the design space in such a way that information associated with model parameters is maximized and the variances of the mean responses are minimized. Specifically, the D-optimality criterion discussed in this thesis minimizes the determinant of the asymptotic variance-covariance matrix of the estimates of the model parameters. The number of support points of the D-optimal designs for the two-variable binary logistic model without interaction varies from 3 to 6. Support points are equally weighted only in case of the 3-point designs and in some special cases of the 4-point designs. The number of support points of the D-optimal designs for the two-variable binary logistic model with interaction varies from 4 to 8. Support points are equally weighted only in case of the 4-point designs and in some special cases of 8-point designs. Numerous examples are given to illustrate theoretical results.

Contents

List of Tables

List of Figures

1

Introduction

1.1. Background

Experiments on drug interaction are conducted in many research areas such as agriculture, pharmacology and medicine. Statistical modelling of responses to drug interactions is well developed in literature, and the models are now easily fitted using statistical software packages such as SAS, STATA, SPSS, GENSTAT and R. Errors in statistical analysis such as choosing the wrong model can be variously corrected. However, a poorly designed experiment such as a clinical trial is difficult or impossible to correct at later stages, and thus as a consequence results from a weakly planned experiment can be misleading. The focus of this thesis is on the construction of optimal designs for detecting drug interactions from which synergy is the desirable effect. The advantage of optimal experimental designs over classical experimental designs, such as factorial designs, is that they provide accurate or precise answers to experimental questions with minimum experimental effort or cost (see for example Atkinson, Donev and Tobias (2007, p. 7). There are many optimality criteria in the theory of optimal designs. The most applied optimality criterion is the D-optimality which maximizes information about all the parameters of a statistical model by minimizing in some sense the variance-covariance structure of the estimates of the parameters.

Even though the theory of optimal designs originated from the work of Smith (1918), it was not fruitful until the 1950s when much research work started to be published (Elfving (1952), Chernoff (1953), Guest (1958), Hoel (1958), Box and Lucas (1959), Kiefer (1958), Kiefer (1959), and Kiefer and Wolfowitz (1960)). Since the 1970s, a large number of articles on optimal designs emerged. Most of these were related to Kiefer's work, generally treating optimal design for linear fixed effect models. Kiefer (1974) gives a good summary of the findings.

There are quite a few textbooks in the English literature written on optimal designs. These are subdivided into two groups, textbooks and conference proceedings. As far as textbooks are concerned, a comprehensive pioneering textbook on optimal design is that by Fedorov (1972). The book by Silvey (1980) gives a very good summary while the books by Atkinson and Donev (1992), and Atkinson et al. (2007) give a more accessible account to the theory of optimal design. Other books are those by Pázman (1986), Shah and Sinha (1989), Pukelsheim (1993) and Fedorov and Hackl (1997). All these books focus on optimal designs for linear models. Some other general books that treat optimal design for linear and nonlinear models are those by Chernoff (1979), Ghosh and Rao (1996), Schwabe (1996), Dette and Studden (1997), Müller (1998), Liski, Mandal, Shah and Sinha (2001), and Seber and Wild (2003, Section 5.13).

There are also a number of published MODA (Model-Oriented Design and Analysis) conference proceedings. MODA proceedings go back over many years, and include: Dodge, Fedorov and Wynn (1988), Fedorov, Müller and Vuchnov (1992), Müller, Wynn and Zhigljavsky (1993), Kitsos and Müller (1995), Atkinson, Pronzato and Wynn (1998), Flournoy, Rosenberger and Wong (1998), Atkinson, Bogacka and Zhigljavsky (2001), Atkinson, Hackl and Muller (2001), Bucchianico, Lauter and Wynn (2004), and López-Fidalgo, Rodriguez-Diaz and Torsney (2007).

The theory of optimal design has been developed extensively for linear models because of the relative ease of constructing optimal designs for these models (see for example Fedorov (1972) and Silvey (1980)), and for generalized linear and nonlinear models but with one explanatory variable (see for example Abdelbasit and Plackett (1983), Wu (1988), Ford, Torsney and Wu (1992), Atkinson and Donev (1992, Chapter 22), Sitter (1992), Sitter and Wu (1993), Torsney and Musrati (1993), Haines (1993), Haines (1995), Sitter and Fainaru (1997), Sitter and Forbes (1997), and Smith and Ridout (2003)). However, very little research has been conducted on optimal designs for nonlinear and generalized linear models with more than one explanatory variable because the construction of optimal designs for such models is challenging. Some work has been done for generalized linear and nonlinear models for two variables without interaction (see for example Burridge and Sebastiani (1992), Burridge and Sebastiani (1994), Sitter and Torsney (1995a), Atkinson and Haines (1996), Myers, Montgomery and Vining (2002), Atkinson (2006), and Atkinson et al. (2007)), and for more than two variables without interaction (see for example Sitter and Torsney (1995b), and Torsney and Gunduz (2001)). Very little research work was done on the construction of optimal design for generalized linear and nonlinear models with two or more explanatory variables with interaction (see for example

Kupchak (2000), Jia and Myers (2001), Wang (2002), and Wang, Myers, Smith and Ye (2006)).

1.2. Research problem and objectives

Problem: Responses to drug interactions are generally modelled using generalized linear models (e.g. logistic and Poisson models) with at least two explanatory variables. This thesis is focused on constructing optimal designs for the two-variable binary logistic models with and without interaction. Once the theory is well established, generalization to binary logistic models with more than two explanatory variables or to some other models can be considered in further research.

Overall objective: The main aim of the study is to develop a new approach to analytically and numerically construct D-optimal designs for the two-variable binary logistic models with and without interaction on an unbounded and on a bounded design (dose) space.

Specific objectives

1. To construct D-optimal designs for the two-variable binary logistic model without interaction $\text{logit}(p) = \beta_0 + \beta_1 d_1 + \beta_2 d_2$ on the design space $\mathcal{D} = [0, \infty) \times [0, \infty)$ and establish conditions on the parameters β_0, β_1 and β_2 and design space $\mathcal{D}$ for the optimal design to be supported on 3 or 4 points.

2. To construct D-optimal designs for the two-variable binary logistic model without interaction $\text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the rectangular design space $\mathfrak{X} = [a, b] \times [c, d]$ where a, b, c and d are real numbers and set up conditions on the parameters β_0, β_1 and β_2 and design space $\mathfrak{X}$ for the optimal design to be supported on 3 to 6 points.

3. To construct D-optimal designs for the two-variable binary logistic model with interaction $\text{logit}(p) = \beta_0 + \beta_1 d_1 + \beta_2 d_2 + \beta_{12} d_1 d_2$ on the design spaces $\mathcal{D} = [0, \infty) \times [0, \infty)$ and provide conditions on the parameters β_0, β_1, β_2 and β_{12}, and on the design space $\mathcal{D}$ for the optimal design to be supported on 4 to 6 points.

4. To construct D-optimal designs for the two-variable binary logistic model with interaction $\text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $\mathfrak{X} = [a, b] \times [c, d]$ and provide conditions on the parameters β_0, β_1, β_2 and β_{12}, and on the design space $\mathfrak{X}$ for the optimal design to be supported on 4 to 8 points.

5. To illustrate the theory of D-optimal designs for the two-variable binary logistic model

without and with interaction using real world data relating to the detection of drug synergy.

1.3. Overview of the thesis

This thesis is subdivided into eight chapters. In Chapter 2 some models for drug interaction are reviewed. For each model, mathematical expressions of quantities termed relative potency and interaction indices are derived. In Chapter 3, a general review of the notion of optimal design is presented. Specifically the information matrices for generalized linear and nonlinear models are introduced, then the optimality criteria and associated equivalence theorems are presented. The equivalence theorem for each specific criterion is formulated. Since the focus of the present research work is to construct optimal designs for estimating all parameters in a given model, special attention is made on the D-optimality criterion. A general review, and extension from the one- to the two-variable binary logistic model without interaction is developed in Chapter 4. The approaches of constructing D-optimal designs for the two-variable binary logistic model without interaction discussed are those introduced in Sitter and Torsney (1995a), Atkinson and Haines (1996) and Jia and Myers (2001). These are shown to give similar results under certain conditions. The extension of existing work is particularly focussed on a new proof of D-optimality for the one-variable binary logistic model. In Chapter 5 a new approach for constructing D-optimal designs for the two-variable binary logistic model without interaction on an unbounded design space is introduced, and extends the approaches of Sitter and Torsney (1995a), and Jia and Myers (2001). Empirical constructions of D-optimal designs for the two-variable binary logistic model without interaction on a rectangular design space are introduced in Chapter 6. The purpose of the chapter is also to extend results found in Sitter and Torsney (1995a), and Jia and Myers (2001) and to rationalize the numerical results found in Atkinson and Haines (1996). Some results of the Chapter are used to illustrate the usefulness of optimal designs over non-optimal but widely used designs. The construction of D-optimal designs for the two-variable binary logistic model with interaction on unbounded and bounded design spaces is discussed in Chapter 7. The main focus is to introduce a new design approach which extends the methodology of constructing the 4-point D-optimal designs introduced by Jia and Myers (2001) to 4- up to 8-point D-optimal designs. Finally, conclusions are drawn in Chapter 8.

2

Review of Models for Detecting Drug Interaction

2.1. Introduction

This chapter reviews some models for detecting drug interaction. Section 2.2 presents the types of interaction that jointly administered drugs can exhibit. Section 2.3 reviews the notions of relative potency, isobologram and combination index usually used by pharmacologists to assess efficacy of drugs. Section 2.4 reviews models used in statistical literature for the detection of drug synergy. For each model, the expressions of relative potency and combination index are calculated when possible.

2.2. Types of interactions

The assessment of the joint action of two or more compounds, such as drugs, insecticides, herbicides, fungicides and diverse poisons, is of great interest in many biological studies. In this thesis, the terms compounds or drugs will be used invariably.

In practical investigations an experimenter has access to two or more compounds and may wish to examine the response or the effect on individuals or on any experimental material when these compounds are used jointly. In a medical framework, the effect can be the recovery or the death of a number of patients, while in agricultural studies the effect can be the death of a number of insects or the destruction of a certain quantity of weeds or other undesirable plants.

A number of terms is employed to differentiate the effects of compounds when administrated jointly to individuals. The most commonly used are additivity, independence, no interaction, synergy and antagonism. These terms are described in Plackett and Hewlett (1952), Hewlett and Plackett (1959), Plackett and Hewlett (1963), Hewlett (1969), Finney (1971, pp. 232-252), Abdelbasit and Plackett (1983), and a general review is given in Greco, Bravo and Parsons (1995).

Two compounds are said to be independent or additive if they do not interact when administered jointly. Their effect or response is the same as if each compound was used alone. The term additivity and independence are used respectively in reference to Loewe (1953) and Bliss (1939) for the concept of no interaction.

Two compounds are said to be interacting when their combined effect or response differs from that expected from additivity/independence. Synergy occurs if the joint effect is greater than that of the sum of the expected individual effects. Conversely antagonism occurs when the joint effect of the compounds is less than the sum of the expected individual effects (see for example Finney (1971, p. 231), Giltinan et al. (1988), Greco and Lawrence (1988) and Greco et al. (1995)). Tallarida (2000) gives a detailed account for drug interactions.

2.3. Relative potency, isobologram and combination index

The concepts of additivity, synergism and antagonism are better described through the notions of relative potency, isobologram, and combination index.

2.3.1 Relative potency

Relative potency when no model is specified

Two compounds, say A and B, are said to exhibit similar action if they behave as if they were the same except dilution or concentration effects. In other words, administering doses d_1 of A and d_2 of B is the same as administering dose $d_1 + \rho d_2$ of A alone. The coefficient ρ is called the relative potency of drug B compared to drug A. It represents the dose of the drug A required to produce the same effect as one unit dose of the drug B. Hence, if $p \in [0, 1]$ is the proportion of dose effect of equally effective doses of A and B, the relative potency of B with respect to A is mathematically expressed as:

$$\rho_p = \frac{ED_{100p,1}}{ED_{100p,2}} \tag{2.1}$$

where $ED_{100p,1}$ and $ED_{100p,2}$ are respective dose concentrations of drugs A and B which offer $100 \times p$ percent of response (see for example Finney (1971, p. 231), Abdelbasit and Plackett (1982), and Stokes, Davis and Koch (2000, p. 331)). If the relative potency ρ_p is the same at any equally effective doses, it is often taken to be the ratio of the median effective doses, denoted $ED_{50,1}$ and $ED_{50,2}$, of the compounds. By ED_{50} it is meant the amount of dose or concentration that causes a half effect of the range of the response. In medical studies, it is often denoted by LD_{50} (LD standing for lethal dose) and originally indicated the amount of drug or the concentration of drug that kills 50% of subjects in the sample under study. Hence, the relative potency of B with respect to A is:

$$\rho_{\frac{1}{2}} = \frac{ED_{50,1}}{ED_{50,2}} = \frac{\theta_{21}}{\theta_{22}} \tag{2.2}$$

where θ_{21} and θ_{22} are the ED_{50} of drug A and B respectively.

Relative potency with generalized linear or nonlinear models

For a class of models termed generalized linear models (GLM) (to be defined later), the respective general link functions between the probability of response p and the dose values d_1 and d_2 of drugs A and B are given by

$$g(p) = \beta_1(\ln d_1 - \ln \mu_1), \tag{2.3}$$

and

$$g(p) = \beta_2(\ln d_2 - \ln \mu_2) \tag{2.4}$$

where β_1 and β_2 are respective slope parameters for drugs A and B (see O'Brien (2004)). The median effective dose ED_{50} for each drug is obtained by noting that $g(p) = 0$ at $p = \frac{1}{2}$. As a consequence, the $ED_{50}'s$ for drugs 1 and 2 are respectively μ_1 and μ_2. Therefore, according to (2.2), the relative potency of drug 2 with respect to drug 1 is

$$\rho_{\frac{1}{2}} = \frac{\mu_1}{\mu_2} \tag{2.5}$$

In general, from

$$g(p) = \beta_i(\ln d_i - \ln \mu_i), \tag{2.6}$$

the ED_{100p}, denoted μ_{ip} for the two drugs are given by

$$\ln d_{i(p)} = \ln \mu_{ip} = \frac{g(p)}{\beta_i} + \ln \mu_i, \tag{2.7}$$

where $i = 1, 2$, and p is the proportion of response with $0 \le p \le 1$. As a consequence,

$$\ln\left(\frac{\mu_{1p}}{\mu_{2p}}\right) = g(p)\left(\frac{1}{\beta_1} - \frac{1}{\beta_2}\right) + \ln\left(\frac{\mu_1}{\mu_2}\right),$$

which is equivalent to

$$\ln(\rho_p) = g(p)\left(\frac{1}{\beta_1} - \frac{1}{\beta_2}\right) + \ln(\rho_{\frac{1}{2}}). \tag{2.8}$$

The condition that $\rho_p = \rho_{\frac{1}{2}} = \rho$, where ρ is constant, is that $\beta_1 = \beta_2$ in (2.8). Thus, the relative potency is defined only when the curves of the responses at each dose for the two drugs have equal slope. For example, if the logistic model is used the dilution assumption implies parallel lines on the logit scale. Assays that assume parallel relationship on the linear predictor are called parallel lines assays (see Stokes et al. (2000, p. 331)).

Note that the results (2.5) and (2.8) remain the same if d_i is used instead of $\ln d_i$, $i = 1, 2$.

2.3.2 Isobologram

It is convenient to consider the suggestion of Giltinan et al. (1988) that if the potency of B relative to A is ρ, a mixture containing d_1 units of drug A and d_2 units of drug B has an effect:
(1) equivalent to that for $d_1 + \rho d_2$ units of A, then the two drugs are said to exhibit an additive behavior;
(2) greater than that for $d_1 + \rho d_2$ units of A, then the two drugs are said to exhibit synergism;
(3) less than that for $d_1 + \rho d_2$ units of A, then the two drugs are said to exhibit antagonism.

One way of assessing relative effects of two drugs is drawing and examining isobolograms. The isobologram is a visual method generally used by pharmacologists to detect departure from additivity. The literature indicates that this method was introduced by Fraser in the 1870's (see for example Meadows, Gennings, Carter and Bae (2002)). An isobologram is a plot of the contours of constant responses, termed isoboles. As an example, in medical terminology consider the case where a dose combination (d_1, d_2) of drugs A and B kills a proportion p of germs in the target population with $0 \le p \le 1$. Then, the isobole associated with additivity or no interaction is the straight line joining the points at $100p$ percent effective doses $(ED_{100p,1}, 0)$ and $(0, ED_{100p,2})$ respectively, as in Figure 2.1. Synergy is claimed when the isobole is on the left side of the additivity line and antagonism is claimed otherwise.

If d_1 and d_2 are doses of drug A and B respectively that produce additive effect, Figure 2.1 indicates that additivity, synergy and antagonism are respectively observed when:

$$d_1 \ge 0, d_2 \ge 0, \frac{d_1}{ED_{100p,1}} + \frac{d_2}{ED_{100p,2}} = 1, \tag{2.9}$$

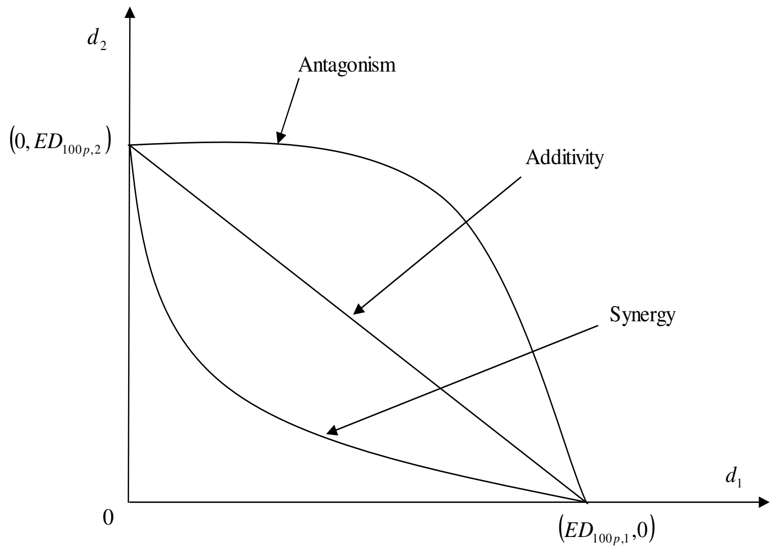


Figure 2.1: Illustration of an isobologram: the straight line corresponds to additivity, the left bowed line indicates synergy and the right bowed line indicates antagonism.

$$d_1 \geq 0, d_2 \geq 0, \frac{d_1}{ED_{100p,1}} + \frac{d_2}{ED_{100p,2}} < 1, \tag{2.10}$$

and

$$d_1 \geq 0, d_2 \geq 0, \frac{d_1}{ED_{100p,1}} + \frac{d_2}{ED_{100p,2}} > 1. \tag{2.11}$$

The method of the isobologram presents two major disadvantages. Firstly, as a graphical method, it is difficult to draw isoboles for assessing the relative effects of three or more drugs.

2.3.3 Interaction index or combination index

Consider two drugs A, B jointly given in amounts d_1 and d_2 to subjects for a defined purpose. Suppose that $ED_{100p,1}$ and $ED_{100p,2}$, with $0 \leq p \leq 1$, are the dose of A and B respectively giving $100p$ percent response of the desired effect. A measure of interaction introduced and termed "interaction index" by Berenbaum (1977), or "combination index" by Chou and Talalay (1984), usually denoted by κ is given by

$$\kappa = \frac{d_1}{ED_{100p,1}} + \frac{d_2}{ED_{100p,2}}. \tag{2.12}$$

It follows from equation (2.12) and the equations (2.9), (2.10) and (2.11) that two drugs interact additively, synergically or antagonistically respectively when $\kappa = 1$, $\kappa \leq 1$, or $\kappa \geq 1$.

Furthermore, Berenbaum (1981) showed that the interaction index is directly related to the isoblogram. In parallel lines assays, p is chosen as $p = \frac{1}{2}$ in (2.12) so that $ED_{100p,1}$ and $ED_{100p,2}$ are the median effective doses, ED_{50}s, of drugs A and B. A large review on the isobologram and combination index methodology is given in Berenbaum (1989).

An advantage of the interaction or combination index over the isobologram, is that it is defined for any number of drugs. The disadvantage of the interaction or combination index is that it does not take into account the variability inherent to the data. In his Ph.D. thesis, Kupchak (2000) derived some inference technique for assessing the interaction or combination index for the two- variable binary logistic model. In the following section, expressions of κ for different models will be given. The main objective is to show that all the values of κ have similar analytical expressions. As a consequence only one model will be studied in subsequent chapters for the construction of optimal designs.

2.4. Models for drug interaction

2.4.1 Effective dose models of Finney

(i) Continuous response

Suppose that two drugs A and B, in amounts d_1 and d_2, are applied jointly to subjects in a clinical trial study. O'Brien (2004) considered the model of Finney (1971, p. 262). One consideration in O'Brien (2004) was for a continuous response and it was stated that the effective dose for a mixture of the two drugs is defined as

$$z = d_1 + \theta_1 d_2 + \theta_2 \sqrt{\theta_1 d_1 d_2}, \tag{2.13}$$

where θ_1 is the potency of drug B relative to drug A, and θ_2 is the interaction parameter. Thus, additivity, synergy and antagonism occur respectively when $\theta_2 = 0$, $\theta_2 > 0$ and $\theta_2 < 0$. The model relating the continuous response y with the effective dose z is the 3-parameter log-logistic model given by

$$y = \frac{\beta_1}{1 + (\frac{z}{\beta_2})^{\beta_3}} + \epsilon, \tag{2.14}$$

where z is defined in equation (2.13) and ϵ is assumed to follow the normal distribution with zero mean and constant variance σ^2, i.e. $\epsilon \sim N(0, \sigma^2)$.

In model (2.14) with $\beta_3 > 0$, the parameter β_1 is the expected response at zero dose while β_2 and β_3 are the ED_{50} and the slope parameter respectively. When $\beta_3 < 0$, β_1 is the expected response at "infinity dose".

Taking the expectation of y, in (2.14) and denoting it by η, gives

$$\eta(z, \boldsymbol{\beta}) = \frac{\beta_1}{1 + (\frac{z}{\beta_2})^{\beta_3}} \tag{2.15}$$

which is a decreasing sigmoid when $\beta_3 > 0$ and an increasing sigmoid when $\beta_3 < 0$. This can be checked by noting that β_1 and β_2 are positive, and at any value of z,

$$\frac{\partial \eta}{\partial z} = \frac{-\beta_1 \beta_3 (\frac{z}{\beta_2})^{\beta_3}}{z[1 + (\frac{z}{\beta_2})^{\beta_3}]^2}.$$

Thus, the sign of the parameter β_3 determines the sign of the slope. Figure 2.2 illustrates the two situations. Also, by performing simple algebra, model (2.15) can be written as

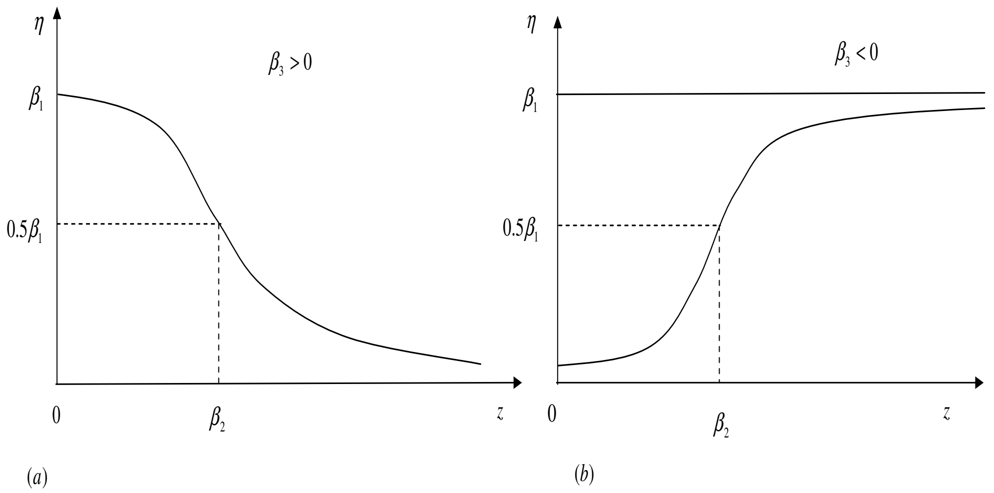


Figure 2.2: Graph of $\eta(z, \beta) = \frac{\beta_1}{1 + \left(\frac{z}{\beta_2}\right)^{\beta_3}}$: (a) $\beta_3 > 0$; (b) $\beta_3 < 0$.

$$\eta(x, \boldsymbol{\beta}) = \frac{\beta_1}{1 + \exp\{\beta_3(x - \mu)\}}, \tag{2.16}$$

where $x = \ln z$ and $\mu = \ln \theta_2$. In this case the concentration or dosage is measured on a log-scale.

A researcher may need to add a parameter β_4 to model (2.16) and get:

$$\eta(z, \boldsymbol{\beta}) = \beta_4 + \frac{\beta_1}{1 + (\frac{z}{\beta_2})^{\beta_3}}. \tag{2.17}$$

In this case $\eta = \beta_4$ is a lower horizontal asymptote at the right if $\beta_3 > 0$ or a lower horizontal asymptote at the left if $\beta_3 < 0$. Figure 2.3 illustrates the two cases. The ED_{50} for model

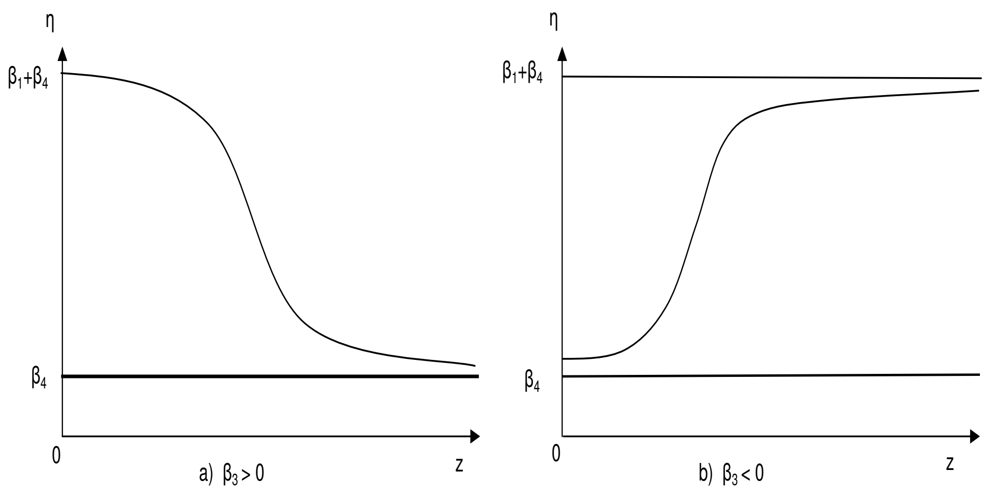


Figure 2.3: Graph of $\eta(z, \boldsymbol{\beta}) = \beta_4 + \dfrac{\beta_1}{1 + \left(\frac{z}{\beta_2}\right)^{\beta_3}}$: (a) $\beta_3 > 0$; (b) $\beta_3 < 0$.

(2.17), found by taking $\eta(z, \beta) = \dfrac{\beta_1 + \beta_4}{2}$, is

$$ED_{50} = \beta_2 \left(\frac{\beta_1 + \beta_4}{\beta_1 - \beta_4} \right)^{\frac{1}{\beta_3}}. \tag{2.18}$$

It follows from equation (2.18) that $ED_{50} = \beta_2$ if and only if $\beta_4 = 0$. The Finney model (2.15) in the continuous case was also discussed by Geric, Blum and Meier (1988) but they replaced the interaction parameter θ_2 by $2\theta_2$ in (2.13).

The potency of drug B relative to drug A with model (2.15) can be calculated as follows. Suppose that $d_2 = 0$, so that the effective dose given in (2.13) is $z = d_1$. Then, (2.6) can be written as

$$0 = \text{logit}\left(\frac{1}{2}\right) = \beta_3(\ln d_1 - \ln \mu)$$

so that the ED_{50} for drug 1 is $\mu_1 = \mu$. Likewise, taking $d_1 = 0$ and $z = \theta_1 d_2$ gives

$$0 = \text{logit}\left(\frac{1}{2}\right) = \beta_3(\ln \theta_1 d_2 - \ln \mu)$$

which implies that the ED_{50} for drug 2 is $\mu_2 = \dfrac{\mu}{\theta_1}$. Hence, according to (2.5), the relative potency is

$$\rho_{\frac{1}{2}} = \frac{\mu_1}{\mu_2} = \theta_1.$$

The combination index for model (2.15) can be calculated as follows. As θ_1 is the relative potency of drug 2 with respect to drug 1, the effective dose (2.13) written in terms of the dose of drug 1 alone at ED_{50} is

$$d_1 + \theta_1 d_2 + \theta_2 \sqrt{\theta_1 d_1 d_2} = \mu_1,$$

or equivalently as

$$\frac{d_1}{\mu_1} + \frac{\theta_1 d_2}{\mu_1} + \frac{\theta_2 \sqrt{\theta_1 d_1 d_2}}{\mu_1} = 1,$$

or, since $\mu_2 = \dfrac{\mu_1}{\theta_1}$, as

$$\frac{d_1}{\mu_1} + \frac{d_2}{\mu_2} + \theta_2 \sqrt{\frac{d_1 d_2}{\mu_1 \mu_2}} = 1. \tag{2.19}$$

Results (2.12) and (2.19) imply that the combination index for the Finney model is given by

$$\kappa = \frac{d_1}{\mu_1} + \frac{d_2}{\mu_2} = 1 - \theta_2 \sqrt{\frac{d_1}{\mu_1} \frac{d_2}{\mu_2}}. \tag{2.20}$$

Recall that in Finney models (2.15) with z given by (2.13), additivity, synergy and antagonism occur if $\theta_2 = 0$, $\theta_2 > 0$ and $\theta_2 < 0$ respectively. As a consequence, equation (2.20) indicates that additivity, synergy and antagonism occur when $\kappa = 1$, $\kappa < 1$ and $\kappa > 1$ respectively. Thus, the results obtained using the interaction parameter and the combination index are in agreement

(ii) Binary response

The binary response equivalent to (2.15), also discussed by O'Brien (2004) assumes that the response variable is binary with expected probability of success given by

$$p = \frac{(\frac{z}{\beta_2})^{\beta_3}}{1 + (\frac{z}{\beta_2})^{\beta_3}} \tag{2.21}$$

$$= \frac{1}{1 + \exp\{-\beta_3(x - \mu)\}} \tag{2.22}$$

where z is the effective dose defined in (2.13), $x = \ln z$ and $\mu = \ln \beta_2$.
If $\beta_3 > 0$, then p takes the maximum value of 1 as $z \to \infty$. Conversely, p approaches the minimum value of 0 as $z \to \infty$. These two cases are illustrated in Figure 2.4.

Comments regarding the combination indices for models (2.21) and (2.22) are similar to those of model (2.15). Model (2.21) or its equivalent (2.22) was discussed by Finney (1971) in the framework of probit analysis.

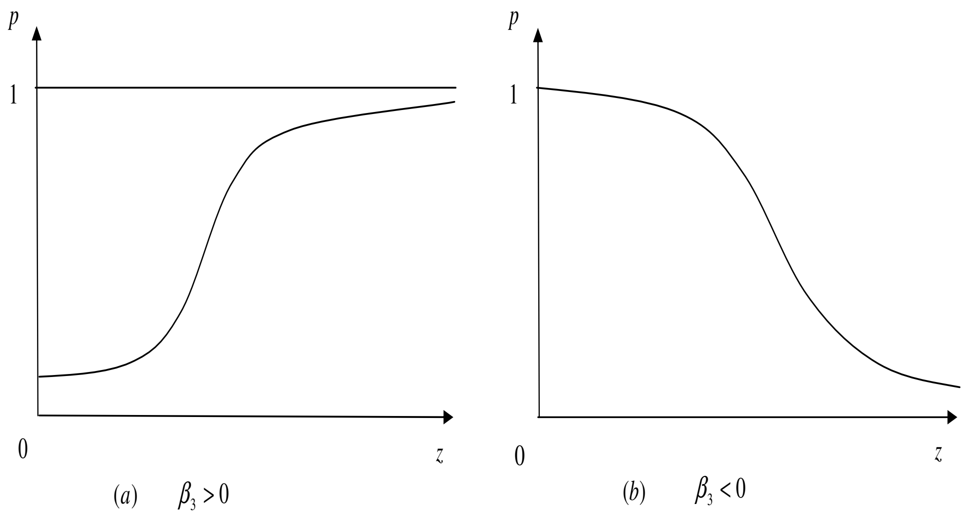


Figure 2.4: Finney model, binary response: (a) $\beta_3 > 0$, (b) $\beta_3 < 0$.

Giltinan et al. (1988) analyzed the Finney models (2.15) and (2.21) as generalized nonlinear models. They consider a mixture of d_1 units of drug A and d_2 units of drug B, where the relative potency of drug B with respect to drug A is ρ. Then, if Y is the response variable, Giltinan et al. (1988) stated that a model of simple similar action, i.e. no interaction, is

$$g(\mu) = \alpha + \beta \log(d_1 + \rho d_2) \tag{2.23}$$

where $\mu = E(Y)$ and g is the link function of the mean response variable Y and the explanatory variables D_1 and D_2 with values d_1 and d_2 respectively. The detection of synergy was obtained by comparing model (2.23) with the alternative model

$$g(\mu) = \alpha + \beta \log(d_1 + \rho d_2 + \gamma(\rho d_1 d_2)^{\frac{1}{2}}). \tag{2.24}$$

Additivity, synergy and antagonism hold respectively when $\gamma = 0$, $\gamma > 0$ and $\gamma < 0$.

2.4.2 Greco models

(i) Continuous response

Greco and Tung (1991) considered the case when two drugs A and B in amounts d_1 and d_2 are administered to subjects and the response of interest is a continuous random variable Y

with expected value E. The expected value of the effect of d_1 and d_2 was modelled as

$$\frac{d_1}{\mu_1[\frac{E}{E_c-E}]^{\frac{1}{m_1}}} + \frac{d_2}{\mu_2[\frac{E}{E_c-E}]^{\frac{1}{m_2}}} + \frac{\alpha d_1 d_2}{\mu_1\mu_2[\frac{E}{E_c-E}]^{(\frac{1}{2m_1}+\frac{1}{2m_2})}} = 1, \tag{2.25}$$

where μ_1 and μ_2 are median effective doses for drug A and B respectively, m_1 and m_2 are slope parameters for drug A and drug B respectively, E_c is the maximum value of E and α is the interaction parameter of the model.

Clearly, if $m_1 \neq m_2$ in model (2.25), it is not possible to find an explicit expression of E in terms of d_1 and d_2. However, if m_1=m_2=m, model (2.25) can be written as

$$E = \frac{E_c z^m}{1+z^m} \tag{2.26}$$

where $z = \dfrac{d_1}{\mu_1} + \dfrac{d_2}{\mu_2} + \alpha\dfrac{d_1 d_2}{\mu_1\mu_2}$. Note that model (2.26) is similar to the effective dose model of Finney given in (2.16) with $\beta_1 = E_c$, $\beta_2 = 1$ and $\beta_3 = m$.

The value of the combination index for model (2.26) is obtained by setting $E = \dfrac{E_c}{2}$. This gives

$$\frac{d_1}{\mu_1} + \frac{d_2}{\mu_2} + \alpha\frac{d_1}{\mu_1}\frac{d_2}{\mu_2} = 1. \tag{2.27}$$

Then, using the definition (2.12) and result (2.27), the combination index for the continuous model (2.26) is

$$\kappa = \frac{d_1}{\mu_1} + \frac{d_2}{\mu_2} = 1 - \alpha\frac{d_1}{\mu_1}\frac{d_2}{\mu_2}. \tag{2.28}$$

This result is similar to result (2.20) obtained for the combination index using the Finney model where the square root sign is omitted and θ_2 replaced by α.

It follows from equation (2.28) that if $\alpha = 0$, $\alpha > 0$ and $\alpha < 0$, then $\kappa = 1$, $\kappa < 1$ and $\kappa > 1$ respectively corresponding to additivity, synergy and antagonism.

(ii) Binary response

Greco and Lawrence (1988) considered the case when the response variable Y is binary, i.e. taking the coded values 0 and 1. If p is the probability of success then, $E(Y) = p$ since this is considered as a Bernoulli experiment. Similarly to the continuous model (2.25), Greco and Lawrence (1988) stated that the binary response model for the detection of drug interaction is

$$\frac{d_1}{\mu_1[\frac{p}{1-p}]^{\frac{1}{m_1}}} + \frac{d_2}{\mu_2[\frac{p}{1-p}]^{\frac{1}{m_2}}} + \frac{\alpha d_1 d_2}{\mu_1\mu_2[\frac{p}{1-p}]^{(\frac{1}{2m_1}+\frac{1}{2m_2})}} = 1 \tag{2.29}$$

where μ_1 and μ_2 are median effective doses for drug A and B respectively, m_1 and m_2 are slope parameters for drug A and drug B respectively, and α is the interaction parameter of the model. As in model (2.25), if $m_1 \neq m_2$ in model (2.29), it is not possible to find an explicit expression of p in terms of d_1 and d_2. However, if m_1=m_2=m, model (2.29) can be written explicitly as

$$p = \frac{z^m}{1 + z^m} \tag{2.30}$$

where $z = \dfrac{d_1}{\mu_1} + \dfrac{d_2}{\mu_2} + \alpha\dfrac{d_1 d_2}{\mu_1 \mu_2}$. Note that model (2.30) is similar to the binary response effective dose model of Finney (2.21) with $\beta_2 = 1$ and $\beta_3 = m$.

The value of the combination index for model (2.30) is obtained by setting $p = \dfrac{1}{2}$. This gives

$$\frac{d_1}{\mu_1} + \frac{d_2}{\mu_2} + \alpha\frac{d_1}{\mu_1}\frac{d_2}{\mu_2} = 1 \tag{2.31}$$

and, thus the combination index is

$$\kappa = \frac{d_1}{\mu_1} + \frac{d_2}{\mu_2} = 1 - \alpha\frac{d_1}{\mu_1}\frac{d_2}{\mu_2} \tag{2.32}$$

which is the same as (2.28) found for the continuous response case.

2.4.3 Plackett-Hewlett model

Giltinan et al. (1988) suggested that an alternative model to (2.24) is the model introduced by Plackett and Hewlett (1952) defined as

$$g(\mu) = \alpha + \beta\eta^{-1}\log(d_1^\eta + \rho^\eta d_2^\eta), \tag{2.33}$$

where η is the interaction parameter and $\rho = \dfrac{\mu_1}{\mu_2}$ is the relative potency of drug B with respect of drug A, μ_1 and μ_2 being the median effective doses of drugs A and B respectively. Additivity, synergy and antagonism are observed when $\eta = 1$, $\eta < 1$ and $\eta > 1$ respectively. The nature of interaction can be explained as follows. Since μ_1 is obtained by taking $g(\mu) = 0$, and $d_2 = 0$ in model (2.33), it follows that $\mu_1 = e^{-\frac{\alpha}{\beta}}$. As a consequence, taking $g(\mu) = 0$, equation (2.33) becomes

$$\left(\frac{d_1}{\theta_{21}}\right)^\eta + \left(\frac{d_2}{\theta_{22}}\right)^\eta = 1. \tag{2.34}$$

Equation (2.34) is referred to in many articles as the Hewlett-Plackett model (see for example Hewlett (1969), and Machado and Robinson (1994)). Clearly, additivity, synergy and antagonism occur respectively when $\eta = 1$, $\eta < 1$ and $\eta > 1$ because these three cases are represented by a straight line, a hyperbolic curve, and an elliptic curve, and hence agree with the isobologram of Figure 2.1. However, it is not possible to express (2.34) as the combination indices given by (2.20) and by (2.32).

2.4.4 Response surface models

The most popular response surface model for drug interaction is the logistic model. This involves the continuous response case and the binary response case.

(i) Continuous Response

Suppose that two drugs A and B are jointly administered to subjects in doses d_1 and d_2 respectively. The expected response of interest in a continuous scale is the semi-nonlinear model defined as

$$\eta(\boldsymbol{d}, \boldsymbol{\beta}) = \frac{\beta_4}{1 + \exp(-\beta_0 - \beta_1 d_1 - \beta_2 d_2 - \beta_{12} d_1 d_2)}, \tag{2.35}$$

or, equivalently as

$$\log \frac{\eta}{\beta_4 - \eta} = \beta_0 + \beta_1 d_1 + \beta_2 d_2 + \beta_{12} d_1 d_2. \tag{2.36}$$

Model (2.35) or its equivalent (2.36) is known as the continuous logistic model. The interaction parameter is β_{12}. In the model (2.36) the interaction term $d_1 d_2$ is not in the same units as the individual doses d_1 and d_2. In addition, the model includes a placebo effect, obtained by setting $d_1 = d_2 = 0$, and its value is

$$\eta(\mathbf{0}, \boldsymbol{\beta}) = \frac{\beta_4}{1 + \exp(-\beta_0)}. \tag{2.37}$$

Since $\exp(-\beta_0) > 0$, the quantity (2.37) is always less than β_4. Large positive values of β_0 renders (2.37) close to β_4, and large negative values of β_0 renders (2.37) close to 0. Figure 2.5 is a graphical illustration of the value of $\eta(\mathbf{0}, \boldsymbol{\beta})$ defined by (2.37) for all possible values of β_0. Since, for active drugs, the placebo effect must be negligible, β_0 must be negative and large in absolute value.

The potency of drug B relative to drug A can be calculated as follows. Taking $\eta = \frac{\beta_4}{2}$ in model (2.36), $d_2 = 0$, or $d_1 = 0$ gives the ED_{50}'s for drugs A and B respectively, i.e.

$$\mu_1 = ED_{50,1} = -\frac{\beta_0}{\beta_1} \qquad \mu_2 = ED_{50,2} = -\frac{\beta_0}{\beta_2}. \tag{2.38}$$

These results assumed to be positive have a meaning when β_0 has opposite sign to that of β_1 and β_2. As discussed that $\beta_0 < 0$, then $\beta_1 > 0$ and $\beta_2 > 0$. However, in real world situations, cases when $\beta_0 > 0$, $\beta_1 < 0$ and $\beta_2 < 0$ can also exist. In such cases the placebo effect, response at zero doses, has the highest value. An example is when the response variable is the amount of bacteria colonies remaining in a patient's body after a certain period with antibiotic treatments. The amount of colonies is expected to be higher when no treatment is given, then decrease with the use of treatments.

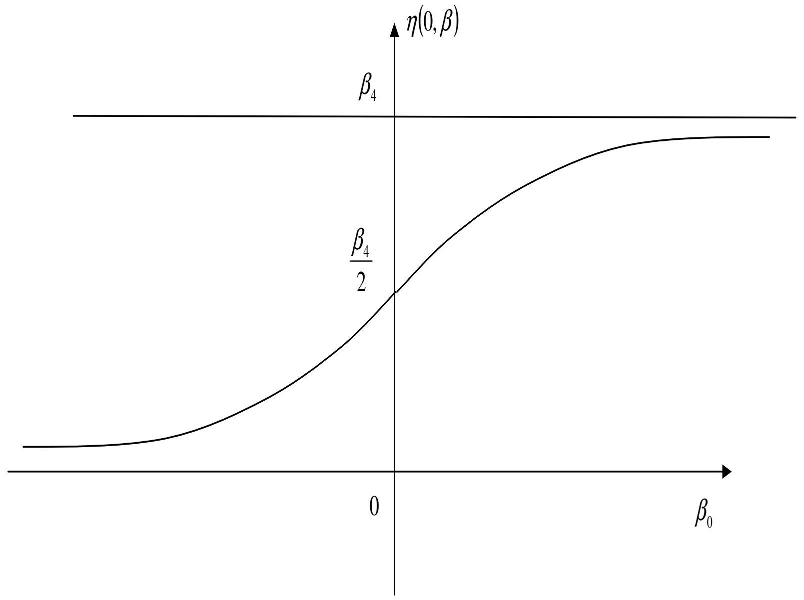


Figure 2.5: Graph of $\eta(\mathbf{0}, \boldsymbol{\beta}) = \dfrac{\beta_4}{1+\exp(-\beta_0)}$.

Now, according to (2.2), the potency of drug B relative to drug A is

$$\rho = \frac{\mu_1}{\mu_2} = \frac{\beta_2}{\beta_1} \tag{2.39}$$

The combination index for model (2.35) or (2.36) can be calculated as follows. Setting η to $\frac{\beta_4}{2}$ in model (2.36) gives

$$\frac{d_1}{\mu_1} + \frac{d_2}{\mu_2} + \alpha \frac{d_1}{\mu_1}\frac{d_2}{\mu_2} = 1, \tag{2.40}$$

where μ_1 and μ_2 are the respective median effective doses given by (2.38) and $\alpha = -\dfrac{\beta_0\beta_{12}}{\beta_1\beta_2}$ is the interaction coefficient. Thus, according to (2.12), the combination index for model (2.35) is

$$\kappa = \frac{d_1}{\mu_1} + \frac{d_2}{\mu_2} = 1 - \alpha \frac{d_1}{\mu_1}\frac{d_2}{\mu_2} \tag{2.41}$$

which is the same expression as those given in (2.28) and (2.32). Recall from Section 2.3.3 that additivity, synergy and antagonism are observed if $\kappa = 1$, $\kappa < 1$ and $\kappa > 1$ respectively which respectively correspond to $\alpha = 0$, $\alpha > 0$ and $\alpha < 0$. Therefore, if $\beta_1 > 0$ and $\beta_2 > 0$, results (2.41) where $\alpha = -\dfrac{\beta_0\beta_{12}}{\beta_1\beta_2}$ lead to the following two cases.

(1) If $\beta_0 < 0$, then additivity, synergy and antagonism are respectively observed when $\beta_{12} = 0$, $\beta_{12} > 0$ and $\beta_{12} < 0$.

(2) If $\beta_0 > 0$, then additivity, synergy, and antagonism are observed when $\beta_{12} = 0$, $\beta_{12} < 0$,

and $\beta_{12} > 0$ respectively. These results indicate that there are two cases of synergy for model (2.35) corresponding to (1) $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_{12} > 0$, and (2) $\beta_0 > 0$, $\beta_1 < 0$, $\beta_2 < 0$ and $\beta_{12} < 0$, and two cases of antagonism corresponding to (1) $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_{12} < 0$, and (2) $\beta_0 > 0$, $\beta_1 < 0$, $\beta_2 < 0$ and $\beta_{12} > 0$. Since it was agreed to only discuss the case when $\beta_0 < 0$, in this subsequent chapters synergy will be discussed when $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_{12} > 0$, and antagonism will be discussed when $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_{12} < 0$.

(ii) Binary response

If the response variable Y is binary with probability of response p, the continuous logistic model (2.35) is replaced by the binary logistic model

$$p = \frac{1}{1 + \exp(-\beta_0 - \beta_1 d_1 - \beta_2 d_2 - \beta_{12} d_1 d_2)}, \tag{2.42}$$

or, equivalently,

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 d_1 + \beta_2 d_2 + \beta_{12} d_1 d_2 \tag{2.43}$$

with $0 \leq p \leq 1$. Note that this model is similar to (2.35), or equivalently to (2.36), except that the parameter β_4 is fixed to 1. The placebo effect, the median effective doses, the relative potency and the combination index are calculated in a similar way as for the continuous case. The results are the same as those found for the continuous logistic model, but β_4 is taken to be 1. In his PhD thesis, Kupchak (2000) used model (2.42) for modelling drug interaction, but assumed $\beta_{12} = 0$ in almost all the cases for the construction of optimal designs.

2.5. Conclusions

This chapter has reviewed some measures and models for detecting drug interaction. Drug interaction can be detected in different ways. One way is a visual approach by observing an isobologram which in turn is related to a descriptive measure termed combination index or interaction index. A statistical approach of assessing drug interaction is to calculate the coefficient of interaction or the interaction parameter. The relationship between the interaction index and the interaction parameter was derived using various models. Specifically, four models were discussed: (1) the effective dose models of Finney; (2) the Greco models; (3) the Plackett-Hewlett model; and (4) the response surface models. It was shown that these models give similar information with respect to values of the combination index and the interaction parameter. Observed difficulties were that the Plackett-Hewlett model is constructed differently from others, and it is in general difficult to write the mean response for the Greco model. In this thesis, drug interaction will be detected using the interaction parameter rather than the combination index since fitting models which incorporate interaction terms is very well developed in statistical literature and statistical packages are now widespread. As the models discussed in this chapter have many similarities, the theory of optimal designs will be illustrated using the binary response case of the response surface models.

3

Review of Optimal Designs

3.1. Introduction

This chapter reviews the theory of optimal designs. The aim of optimal design theory is to provide values of explanatory variables in a model such that the information about the parameters is provided as precisely as possible by reducing noise on these parameters. The theory of optimal designs is well developed for linear models (see for Example Fedorov (1972) and Silvey (1980)). However, generalized linear and nonlinear models are known to be suitable for fitting some kind of data (see for example McCullagh and Nelder (1989), Dobson (2002), Myers et al. (2002), and also Seber and Wild (2003)). Specifically, generalized linear models and nonlinear models are useful to fit data for the detection of drug interaction (see for example Finney (1971) and Tallarida (2000)). This chapter reviews the theory of optimal designs for linear, nonlinear and generalized linear models and is structured as follows. Section 3.2 gives a brief review of the definition of a generalized linear model, and associated information matrix for the model parameters. Section 3.3 provides an extension of generalized linear models to generalized nonlinear models. Section 3.4 gives examples of models which belong to generalized linear and nonlinear models. Section 3.5 gives an overview of the theory of optimal designs, and specifically presents popular optimality criteria. Section 3.6 discusses the construction of optimal designs for linear models. Section 3.8 reviews the methodology of constructing optimal designs for nonlinear models. Finally, Section 3.9 provides concluding remarks on the Chapter.

3.2. Generalized linear models and information matrix

3.2.1 Exponential family of distributions

Consider a random variable Y with probability density function (pdf) $f(y;\theta)$ that depends on a single parameter θ. The distribution of Y belongs to the exponential family if $f(y;\theta)$ can be written in the following form

$$f(y,\theta) = \exp[a(y)b(\theta) + c(\theta) + d(y)] \tag{3.1}$$

where a, b, c, and d are known functions. If $a(y) = y$, the distribution of Y is said to be in canonical form and $b(\theta)$ is called the natural parameter of the distribution. The normal distribution, the poisson distribution, and the binomial distribution belong to the exponential family of distributions since the pdf of each of them can be written in the form (3.1) (see for example Dobson (2002, pp. 44-46)). The mean and the variance of the variable $a(Y)$ are given by

$$\mu = E[a(Y)] = -\frac{c'(\theta)}{b'(\theta)} \quad \text{and} \quad \mathrm{Var}[a(Y)] = \frac{b''(\theta)c'(\theta) - c''(\theta)b'(\theta)}{[b'(\theta)]^3}$$

(Dobson (2002, p. 47)).

3.2.2 Generalized linear models

Consider a set of N independent random variables Y_1, Y_2, …, Y_N with a distribution that belongs to the exponential family. The following features define a generalized linear model (GLM) (Nelder and Wedderburn (1972)).

1. *Random component:* The random variables Y_i belong to the same exponential family in the canonical form and are such that each of them has its own single parameter θ_i. That is

   $$f(y_i;\theta_i) = \exp[y_i b(\theta_i) + c(\theta_i) + d(y_i)],\ i = 1, 2, \dots, N.$$

   In principle not all the parameters θ_i are involved, but a set of parameters of interest $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_p)^T$ with $p < N$.

2. *Deterministic or systematic component:* The linear predictor

   $$\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip} = \boldsymbol{x}_i^T \boldsymbol{\beta}$$

gives a linear combination of values of the explanatory variables X_1, X_2, ..., X_p for $i = 1, 2, \ldots, N$. The involvement of the linear predictor is the source of the terminology *generalized linear models*. It should be noted that if Y_i belongs to the exponential family and is in canonical form then $b(\theta_i) = \boldsymbol{x}_i^T \boldsymbol{\beta}$ for $i = 1, 2, \ldots, N$.

3. *Link function:* The association between the mean response $\mu_i = E(Y_i)$ and the linear predictor η_i is

$$g(\mu_i) = \eta_i = \boldsymbol{x}_i^T \boldsymbol{\beta}$$

where g is a monotone, differentiable function called the *link function* (see McCullagh and Nelder (1989, p. 27), Dobson (2002, p. 49), and Myers et al. (2002, p. 161)). The link functions for common distributions are listed for example in McCullagh and Nelder (1989, p. 31), and in (Myers et al., 2002, p. 162).

3.2.3 Information matrix

Consider independent random variables Y_1, Y_2, ..., Y_N which satisfy the conditions of a GLM, and suppose that $\mu_i = E(Y_i)$. The likelihood function is defined by

$$L(\boldsymbol{\beta}; \boldsymbol{y}) = \prod_{i=1}^{N} f(y_i; \theta_i)$$

where $\boldsymbol{y}$ is the $N \times 1$ vector of observed responses, and the elements of the parameter vector $\boldsymbol{\beta}$ are related to the parameters θ_i through the link function. Thus, the log-likelihood can be expressed as

$$l(\boldsymbol{\beta}; \boldsymbol{y}) = \ln L(\boldsymbol{\beta}; \boldsymbol{y}) = \sum_{i=1}^{N} \{y_i b(\theta_i) + c(\theta_i)\}.$$

The information matrix for $\boldsymbol{\beta}$ at a single vector $\boldsymbol{x} = (1, x_1, x_2, \ldots, x_p)^T$ is defined as

$$M(\boldsymbol{x}; \boldsymbol{\beta}) = -E\left(\frac{\partial^2 l(\boldsymbol{\beta}; \boldsymbol{y})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T}\right)$$

and is readily shown to be

$$M(\boldsymbol{x}; \boldsymbol{\beta}) = \frac{1}{Var(Y)} \left(\frac{\partial \mu}{\partial \boldsymbol{\beta}}\right) \left(\frac{\partial \mu}{\partial \boldsymbol{\beta}}\right)^T = \frac{1}{Var(Y)} \left(\frac{\partial \mu}{\partial \eta}\right)^2 \left(\frac{\partial \eta}{\partial \boldsymbol{\beta}}\right) \left(\frac{\partial \eta}{\partial \boldsymbol{\beta}}\right)^T = \frac{1}{Var(Y)} \left(\frac{\partial \mu}{\partial \eta}\right)^2 \boldsymbol{x}\boldsymbol{x}^T \tag{3.2}$$

where $\left(\dfrac{\partial \mu}{\partial \boldsymbol{\beta}}\right)$ is a $(p+1) \times 1$ vector of partial derivatives of μ with respect to each of component of the vector $\boldsymbol{\beta}^T = (\beta_0, \beta_1, \ldots, \beta_p)$, $\left(\dfrac{\partial \mu}{\partial \eta}\right)$ is a scalar and $\mu = g^{-1}(\eta)$ since g is a monotone function (Dobson (2002, p. 62-63). Thus, for the independent response variables Y_i at $\boldsymbol{x}_i$ with $i = 1, 2, \ldots, N$, the information matrix for $\boldsymbol{\beta}$ is given by

$$M(\boldsymbol{X}; \boldsymbol{\beta}) = \sum_{i=1}^{N} \frac{1}{Var(Y_i)} \left(\frac{\partial \mu_i}{\partial \eta_i} \right)^2 \boldsymbol{x}_i \boldsymbol{x}_i^T. \tag{3.3}$$

The information matrix (3.3) can be written in succinct form as

$$M(\boldsymbol{X}; \boldsymbol{\beta}) = \boldsymbol{X}^T \boldsymbol{W} \boldsymbol{X} \tag{3.4}$$

where $\boldsymbol{W}$ is the diagonal matrix with (i, i)th element $\frac{1}{Var(Y_i)} \left(\frac{\partial \mu_i}{\partial \eta_i} \right)^2$ and $\boldsymbol{X}$ is the $N \times (p+1)$ design matrix with ith row $\boldsymbol{x}_i^T = (1, x_{1i}, \ldots, x_{pi})$ (Dobson (2002, p. 64)).

3.3. Generalized nonlinear models

The definition of the generalized nonlinear model (GnLM) is the same as that of the generalized linear model except that the predictor $\eta_i = \eta(\boldsymbol{x}_i; \boldsymbol{\beta})$ for $i = 1, 2, \ldots, N$ is introduced with η a nonlinear function in some or all the parameters $\boldsymbol{\beta}$. Writing $g(\mu_i) = \eta(\boldsymbol{x}_i; \boldsymbol{\beta})$ gives $\mu_i = g^{-1}(\eta(\boldsymbol{x}_i; \boldsymbol{\beta})) = h(\boldsymbol{x}_i; \boldsymbol{\beta})$. Therefore, μ_i is itself a nonlinear function. Similarly to (3.2), the information matrix for $\boldsymbol{\beta}$ at a single point $\boldsymbol{x}$ is given by

$$M(\boldsymbol{x}; \boldsymbol{\beta}) = \frac{1}{Var(Y)} \left(\frac{\partial \mu}{\partial \boldsymbol{\beta}} \right) \left(\frac{\partial \mu}{\partial \boldsymbol{\beta}} \right)^T = \frac{1}{Var(Y)} \left(\frac{\partial \mu}{\partial \eta} \right)^2 \left(\frac{\partial \eta}{\partial \boldsymbol{\beta}} \right) \left(\frac{\partial \eta}{\partial \boldsymbol{\beta}} \right)^T \tag{3.5}$$

where $\left(\frac{\partial \eta}{\partial \boldsymbol{\beta}} \right)$ is a $(p+1) \times 1$ vector of partial derivatives of η with respect to the $(p+1)$ parameters β_i, and $\frac{1}{Var(Y)} \left(\frac{\partial \mu}{\partial \eta} \right)^2$ is a scalar. Thus for the independent responses Y_i at $\boldsymbol{x}_i$ with $i = 1, 2, \ldots, N$, the information matrix for $\boldsymbol{\beta}$ is given by

$$M(\boldsymbol{X}; \boldsymbol{\beta}) = \sum_{i=1}^{N} M(\boldsymbol{x}_i; \boldsymbol{\beta}) = \sum_{i=1}^{N} \frac{1}{Var(Y_i)} \left(\frac{\partial \mu_i}{\partial \eta_i} \right)^2 \left(\frac{\partial \eta_i}{\partial \boldsymbol{\beta}} \right) \left(\frac{\partial \eta_i}{\partial \boldsymbol{\beta}} \right)^T = \boldsymbol{F}^T \boldsymbol{W} \boldsymbol{F} \tag{3.6}$$

where $\boldsymbol{W}$ is the $N \times N$ diagonal matrix with the (i, i)th element $\frac{1}{Var(Y_i)} \left(\frac{\partial \mu_i}{\partial \eta_i} \right)^2$ and $\boldsymbol{F}$ is the $N \times (p+1)$ matrix with the ith row $\left(\frac{\partial \eta_i}{\partial \boldsymbol{\beta}} \right)^T = \left(\frac{\partial \eta_i}{\partial \beta_0}, \frac{\partial \eta_i}{\partial \beta_1}, \ldots, \frac{\partial \eta_i}{\partial \beta_p} \right)$. Clearly, the information matrix (3.4) is a special case of the information matrix (3.6).

3.4. Examples

Example 3.1. *Normal linear model*

In this case,
(a) $Y_i \sim N(\mu_i; \sigma^2)$ and the normal distribution is a member of the exponential family (Dobson (2002, p. 45)).
(b) The linear predictor can be expressed as $\eta = \boldsymbol{x}^T\boldsymbol{\beta}$.
(c) The link function g is the identity, i.e. $\mu_i = \eta_i = \boldsymbol{x}_i^T\boldsymbol{\beta}$.
The information matrix (3.2) at a single point $\boldsymbol{x}$ follows immediately from the fact that $Var(Y_i) = \sigma^2$, $\left(\dfrac{\partial \mu_i}{\partial \eta_i}\right) = 1$ and $\dfrac{\partial \eta_i}{\partial \boldsymbol{\beta}} = \boldsymbol{x}$, and thus the information matrix at point $\boldsymbol{x}$ can be written as $M(\boldsymbol{x}) = \dfrac{1}{\sigma^2}\boldsymbol{x}\boldsymbol{x}^T$. Clearly, this matrix does not depends on $\boldsymbol{\beta}$, and thus the notation $M(\boldsymbol{x};\boldsymbol{\beta})$ can be replaced by the notation $M(\boldsymbol{x})$. The information matrix for N observations at $\boldsymbol{x}_i$, $i = 1, 2, \ldots, N$ is then

$$M(\boldsymbol{X}) = \frac{1}{\sigma^2}\boldsymbol{X}^T\boldsymbol{X} \tag{3.7}$$

where $\boldsymbol{X}$ is an $N \times (p+1)$ matrix whose ith row is $\boldsymbol{x}_i^T = (1, x_{i1}, x_{i2}, \ldots, x_{ip})$.

Example 3.2. *Linear binary logistic model*

(a) Suppose that the response variable Y is binary, taking values such as success and failure coded 1 and 0, i.e. for a single observation, the random variable Y follows a Bernoulli distribution with probability of success π. Therefore, for n_i observations at $\boldsymbol{x}_i$, the random variable Y_i, with $i = 1, 2, \ldots, N$, follows a binomial distribution with probability of success π_i, that is $Y_i \sim \text{Binomial}(n_i, \pi_i)$. In consequence, $E(Y_i) = n_i\pi_i$ and n_i is assumed known. In addition, the binomial distribution is a member of the exponential family (Dobson (2002, p. 46)).
(b) The linear predictor is $\eta = \boldsymbol{x}^T\boldsymbol{\beta}$.
(c) The link function for π is the logit and is the natural parameter, namely

$$\eta = \text{logit}(\pi) = \ln\frac{\pi}{1-\pi} = \boldsymbol{x}^T\boldsymbol{\beta}$$

(Dobson (2002, p. 46)). This equation implies that $\pi = \dfrac{e^\eta}{1+e^\eta}$. The information matrix follows from the fact that $Var(Y) = n\pi(1-\pi)$, $\mu = E(Y) = n\pi$ and for the logit link

$$\left(\frac{\partial \mu}{\partial \eta}\right) = \left(\frac{\partial \mu}{\partial \pi}\right)\left(\frac{\partial \pi}{\partial \eta}\right) = n\pi(1-\pi).$$

Thus, for a single observation $\boldsymbol{x}$, the information matrix (3.2) for $\boldsymbol{\beta}$ is

$$M(\boldsymbol{x};\boldsymbol{\beta}) = n\pi(1-\pi)\boldsymbol{x}\boldsymbol{x}^T$$

and for N observations, the information matrix (3.4) is

$$M(\boldsymbol{X};\boldsymbol{\beta}) = \boldsymbol{X}^T\boldsymbol{W}\boldsymbol{X} \tag{3.8}$$

where $\boldsymbol{W} = \text{diag}\{n_i\pi_i(1-\pi_i)\}$ and $\boldsymbol{X}$ has the ith row $\boldsymbol{x}_i^T = (1, x_{1i}, \ldots, x_{pi})$ with $i = 1, 2, \ldots, N$. The information matrix (3.8) depends on the parameters $\boldsymbol{\beta}$ through $\boldsymbol{W}$ which is a function of $\pi_i = \dfrac{1}{1+\exp\{-\boldsymbol{x}^T\boldsymbol{\beta}\}}$.

Example 3.3. *Normal nonlinear models*

Here
(a) $Y_i \sim N(\mu_i;\sigma^2)$ and the normal distribution is a member of the exponential family.
(b) The predictor expressed as $\eta = \eta(\boldsymbol{x};\boldsymbol{\beta})$ is not linear.
(c) The link function is the identity, i.e. $\mu_i = \eta_i$ with $i = 1, 2, \ldots, N$.
The information matrix (3.5) for $\boldsymbol{\beta}$ follows immediately from the fact that $Var(Y_i) = \sigma^2$ and $\left(\dfrac{\partial\mu_i}{\partial\eta_i}\right) = 1$ and thus for a single observation at $\boldsymbol{x}$, the information matrix (3.5) can be written as

$$M(\boldsymbol{x};\boldsymbol{\beta}) = \frac{1}{\sigma^2}\left(\frac{\partial\eta}{\partial\boldsymbol{\beta}}\right)\left(\frac{\partial\eta}{\partial\boldsymbol{\beta}}\right)^T. \tag{3.9}$$

Thus for N observations at $\boldsymbol{x}_i$, $i = 1, 2, \ldots, N$, the information matrix (3.6) for $\boldsymbol{\beta}$ is given by

$$M(\boldsymbol{X};\boldsymbol{\beta}) = \frac{1}{\sigma^2}\boldsymbol{F}^T\boldsymbol{F} \tag{3.10}$$

where $\boldsymbol{F}$ has the ith row $\left(\dfrac{\partial\eta_i}{\partial\beta_0}, \dfrac{\partial\eta_i}{\partial\beta_1}, \ldots, \dfrac{\partial\eta_i}{\partial\beta_p}\right)$. The information matrix (3.10) depends on $\boldsymbol{\beta}$ through the elements of $\boldsymbol{F}$.

Example 3.4. *Nonlinear binary logistic model*

Here
(a) $Y_i \sim \text{Binomial}(n_i;\pi_i)$ and the binomial distribution is a member of the exponential family.
(b) The predictor can be expressed as $\eta = \eta(\boldsymbol{x},\boldsymbol{\beta})$ and is not linear.
(c) The link for π is the logit and is the natural parameter, namely

$$\text{logit}(\pi) = \ln\frac{\pi}{1-\pi} = \eta(\boldsymbol{x};\boldsymbol{\beta}).$$

The information matrix follows from the fact that $Var(Y) = n\pi(1-\pi)$, $\mu = n\pi$ and for the logit link $\left(\dfrac{\partial\mu}{\partial\eta}\right) = \left(\dfrac{\partial\mu}{\partial\pi}\right)\left(\dfrac{\partial\pi}{\partial\eta}\right) = n\pi(1-\pi)$. For a single observation $\boldsymbol{x}$, the information matrix is given by

$$M(\boldsymbol{x};\boldsymbol{\beta}) = n\pi(1-\pi)\left(\frac{\partial\eta}{\partial\boldsymbol{\beta}}\right)\left(\frac{\partial\eta}{\partial\boldsymbol{\beta}}\right)^T. \tag{3.11}$$

Thus for N observations at $\boldsymbol{x}_i$ with $i = 1, 2, \dots, N$ the information matrix for $\boldsymbol{\beta}$ is

$$M(\boldsymbol{x};\boldsymbol{\beta}) = \boldsymbol{F}^T\boldsymbol{W}\boldsymbol{F} \tag{3.12}$$

where $\boldsymbol{W} = \text{diag}\{n_i\pi_i(1-\pi_i)\}$ and $\boldsymbol{F}$ has the ith row $\left(\frac{\partial\eta_i}{\partial\beta_0}, \frac{\partial\eta_i}{\partial\beta_1}, \dots, \frac{\partial\eta_i}{\partial\beta_p}\right)$. The information matrix (3.12) depends on $\boldsymbol{\beta}$ through $\boldsymbol{W}$ and the elements of $\boldsymbol{F}$.

3.5. Aim, design and optimality criteria

3.5.1 Aim

The aim of optimal design is to choose a set of values of the explanatory variables in a design space, say $\mathfrak{X}$, in order to maximize the information about a set of parameters so that these parameters are estimated as precisely as possible. Hence, the precise estimation of model parameters will lead to reliable inferences and predictions.

3.5.2 Design

A mapping that allocates inputs to suitable points in $\mathfrak{X}$ is called a design measure or a design, in short. The points of $\mathfrak{X}$ at which the inputs are allocated are called support points. Designs can be classified as exact or approximate.

Exact design

Consider an experiment in which N observations are made on the response variable. An exact design measure ξ_N is a mapping which allocates n_i trials to each of k distinct points $\boldsymbol{x}_i$ of the design space $\mathfrak{X}$, where n_i is an integer for $i = 1, 2, \dots, k$ and $\sum_{i=1}^{k} n_i = N$. Such a design is symbolically represented as

$$\xi_N = \left\{ \begin{array}{cccc} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \dots & \boldsymbol{x}_k \\ n_1 & n_2 & \dots & n_k \end{array} \right\}. \tag{3.13}$$

The quantities $\lambda_i = \frac{n_i}{N}$ corresponding to the support points $\boldsymbol{x}_i$ for $i = 1, 2, \ldots, k$ are called design weights. The weights satisfy the condition

$$0 \leq \lambda_i \leq 1 \quad \text{and} \quad \sum_{i=1}^{k} \lambda_i = 1. \tag{3.14}$$

The design (3.13) is called an exact design since $N\lambda_i = n_i$ is an integer (Atkinson and Donev (1992, p. 94), or Atkinson et al. (2007, p. 120)). Recall from Section 3.3 that the information matrix for a generalized nonlinear model at N observations was given in (3.6) so which can be written as

$$M(\boldsymbol{X};\boldsymbol{\beta}) = \sum_{i=1}^{N} w_i \left(\frac{\partial \eta_i}{\partial \boldsymbol{\beta}}\right)\left(\frac{\partial \eta_i}{\partial \boldsymbol{\beta}}\right)^T = \boldsymbol{F}^T \boldsymbol{W} \boldsymbol{F} \tag{3.15}$$

where w_i is the scalar $\frac{1}{Var(Y_i)}\left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2$ and $\boldsymbol{W}$ is a $N \times N$ diagonal matrix with w_i as the (i,i)th element and $\boldsymbol{F}$ has the ith row $\left(\frac{\partial \eta_i}{\partial \beta_0}, \frac{\partial \eta_i}{\partial \beta_1}, \ldots, \frac{\partial \eta_i}{\partial \beta_p}\right)$. For a GLM, $\left(\frac{\partial \eta_i}{\partial \boldsymbol{\beta}}\right) = \boldsymbol{x}$ and $\boldsymbol{F} = \boldsymbol{X}$ and in the normal case $w_i = \frac{1}{\sigma^2}$. As a consequence of (3.15), the standardized or average information matrix corresponding to the exact design (3.13) is given by

$$\frac{1}{N} M(\xi_N;\boldsymbol{\beta}) = \sum_{i=1}^{k} \frac{n_i}{N} w_i \left(\frac{\partial \eta_i}{\partial \boldsymbol{\beta}}\right)\left(\frac{\partial \eta_i}{\partial \boldsymbol{\beta}}\right)^T = \boldsymbol{F}^T \boldsymbol{W} \boldsymbol{F} \tag{3.16}$$

where $\boldsymbol{W}$ is a $N \times N$ diagonal matrix with $\frac{n_i}{N\text{Var}(Y_i)}\left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2$ as the (i,i)th element and $\boldsymbol{F}$ has the ith row $\left(\frac{\partial \eta_i}{\partial \beta_0}, \frac{\partial \eta_i}{\partial \beta_1}, \ldots, \frac{\partial \eta_i}{\partial \beta_k}\right)$ (see Fedorov and Hackl (1997, p. 18)). In this thesis, the scalar $w_i = \frac{n_i}{N\text{Var}(Y_i)}\left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2$ will be called "model weight" in order to distinguish it with the design weight $\lambda_i = \frac{n_i}{N}$.

Approximate design

Suppose that in the design (3.13), the condition that n_i is an integer is relaxed to the ratio $\lambda_i = \frac{n_i}{N}$ with $0 \leq \lambda_i \leq 1$ for $i = 1, \ldots, k$. This introduces the idea of approximate design. Clearly, an approximate design ξ defines a measure $\xi(\boldsymbol{x})$ at each $\boldsymbol{x} \in \mathfrak{X}$. If the design measure ξ is discrete with k support points on $\mathfrak{X}$, then the quantities $\lambda_i = \xi(\boldsymbol{x}_i)$ corresponding to the support points $\boldsymbol{x}_i$ for $i = 1, 2, \ldots, k$ are called design weights. Hence, an approximate design is represented by

$$\xi = \left\{ \begin{array}{cccc} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \ldots & \boldsymbol{x}_k \\ \lambda_1 & \lambda_2 & \ldots & \lambda_k \end{array} \right\}, \tag{3.17}$$

with $0 \leq \lambda_i \leq 1$ and $\sum_{i=1}^{k} \lambda_i = 1$. Similar to (3.16), the standardized or average information matrix corresponding to an approximate design is given, in the discrete case, by

$$M(\xi;\boldsymbol{\beta}) = \sum_{i=1}^{k} \lambda_i w_i \left(\frac{\partial \eta_i}{\partial \boldsymbol{\beta}}\right)\left(\frac{\partial \eta_i}{\partial \boldsymbol{\beta}}\right)^T = \boldsymbol{F}^T\boldsymbol{W}\boldsymbol{F}, \tag{3.18}$$

where $\boldsymbol{W}$ is a $N \times N$ diagonal matrix with (i,i)th element $w_i = \dfrac{\lambda_i}{Var(Y_i)}\left(\dfrac{\partial \mu_i}{\partial \eta_i}\right)^2$ and $\boldsymbol{F}$ has the ith row $\left(\dfrac{\partial \eta_i}{\partial \beta_0}, \dfrac{\partial \eta_i}{\partial \beta_1}, \ldots, \dfrac{\partial \eta_i}{\partial \beta_k}\right)$.

If the design measure is continuous on $\mathfrak{X}$, then the design weights are defined by means of a probability density function which is a continuous function $\xi(\boldsymbol{x})$, $\boldsymbol{x} \in \mathfrak{X}$ such that $\xi(\boldsymbol{x}) \geq 0$ and $\int_{\mathfrak{X}} \xi(\boldsymbol{x})d\boldsymbol{x} = 1$.

For an approximate continuous design, the information matrix of the parameters is given by

$$M(\xi;\boldsymbol{\beta}) = \int_{\mathfrak{X}} M(\boldsymbol{x};\boldsymbol{\beta})\xi(\boldsymbol{x})d\boldsymbol{x} = \int_{\mathfrak{X}} \frac{\xi(\boldsymbol{x})}{\mathrm{var}(Y)}\left(\frac{\partial \mu}{\partial \eta}\right)^2\left(\frac{\partial \eta}{\partial \boldsymbol{\beta}}\right)\left(\frac{\partial \eta}{\partial \boldsymbol{\beta}}\right)^T d\boldsymbol{x}. \tag{3.19}$$

In practice continuous optimal designs are unusual and will not be treated in this thesis.

3.5.3 Optimality criteria

A design is said to be optimal if it maximizes the information on the parameters in such a way that these parameters are evaluated as precisely as possible. This information is summarized in the information matrix. Maximizing the information matrix of the vector of parameters $\boldsymbol{\beta}$ implies minimizing the asymptotic variance-covariance of $\boldsymbol{\beta}$. It is not possible to optimize a matrix, except in the one-parameter case. Therefore functions of the information matrix which are statistically meaningful are introduced. Some criteria, often called alphabetic criteria, have been formulated for this purpose, mainly for linear regression models (see for example Fedorov (1972), Silvey (1980), Atkinson and Donev (1992), and Atkinson et al. (2007)). These criteria form the basis for other more sophisticated criteria such as those for nonlinear models (see for example White (1973)).

The notion of convex or concave functions is important in defining some optimality criteria. Concave and convex functions are defined as follows. Consider a design space $\mathfrak{X}$. A function $\Psi(M(\xi;\boldsymbol{\beta}))$ is convex on $\mathfrak{X}$ if for any $\alpha \in [0,1]$ and for any two designs ξ_1 and ξ_2 defined on $\mathfrak{X}$ the inequality

$$\Psi\left\{\alpha M(\xi_1,\boldsymbol{\beta}) + (1-\alpha)M(\xi_2,\boldsymbol{\beta})\right\} \leq \alpha\Psi\{M(\xi_1,\boldsymbol{\beta})\} + (1-\alpha)\Psi\{M(\xi_2,\boldsymbol{\beta})\} \tag{3.20}$$

holds. Similarly, Ψ is concave on $\mathfrak{X}$, if for any $\alpha \in [0,1]$ and for any two designs ξ_1 and ξ_2 defined on $\mathfrak{X}$ the inequality

$$\Psi\left\{\alpha M(\xi_1,\boldsymbol{\beta}) + (1-\alpha)M(\xi_2,\boldsymbol{\beta})\right\} \geq \alpha\Psi\{M(\xi_1,\boldsymbol{\beta})\} + (1-\alpha)\Psi\{M(\xi_2,\boldsymbol{\beta})\} \tag{3.21}$$

holds (see for example Fedorov (1972, p. 70) and Silvey (1980, p. 17)). The optimality criteria are classified with respect to minimizing a convex design criterion $\Psi\{M(\xi;\boldsymbol{\beta})\}$ or to maximizing a concave design criterion $\Psi\{M(\xi;\boldsymbol{\beta})\}$. Optimality criteria are subdivided in three main categories, namely determinant based, linear based, and minimax based. The following definitions can be found in any of the monographs on optimal designs, such as Fedorov (1972), Silvey (1980), Pázman (1986), Atkinson and Donev (1992), and Atkinson et al. (2007) as well as in a vast number of articles on optimal designs.

3.5.4 Determinant based criteria

D-optimality

Definition 3.1. A design ξ^* is said to be D-optimal if and only if it maximizes $\ln|M(\xi;\boldsymbol{\beta})|$ or equivalently if and only if it minimizes $\ln|M^{-1}(\xi;\boldsymbol{\beta})| = -\ln|M(\xi;\boldsymbol{\beta})|$, where $|M(\xi;\boldsymbol{\beta})|$ is the determinant of information matrix $M(\xi;\boldsymbol{\beta})$ for the parameters of a given model and ξ is an element of the set Ξ of all possible designs on a design space $\mathfrak{X}$.

The letter D stands for determinant and the design criterion to be minimized is

$$\Psi_D\{M(\xi;\boldsymbol{\beta})\} = -\ln|M(\xi;\boldsymbol{\beta})|. \tag{3.22}$$

D-optimality criterion has the following important properties:

- The function $M(\xi;\boldsymbol{\beta}) \to -\ln|M(\xi;\boldsymbol{\beta})|$ is a strictly convex function since it is the reciprocal of $\ln|M(\xi;\boldsymbol{\beta})|$ which is a strictly concave function Fedorov (1972, p. 71, Lemma 2.2.2). Thus $\Psi_D\{M(\xi;\boldsymbol{\beta})\}$ corresponds to a global optimality criterion.

- If λ_1, λ_2,…,λ_p are the eigenvalues of $M(\xi;\boldsymbol{\beta})$, then the D criterion maximizes the determinant $|M(\xi;\boldsymbol{\beta})| = \prod_{i=1}^{p}\lambda_i$ or minimizes $|M^{-1}(\xi;\boldsymbol{\beta})| = \prod_{i=1}^{p}\frac{1}{\lambda_i}$ called the generalized variance of $\hat{\boldsymbol{\beta}}$ (see for example Silvey (1980, p. 10), Atkinson and Donev (1992, p. 42), and Atkinson et al. (2007, p. 53)). In fact in the generalized linear and nonlinear case, $|M^{-1}(\xi;\boldsymbol{\beta})|$ is asymptotically the variance-covariance matrix of the vector $\boldsymbol{\beta}$, that is $\widehat{\boldsymbol{\beta}}$ is approximately distributed as $N(\boldsymbol{\beta}, M^{-1}(\xi;\boldsymbol{\beta}))$. In the linear regression case with error term normally distributed with mean zero and variance σ^2, the distribution $\widehat{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}, M^{-1}(\xi;\boldsymbol{\beta}))$ is exact (Dobson (2002, p. 73)). The interpretation of this result is that the D-optimality criterion minimizes the volume of the confidence ellipsoid of the vector $\boldsymbol{\beta}$ of parameters given by

$$\left\{\boldsymbol{\beta} : (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})^T M(\xi;\boldsymbol{\beta})(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) \leq \text{ constant}\right\} \tag{3.23}$$

where $\hat{\boldsymbol{\beta}}$ is an estimate of $\boldsymbol{\beta}$ (Silvey (1980, p. 10)).

- The D criterion is invariant with respect to any linear transformation of the explanatory variables (Atkinson and Donev (1992, p. 107), and (Atkinson et al., 2007, p. 136)).

The above properties make the D-optimality criterion the most used optimality criterion in practical applications (Silvey (1980, p. 41)).

D_A and D_s-optimality criteria

Suppose that A is a $p \times s$ matrix with $s < p$ and interest is on estimating the s linear combinations $A^T\boldsymbol{\beta}$ of the parameters $\boldsymbol{\beta}$. Since the asymptotic variance-covariance matrix of $\hat{\boldsymbol{\beta}}$ is $M^{-1}(\xi;\boldsymbol{\beta})$, then the asymptotic variance-covariance matrix of $A^T\hat{\boldsymbol{\beta}}$ is $A^T M^{-1}(\xi;\boldsymbol{\beta})A$. As a consequence, a criterion for estimating $A^T\boldsymbol{\beta}$ similar to (3.22), is to minimize the following criterion

$$\Psi_{D_A}\{M(\xi;\boldsymbol{\beta})\} = \ln |A^T M^{-1}(\xi;\boldsymbol{\beta})A|. \tag{3.24}$$

This criterion was called the D_A-optimality by Sibson (1974) to emphasize its dependency on the matrix A (also see Atkinson and Donev (1992, p. 108), and Atkinson et al. (2007, p. 137)).

One special case of the D_A optimality criterion is that for a single linear combination $\boldsymbol{a}^T\boldsymbol{\beta}$. In this case the determinant in (3.24) is not needed because $\boldsymbol{a}^T M^{-1}(\xi;\boldsymbol{\beta})\boldsymbol{a}$ is a scalar. The criterion which minimizes the quadratic form $\boldsymbol{a}^T M^{-1}(\xi;\boldsymbol{\beta})\boldsymbol{a}$ is called c-optimality and will be discussed in Section 3.5.5.

Another special case of D_A optimality is that for a linear combination of a subset of the parameters of the parameter vector $\boldsymbol{\beta}$. In that regards, consider a $p \times 1$ parameter vector $\boldsymbol{\beta}$ and suppose that, without loss of generality, the interest of the experimenter is on estimating the vector $\boldsymbol{\beta}_1$ of the first s parameters $\beta_1, \beta_2, \ldots, \beta_s$ where $s < p$ and the remaining $p - s$ parameters $\beta_{s+1}, \beta_{s+2}, \ldots, \beta_p$ are taken to be nuisance parameters. This is a special case of

D_A-optimality where $A^T = (I_s \quad \mathbf{0})$, I_s being the $s \times s$ identity matrix and $\mathbf{0}$ the $s \times (p-s)$ null matrix. Here, the letter s stands for subset. The information matrix for the parameters can be conformably partitioned as

$$M(\xi;\boldsymbol{\beta}) = \begin{pmatrix} M_{11}(\xi;\boldsymbol{\beta}) & M_{12}(\xi;\boldsymbol{\beta}) \\ M_{12}^T(\xi;\boldsymbol{\beta}) & M_{22}(\xi;\boldsymbol{\beta}) \end{pmatrix}.$$

The variance-covariance matrix of the estimates of $\boldsymbol{\beta}_1$ is the upper $s \times s$ sub-matrix of $M^{-1}(\xi;\boldsymbol{\beta})$ given by

$$M^{11}(\xi;\boldsymbol{\beta}) = \{M_{11}(\xi;\boldsymbol{\beta}) - M_{12}(\xi;\boldsymbol{\beta})M_{22}^{-1}(\xi;\boldsymbol{\beta})M_{12}^T(\xi;\boldsymbol{\beta})\}^{-1}$$

and the D_s-optimal design maximizes the determinant

$$\ln|\{M^{11}(\xi;\boldsymbol{\beta})\}^{-1}| = \ln|M_{11}(\xi;\boldsymbol{\beta}) - M_{12}(\xi;\boldsymbol{\beta})M_{22}^{-1}(\xi;\boldsymbol{\beta})M_{12}^T(\xi;\boldsymbol{\beta})| = \ln\frac{|M(\xi;\boldsymbol{\beta})|}{|M_{22}(\xi;\boldsymbol{\beta})|} \quad (3.25)$$

(see Silvey (1980, p. 11), Atkinson and Donev (1992, p. 109) and Atkinson et al. (2007, p. 139)).

3.5.5 Linear optimality

Definition 3.2. Suppose that L is a $p \times q$ nonnegative definite matrix of real coefficients. A design ξ_L^* is termed an L-optimal design if it minimizes a criterion of the form

$$\Psi_L\{M(\xi;\boldsymbol{\beta})\} = tr\{M^{-1}(\xi;\boldsymbol{\beta})L\} \quad (3.26)$$

where $tr\{M^{-1}(\xi;\boldsymbol{\beta})L\}$ is the trace of the matrix $M^{-1}(\xi;\boldsymbol{\beta})L$, that is the sum of the (i,i)th elements of $M^{-1}(\xi;\boldsymbol{\beta})L$. The letter L stands for linear. The following two cases are possible.

1. Suppose that s is the rank of L with $s \leq q$. Then L can be written as $L = AA^T$ where A is a $p \times s$ matrix of rank s. In this case, the criterion (3.26) can be written as

   $$\Psi_{A_A}\{M(\xi;\boldsymbol{\beta})\} = \mathrm{tr}\{M^{-1}(\xi;\boldsymbol{\beta})L\} = tr\{M^{-1}(\xi;\boldsymbol{\beta})AA^T\} = tr\{A^TM^{-1}(\xi;\boldsymbol{\beta})A\} \quad (3.27)$$

   and this criterion is called A_A-optimality criterion (see Atkinson and Donev (1992, p. 114). Some special cases are the following

A-optimality

The A-optimality is obtained by taking $A = I$, where I is the $p \times p$ identity matrix, to give the following definition.

Definition 3.3. A design ξ_A^* is said to be A-optimal if and only if it minimizes the criterion $\Psi_A M(\xi;\boldsymbol{\beta}) = tr\left\{M^{-1}(\xi;\boldsymbol{\beta})\right\}$ where $\mathrm{tr}\left\{M^{-1}(\xi;\boldsymbol{\beta})\right\}$ is the trace of the inverse of the information matrix of the parameters.

Note that

$$\mathrm{tr}\{M^{-1}(\xi;\boldsymbol{\beta})\} = \sum_{i=1}^{p}\frac{1}{\lambda_i} = \sum_{i=1}^{p}\mathrm{var}(\hat{\boldsymbol{\beta}}_i) \tag{3.28}$$

where $\lambda_i, i = 1, 2, \ldots, p$, are the eigenvalues of the information matrix $M(\xi;\boldsymbol{\beta})$. Clearly, minimizing $\Psi_A M(\xi;\boldsymbol{\beta})$ is equivalent to minimizing

$$\Psi\{M(\xi;\boldsymbol{\beta})\} = \frac{1}{p}\mathrm{tr}\left\{M^{-1}(\xi;\boldsymbol{\beta})\right\} = \frac{1}{p}\sum_{i=1}^{p}\frac{1}{\lambda_i} = \frac{1}{p}\sum_{i=1}^{p}\mathrm{var}(\hat{\boldsymbol{\beta}}_i) \tag{3.29}$$

so that A stands for "average". The expression (3.29) shows that one disadvantage of the A-optimality criterion is that it minimizes the average of the variances of the parameters but does not take into account their covariances. Another disadvantage of this criterion is that it is not invariant to a linear transformation of the explanatory variables (Atkinson and Donev (1992, p. 107), and Atkinson et al. (2007, p. 136)).

c-optimality

If $s = 1$, then A in (3.27) is the column vector $\boldsymbol{c}$ and the c-optimality is obtained, which is defined as follows.

Definition 3.4. A design ξ_c^* is c-optimal if and only if it minimizes the variance of the estimate of the linear combination $\boldsymbol{c}^T\boldsymbol{\beta}$ of the vector of parameters $\boldsymbol{\beta}$, which means that ξ_c^* minimizes the criterion $\Psi_c\{M(\xi;\boldsymbol{\beta})\} = \boldsymbol{c}^T M^{-1}(\xi;\boldsymbol{\beta})\boldsymbol{c}$ over all designs $\xi \in \Xi$ where Ξ is the set of all possible designs on a given design space.

Note that c-optimal designs are interesting since in some cases, parameters may not be estimable but their linear combinations estimable or interest is on these linear combinations. A disadvantage of c-optimum designs is that in some situations the information matrix and therefore the design is singular. This renders calculations difficult or requires the use of generalized inverses of the information matrices (Silvey (1978), Silvey (1980,

p. 13), Atkinson and Donev (1992, p. 113), and Atkinson et al. (2007, p. 142)). Two important special cases of c-optimality are the following:

(a) *Precision on an individual parameter*
If interest is on the parameter β_i, $i = 1, 2, \ldots, p$, then c^T=$(0, 0, \ldots, 1, 0, \ldots, 0)$ with 1 at the *ith* position and 0 elsewhere.

(b) *Two parameter contrasts*
If interest is now on the precision of some contrasts $\beta_i - \beta_j$, $i \neq j$ and $i, j = 1, 2, \ldots, p$, then c^T=$(0, 0, \ldots, 1, \ldots, 0, 0, \ldots, -1, 0, \ldots, 0)$ with $c_i = 1$, $c_j = -1$ and $c_l = 0$ for $i \neq j \neq l$, and $i, j, l = 1, 2, \ldots, p$.

An interesting feature of c-optimality is that it is invariant to linear transformation of the explanatory variables as for D-optimality.

2. Suppose that interest is on several contrasts, say s contrasts $C^T\boldsymbol{\beta}$, where C is a $p \times s$ matrix of coefficients with elements in each row summing to 0. The criteria to be minimized given by

$$\Psi_C\{M(\xi; \boldsymbol{\beta})\} = tr\{C^T M^{-1}(\xi; \boldsymbol{\beta})C\} \tag{3.30}$$

is called C-optimality (Atkinson and Donev (1992, p. 114), and Atkinson et al. (2007, p. 143)).

3.5.6 Minimax criteria: G-and E-optimality

Minimax criteria are defined in terms of the variance of the predicted response. For a linear model with normal errors, the variance of the predicted response $\boldsymbol{x}^T\hat{\boldsymbol{\beta}}$ is given by $d(\boldsymbol{x}; \xi) = \boldsymbol{x}^T M^{-1}(\xi)\boldsymbol{x}$ and this result is exact. However, for a nonlinear model the approximate asymptotic variance of the predicted response $\eta = \eta(\boldsymbol{x}; \hat{\boldsymbol{\beta}})$ is

$$d(\boldsymbol{x}, \xi; \boldsymbol{\beta}) = \left(\frac{\partial \eta}{\partial \boldsymbol{\beta}}\right)^T M^{-1}(\xi; \widehat{\boldsymbol{\beta}}) \left(\frac{\partial \eta}{\partial \boldsymbol{\beta}}\right) \tag{3.31}$$

(see for examples O'Brien and Funk (2003)).

G-optimality

Definition 3.5. A design ξ^*_G is G-optimal if it minimizes the maximum standardized variance $d(\boldsymbol{x}, \xi; \boldsymbol{\beta})$ of the predicted response over a design space $\mathfrak{X}$. In other words, the criterion

$\Psi_G\{M(\xi;\boldsymbol{\beta})\}$ to be minimized by the choice of design ξ_G^* in the set Ξ of all designs on the design space $\mathfrak{X}$ is given by

$$\Psi_G\{M(\xi;\boldsymbol{\beta})\} = \min_{\xi\in\Xi}\max_{\boldsymbol{x}\in\mathfrak{X}}\{d(\boldsymbol{x},\xi;\boldsymbol{\beta})\} \tag{3.32}$$

where $d(\boldsymbol{x},\xi;\boldsymbol{\beta})$ is defined by (3.31).

E-optimality

Consider looking at the variance of the linear combination $\boldsymbol{c}^T\hat{\boldsymbol{\beta}}$ where $\boldsymbol{c} \in \mathbb{R}^p$ such that $\boldsymbol{c}^T\boldsymbol{c} = 1$, that is the vector $\boldsymbol{c}$ lies on the unit sphere. The E-optimality criterion minimizes the maximum variance of all linear combinations $\boldsymbol{c}^T\hat{\boldsymbol{\beta}}$ for $\boldsymbol{c}$ lying in the unit sphere $\boldsymbol{c}^T\boldsymbol{c} = 1 \subset \mathbb{R}^p$, and thus the following definition.

Definition 3.6. A design ξ_E^* is E-optimal if and only if it minimizes the variance of the linear combination $\boldsymbol{c}^T\hat{\boldsymbol{\beta}}$ subject to $\boldsymbol{c}^T\boldsymbol{c} = 1$. In other words, the criterion Ψ to be minimized by the choice of design ξ_E^* in the set Ξ of all designs on the design space $\mathfrak{X}$ is given by

$$\Psi_E\{M(\xi;\boldsymbol{\beta})\} = \max_{\boldsymbol{c}^T\boldsymbol{c}=1}\left[\boldsymbol{c}^T\{M^{-1}(\xi;\boldsymbol{\beta})\}\boldsymbol{c}\right]. \tag{3.33}$$

Minimizing (3.33) is equivalent to minimizing the maximum eigenvalue of $M^{-1}(\xi;\boldsymbol{\beta})$ or equivalently maximizing the minimum eigenvalue of $M(\xi;\boldsymbol{\beta})$ by choice of designs ξ in Ξ (Silvey (1980, p. 12) and Atkinson and Donev (1992, p. 107)).

3.6. Construction of optimal designs for linear models

3.6.1 Preliminaries

Aim

The aim of this section is to review methods of optimization of the criteria given in Sections 3.5.4 and 3.5.5.

Scope

The information matrix (3.7) for linear models does not depend on the parameters. Thus, the theory flows smoothly and all extensions, as for example to generalized linear and nonlinear models, are based on these ideas.

3.6.2 Construction of exact optimal designs

Problem

The construction of exact optimal designs is difficult. Specifically the problem is rooted in combinatorics. As an example, the D-criterion has been shown to be non-deterministic polynomial-time hard (NP-hard), i.e. no fast solution is known (Welch (1982a)). In certain cases it is possible to enumerate all possible designs and thus find the optimum. However this approach rapidly becomes prohibitive in terms of computer-time as the size of the problem increases (Welch (1982b)). Thus, at least in general, heuristic algorithms are needed for constructing exact optimal designs and these do not guarantee to find the global optimum (Atkinson and Donev (1992, p. 178), and Atkinson et al. (2007, p. 181)).

Algorithms

There is a number of algorithms used in the search for exact designs on a given design space. Some of them are the following:
(a) Enumeration: This consists of giving a complete list of all possible designs so that the optimal one can be located. A more efficient way is to perform a branch-and-bound (Welch, 1982b). The method guarantees to find an optimal design, but computation becomes cumbersome for high sized designs (Atkinson and Donev (1992, p. 178), and Atkinson et al. (2007, p. 181)).
(b) Exchange algorithms: These algorithms are based on adding trials to a smaller candidate design or deleting trials from a larger candidate design or exchanging points from the K initial support points and from the list L of candidate support points. The aim is to obtain an improved N-trial design for a given criterion. Exchange algorithms include DETMAX (Mitchell (1974)) and the Fedorov algorithm (Fedorov (1972, Chap. 3)). The theory related to these algorithms are summarized and extended in Atkinson and Donev (1992, Chapter 15), and in Atkinson et al. (2007, Chapter 12).
(c) Other heuristic procedures: Examples include the Nelder-Mead simplex method (Nelder and Mead (1965), simulated annealing (Haines (1987)) and genetic algorithms (Joshi and Moudgalya (2004, p. 308)).

3.6.3 Construction of approximate optimal designs

Idea

An approximate design was defined as a design of form (3.17). The theory associated with approximate optimal design is powerful in the sense that it is based on a well documented theory and, in particular, leads to the construction of global optimal designs.

Caratheodory's Theorem

One of the important results in the theory of approximate designs is the Caratheodory's Theorem as it provides *a priori* the maximum number of support points of a D-optimal design. The theorem is formulated as follows.

Theorem 3.1. Consider a design space $\mathfrak{X} \subset \mathbb{R}^k$, and Ξ the set of all designs ξ on $\mathfrak{X}$ and a criterion Ψ defined on $\mathfrak{X}$. The Caratheodory's Theorem states there always exist an approximate design ξ^* with number of support points at most equal to $\frac{1}{2}k(k+1)+1$ that satisfies the criterion Ψ. In particular, if the criterion Ψ is concave then the number of support points of the optimal design ξ^* is at most equal to $\frac{1}{2}k(k+1)$ (Silvey (1980, p. 16)).

Special case

Suppose that the number, k, of support points is equal to the number, p, of parameters and consider the D-optimality criterion. The information matrix of the approximate design (3.17) is given by

$$M(\xi) = \sum_{i=1}^{p} \lambda_i M(\boldsymbol{x}_i) = \boldsymbol{X}^T \boldsymbol{W} \boldsymbol{X} \tag{3.34}$$

where the design matrix $\boldsymbol{X}$ with ith row $\boldsymbol{x}_i^T = (1, x_{i1}, x_{i2}, \ldots, x_{i,p-1})$ and the matrix of weights $W = \text{diag}\left(\frac{\lambda_1}{\sigma^2}, \frac{\lambda_2}{\sigma^2}, \ldots, \frac{\lambda_p}{\sigma^2}\right)$ have the same dimension $p \times p$. The determinant of the information matrix (3.34) factorizes to

$$|M(\xi)| = |\boldsymbol{X}^T \boldsymbol{W} \boldsymbol{X}| = |\boldsymbol{X}|^2 |\boldsymbol{W}| = |\boldsymbol{X}|^2 \prod_{i=1}^{p} \lambda_i. \tag{3.35}$$

Thus, the optimization problem separate to maximizing $|\boldsymbol{X}|$ with respect to the support points $\boldsymbol{x}_i$ and $|\boldsymbol{W}| = \prod_{i=1}^{p} \lambda_i$ with respect to the weights λ_i subject to $0 < \lambda_i < 1$ and $\sum_{i=1}^{p} \lambda_i = 1$. The latter maximization leads to $\lambda_i = \frac{1}{p}$, $i = 1, 2, \ldots, p$. Thus, yielding a p-point D-optimum design with weights $\frac{1}{p}$ at each support point (Fedorov (1972, pp. 84-85) and Silvey (1980, p. 42)).

3.7. The general equivalence theorem for approximate designs

The general equivalence theorem for approximate designs is defined using the concept of directional derivative that can, in the present context, be defined as follows.

Definition 3.7. Suppose that Ξ is the set of all designs ξ on the design space $\mathfrak{X}$ and consider a design $\xi_{\boldsymbol{x}} \in \Xi$ which puts weight one on $\boldsymbol{x} \in \mathfrak{X}$. The directional derivative of an optimality criterion $\Psi\{M(\xi)\}$ at ξ in the direction of $\xi_{\boldsymbol{x}}$ is defined as

$$\phi(\boldsymbol{x};\xi) = \lim_{\alpha \to 0} \frac{\Psi\{(1-\alpha)M(\xi) + \alpha M(\xi_{\boldsymbol{x}})\} - \Psi\{M(\xi)\}}{\alpha}. \tag{3.36}$$

The directional derivative (3.36) is a building block for the following theorem known as the general equivalence theorem for optimal designs.

Theorem 3.2. Suppose that Ψ is a convex criterion function on the set of information matrices on the design space $\mathfrak{X}$. The general equivalence theorem states that the following three assertions are equivalent.

1. The design ξ^* minimizes $\Psi\{M(\xi)\}$ for all $\xi \in \Xi$;

2. The directional derivative $\phi(\boldsymbol{x};\xi^*)$ is greater than or equal to zero for all $\boldsymbol{x} \in \mathfrak{X}$;

3. The directional derivative $\phi(\boldsymbol{x};\xi^*)$ attains its minimum at the support points of the design.

This theorem is due to Kiefer and Wolfowitz (1960) and is reported in Atkinson and Donev (1992, p. 96) and in Atkinson et al. (2007, p. 122). The general proof of the theorem is found in Whittle (1973).

3.7.1 Directional derivatives for some criteria

The emphasis of this section is looking at the general equivalence theorem for the determinant-based and the linear-based optimality criteria. For each criterion the discussion falls into two parts: (a) deriving the directional derivative and (b) stating the theorem.

D-optimality

Firstly, recall that the D-optimality criterion is $\Psi_D\{M(\xi)\} = \ln|M^{-1}(\xi)| = -\ln|M(\xi)|$. Secondly note that $\frac{\partial}{\partial\alpha}\{\ln|M|\} = \mathrm{tr}\left\{M^{-1}\left(\frac{\partial M}{\partial\alpha}\right)\right\}$ where M is a nonsingular square matrix and α is a scalar (Fedorov (1972, p. 21)). Therefore, the directional derivative (3.36) for D-optimality criterion is given by

$$\begin{aligned}
\phi(\boldsymbol{x};\xi) &= \lim_{\alpha\to 0}\frac{-\ln|(1-\alpha)M(\xi)+\alpha M(\xi_{\boldsymbol{x}})|+\ln|M(\xi)|}{\alpha} \\
&= -\lim_{\alpha\to 0}\mathrm{tr}\{[(1-\alpha)M(\xi)+\alpha M(\xi_x)]^{-1}(-M(\xi)+M(\xi_{\boldsymbol{x}}))\} \\
&= -\mathrm{tr}\{M^{-1}(\xi)[-M(\xi)+M(\xi_{\boldsymbol{x}})]\} \\
&= p-\mathrm{tr}[M^{-1}(\xi)M(\xi_x)] \\
&= p-\boldsymbol{x}^T M^{-1}(\xi)\boldsymbol{x}
\end{aligned}$$

since $M(\xi_{\boldsymbol{x}}) = \boldsymbol{x}\boldsymbol{x}^T$ and $d(\boldsymbol{x};\xi) = \mathrm{tr}[M^{-1}(\xi)\boldsymbol{x}\boldsymbol{x}^T] = \boldsymbol{x}^T M^{-1}(\xi)\boldsymbol{x}$ (see for example Graybill (1983, Theorem 9. 1. 20)). This result and the General Equivalence Theorem 3.2 imply that General Equivalence Theorem 3.2 for D-optimality for linear models is the following.

Theorem 3.3. The following three assertions are equivalent for a D-optimal design, ξ^*, in a design space $\mathfrak{X}$.

1. The design ξ^* minimizes $-\ln|M(\xi)|$ over all designs $\xi\in\Xi$;
2. $d(\boldsymbol{x};\xi^*)\le p$ for all $\boldsymbol{x}\in\mathfrak{X}$;
3. The support points of the D-optimal design ξ^* are real solutions of equation $d(\boldsymbol{x};\xi^*) = p$

where $d(\boldsymbol{x};\xi^*) = \boldsymbol{x}^T M^{-1}(\xi^*)\boldsymbol{x}$.

It immediately follows from this theorem that for approximate designs D-and G-optimality are equivalent. This equivalence is an important result since minimax designs are in general difficult to construct.

D_A and D_s-optimality

The D_A-optimality criterion is given by the equation (3.24). Thus, for a linear model the directional derivative of Ψ_{D_A} at ξ in the direction of ξ_x is calculated as follows.

$$\begin{aligned}
\phi(\boldsymbol{x};\xi) &= \lim_{\alpha\to 0}\frac{\ln|A^TM^{-1}\{(1-\alpha)\xi+\alpha\xi_{\boldsymbol{x}}\}A| - \ln|A^TM^{-1}(\xi)A|}{\alpha}\\
&= \lim_{\alpha\to 0}\frac{\partial}{\partial\alpha}\{\ln|A^TM^{-1}\{(1-\alpha)\xi+\alpha\xi_{\boldsymbol{x}}\}A|\}\\
&= \operatorname{tr}\left\{\{A^TM^{-1}(\xi)A\}^{-1}\{A^T[-M^{-1}(\xi)(-M(\xi)+M(\xi_{\boldsymbol{x}}))M^{-1}(\xi)]A\}\right\}\\
&= \operatorname{tr}I_s - \operatorname{tr}\left\{\{A^TM^{-1}(\xi)A\}^{-1}A^TM^{-1}(\xi)\boldsymbol{x}\boldsymbol{x}^TM^{-1}(\xi)A\right\}\\
&= s - d(\boldsymbol{x};\xi)
\end{aligned}$$

where $s = \operatorname{rank}(A)$ and $d(\boldsymbol{x};\xi) = \boldsymbol{x}^TM^{-1}(\xi)A\{A^TM^{-1}(\xi)A\}^{-1}A^TM^{-1}(\xi)\boldsymbol{x}$. Thus, the General Equivalence Theorem 3.2 for the D_A optimality can be formulated as follows.

Theorem 3.4. The following three assertions are equivalent for a D_A-optimal design.

1. The design ξ^* minimizes $\ln|A^TM^{-1}(\xi)A|$ for all $\xi \in \Xi$;
2. $d(\boldsymbol{x};\xi^*) \le s$ for all $\boldsymbol{x} \in \mathfrak{X}$;
3. The support points of the D_A optimal design are real solutions of $d(\boldsymbol{x};\xi^*) = s$

where $s = \operatorname{rank}(A)$ and $d(\boldsymbol{x};\xi^*) = \boldsymbol{x}^TM^{-1}(\xi^*)A\{A^TM^{-1}(\xi^*)A\}^{-1}A^TM^{-1}(\xi^*)\boldsymbol{x}$.

The relationship between D_A optimality and D_s optimality discussed in Section 3.5.4 implies that the directional derivative for the D_s-optimality criterion is given by

$$\phi(\boldsymbol{x};\xi) = s - \boldsymbol{x}^TM^{-1}(\xi)\boldsymbol{x} + \boldsymbol{x}_2^TM_{22}^{-1}(\xi)\boldsymbol{x}_2$$

where $\boldsymbol{x}^T$ is partitioned as $\boldsymbol{x}^T = (\boldsymbol{x}_1, \boldsymbol{x}_2)$ and $M(\xi)$ is conformably partitioned as in Section 3.5.4. Thus the equivalence theorem for D_s-optimality is given by the following assertions.

Theorem 3.5. The following three assertions are equivalent.

1. The design ξ^* minimizes $\ln\left|\dfrac{M(\xi)}{M_{22}(\xi)}\right|$ for $\xi \in \Xi$ where $M_{22}(\xi)$ defined in Subsection 3.5.4 is the $(2,2)$ submatrix in $M(\xi;\boldsymbol{\beta}) = \begin{pmatrix} M_{11}(\xi;\boldsymbol{\beta}) & M_{12}(\xi;\boldsymbol{\beta}) \\ M_{12}^T(\xi;\boldsymbol{\beta}) & M_{22}(\xi;\boldsymbol{\beta}) \end{pmatrix}$;
2. $d(\boldsymbol{x},\xi^*) \le s$ for all $\boldsymbol{x} \in \mathfrak{X}$;

3. The support points of the D_s optimal design are solutions of the equation $d(\boldsymbol{x}, \xi^*) = s$

where s is the number of parameters of interest and $d(\boldsymbol{x}, \xi^*) = \boldsymbol{x}^T M^{-1}(\xi^*)\boldsymbol{x} - \boldsymbol{x}_2^T M_{22}^{-1}(\xi^*)\boldsymbol{x}_2$.

Linear criteria

Consider the linear optimality criterion defined in (3.26). The directional derivative is given by

$$\begin{aligned}\phi(\boldsymbol{x};\xi) &= \lim_{\alpha\to 0} \frac{\mathrm{tr}M^{-1}\{(1-\alpha)\xi + \alpha\xi_x\}L - \mathrm{tr}M^{-1}(\xi)L}{\alpha} \\ &= \lim_{\alpha\to 0} \mathrm{tr}\frac{\partial}{\partial\alpha}\{(1-\alpha)M^{-1}(\xi) + \alpha M^{-1}(\xi_x)\}L \\ &= \mathrm{tr}\{M^{-1}(\xi)L\} - \boldsymbol{x}^T M^{-1}(\xi)LM^{-1}(\xi)\boldsymbol{x}. \qquad (3.37)\end{aligned}$$

Hence the General Equivalence Theorem 3.2 for L-optimality is given by the following assertions.

Theorem 3.6. The following three assertions are equivalent.

1. The design ξ^* minimizes $\mathrm{tr}\{M^{-1}(\xi)L\}$ for all $\xi \in \Xi$;

2. $d(\boldsymbol{x};\xi^*) \le \mathrm{tr}\{M^{-1}(\xi^*)L\}$ for all $\boldsymbol{x} \in \mathfrak{X}$;

3. The support points of the L-optimal design are real solutions of the equation $d(\boldsymbol{x};\xi^*) = \mathrm{tr}\{M^{-1}(\xi^*)L\}$

where $d(\boldsymbol{x};\xi^*) = \boldsymbol{x}^T M^{-1}(\xi^*)LM^{-1}(\xi^*)\boldsymbol{x}$.

Special cases

Two important special cases are the following:
(1) *A-optimality*
The directional derivative for the A-optimality is a particular case of (3.37) where L is replaced by the identity matrix I_p, that is:

$$\phi(\boldsymbol{x};\xi) = \mathrm{tr}\{M^{-1}(\xi)\} - \boldsymbol{x}^T \left[M^{-1}(\xi)\right]^2 \boldsymbol{x}.$$

(2) *c-optimality*
If $L = \boldsymbol{c}\boldsymbol{c}^T$, where $\boldsymbol{c}$ is a $p \times 1$ column vector, the directional derivative (3.37) may be written as

$$\phi(\boldsymbol{x},\xi) = \boldsymbol{c}^T M^{-1}(\xi)\boldsymbol{c} - \{\boldsymbol{x}^T M^{-1}(\xi)\boldsymbol{x}\}^2.$$

3.7.2 Efficiency of a design

The efficiency of a design ξ is a real number in the interval $[0,1]$, generally expressed as a percentage, that gives the extent to which the design exhausts the information on the parameters (see Pukelsheim (1993, p. 132)). In other words, the higher the information obtained from the design ξ compared to other competing designs from the same set of designs Ξ, the higher the efficiency of the design ξ. Two measures of efficiency are popular. One is the D-efficiency defined by

$$D_{\text{eff}} = \left\{ \frac{|M(\xi;\boldsymbol{\theta})|}{|M(\xi^*;\boldsymbol{\theta})|} \right\}^{1/p}, \tag{3.38}$$

where ξ is an arbitrary design, ξ^* is the optimal design and p is the number of the parameters of the model. The ratio of the determinants in (3.38) is taken to the power $\frac{1}{p}$ so that the measure of efficiency is proportional to the design size for any dimension of the model (see Atkinson and Donev (1992, p. 116), and Atkinson et al. (2007, p. 152)).

Another measure of efficiency is that of the L-type and minimax-type optimal design given by

$$\Psi_{\text{eff}} = \frac{\Psi(\xi^*;\boldsymbol{\theta})}{\Psi(\xi;\boldsymbol{\theta})} \tag{3.39}$$

where $\Psi(\xi;\boldsymbol{\theta})$ is the optimality criterion of interest. For example, the G-efficiency, is given by

$$G_{\text{eff}} = \frac{\bar{d}(\xi^*)}{\bar{d}(\xi)} = \frac{p}{\bar{d}(\xi)}, \tag{3.40}$$

where $\bar{d}(\xi) = \max_{\boldsymbol{x} \in \mathcal{X}} d(\boldsymbol{x};\xi)$ (see Atkinson and Donev (1992, p. 116), and Atkinson et al. (2007, p. 152)).

3.8. Locally optimal design construction for nonlinear models

3.8.1 Preliminaries

Problem

For generalized linear models (GLM) (but not the normal case) and generalized nonlinear models (GnLM), the information matrix (3.18) or (3.19) is a function of the unknown parameters and this issue needs to be addressed.

Scope

The theory of optimal designs for linear regression models discussed in Section 3.6 forms the basis for handling the design problem for nonlinear models for which GLMs and GnLMs are special cases. One of the approaches usually adopted is that of “locally optimal designs” which is briefly discussed below.

3.8.2 Locally optimal designs

Idea and Equivalence Theorem

One way of addressing the problem of design dependency of the information matrix to the model parameters is *locally optimal design* introduced by Chernoff (1953). The principle is taking a best guess $\boldsymbol{\beta}^0$ of the vector of parameters $\boldsymbol{\beta}$. In this case the information matrices (3.18) and (3.19) for an approximate design do not depend on the parameters of the model, and hence the optimality criteria discussed for the construction of optimal design for linear models in Section 3.6 are applicable for locally optimal designs. The presence of $\boldsymbol{\beta}$ in parentheses following the criterion's notation is used to indicate the dependence on the best guess of $\boldsymbol{\beta}$ for locally optimal design. For example, $D(\boldsymbol{\beta})$- and $c(\boldsymbol{\beta})$- will stand for locally D-optimal and locally c-optimum designs respectively.

Some models can be linear, but interest can be in a nonlinear function of the vector of parameters. Suppose for example that g is a nonlinear function of the vector of parameters $\boldsymbol{\beta}$ with maximum likelihood estimator $g(\hat{\boldsymbol{\beta}})$. The first order Taylor expansion about the true parameter value $\boldsymbol{\beta}$ gives

$$g(\widehat{\boldsymbol{\beta}}) = g(\boldsymbol{\beta}) + \left[\frac{\partial g(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right]^T (\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}).$$

Then,

$$\begin{aligned} \mathrm{Var}[g(\widehat{\boldsymbol{\beta}})] &= \left[\frac{\partial g(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right]^T \mathrm{Var}(\widehat{\boldsymbol{\beta}}) \left[\frac{\partial g(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right] \\ &= \boldsymbol{c}^T(\boldsymbol{\beta}) M^{-1}(\widehat{\boldsymbol{\beta}}) \boldsymbol{c}(\boldsymbol{\beta}), \end{aligned} \tag{3.41}$$

where $\boldsymbol{c}(\boldsymbol{\beta}) = \dfrac{\partial g(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}$ and $M^{-1}(\widehat{\boldsymbol{\beta}})$=var$(\widehat{\boldsymbol{\beta}})$. With a best guess of $\boldsymbol{\beta}$, the optimum design is obtained by referring to c-optimality discussed in Section 3.5.5.

A pioneering work on the calculation of optimal designs for nonlinear models is that of Box and Lucas (1959). The one variable compartmental model with two parameters was used to illustrate the theory. The methodology used was that of an induced design introduced by Elfving (1952), but this approach is complicated and can be unpracticable for a model with more than one variable.

The General Equivalence Theorem 3.2 introduced by Kiefer and Wolfowitz (1960) for the construction of optimal designs for linear models was extended to the case of the constructing optimal designs for nonlinear models by White (1973). One of the elegant results established by White (1973) is the equivalence between $D(\boldsymbol{\beta})$-optimality and $G(\boldsymbol{\beta})$-optimality as summarized in the following theorem.

Theorem 3.7. The following three assertions are equivalent.

1. The design ξ^* is $D(\boldsymbol{\beta})$-optimal;
2. The design ξ^* is $G(\boldsymbol{\beta})$-optimal;
3. $\sup_{\boldsymbol{x}\in\mathcal{X}} d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) = p$ where p is the number of parameters in the model and $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) = \mathrm{tr}\{M(\boldsymbol{x}; \boldsymbol{\beta})M^{-1}(\xi^*, \boldsymbol{\beta})\}$ is the variance of the predicted response similar to (3.31).

The importance of Theorem 3.7 is in handling minimax criteria, such as $G(\boldsymbol{\beta})$-optimality, which are not based on differentiation. In fact, the above Equivalence Theorem permits the evaluation of optimality by means of determinant based criteria, such as $D(\boldsymbol{\beta})$-optimality, which are differentiable criteria. This renders the optimization relatively easy to perform.

In particular, the D-optimality based on Theorem 3.3 can be extended to GLM and GnLM as follows.

Theorem 3.8. For a generalized linear or nonlinear model, the following three assertions are equivalent for a D-optimal design on a design space $\mathfrak{X}$.

1. The design ξ^* minimizes $-\ln|M(\xi; \boldsymbol{\beta})|$ for all $\xi \in \Xi$;
2. $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) \leq p$ for all $\boldsymbol{x} \in \mathfrak{X}$;
3. The support points of the D-optimal design ξ^* are real solutions of the equation $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) = \mathrm{tr}\{M(\boldsymbol{x}; \boldsymbol{\beta})M^{-1}(\xi^*, \boldsymbol{\beta})\} = p$ where p is the number of model parameters.

3.8.3 Algorithms for constructing optimal designs

As stated previously, constructing an exact optimal design is a difficult combinatorial problem. In most situations the construction of optimal designs is for approximate designs discussed in

Section 3.6.3. For a generalized linear or nonlinear model, constrained nonlinear optimization routines can be used in the search of candidate D-optimal designs. The constraints rely on the weights and the design region $\mathfrak{X}$ with

$$w_i > 0, \quad \sum w_i = 1, \quad \text{and} \quad \boldsymbol{x} \in \mathfrak{X}. \tag{3.42}$$

First suppose that the design region $\mathfrak{X}$ is an ordered subset of $\mathbb{R}$ and that the scalar $x_{\min}$ and $x_{\max}$ are respectively the minimum and the maximum values of a scalar x on $\mathfrak{X}$. That is for any $x \in \mathfrak{X}$, $x_{\min} \leq x \leq x_{\max}$. It immediate follows that

$$0 \leq \frac{x - x_{min}}{x_{max} - x_{min}} \leq 1.$$

This expression is satisfied by the one to one transformation

$$\frac{x - x_{min}}{x_{max} - x_{min}} = \sin^2 y,$$

or, equivalently

$$x = x_{min} + (x_{max} - x_{min}) \sin^2 y, \tag{3.43}$$

and

$$y = \arcsin \sqrt{\frac{x - x_{min}}{x_{max} - x_{min}}}, \tag{3.44}$$

where $y \in \left[0, \frac{\pi}{2}\right]$. The same process on support points x is repeated on the weights w. If, for example, the D-optimality criterion is used for a candidate D-optimal design with k distinct support points and associated k weights, where some of the weights can be zeros, the criterion $\Psi_D(\xi; \boldsymbol{\beta}) = \ln |M^{-1}(\xi; \boldsymbol{\beta})|$ is transformed into the criterion $f(y_1, y_2, \ldots, y_{2k})$, i.e. a function of k points and k weights in a long $(2k \times 1)$ vector y_1, y_2, $\ldots$, y_{2k} with y_i constrained to be on $\left[0, \frac{\pi}{2}\right]$ for $i = 1, 2, \ldots, 2k$. Therefore the problem of calculating the values of x_1, x_2, $\ldots$, x_k of $\mathfrak{X}$ and $w_i \geq 0$ with $\sum_{i=1}^{k} w_i = 1$ which minimize the criterion $\Psi_D(\xi, \boldsymbol{\beta}) = \ln |M^{-1}(\xi; \boldsymbol{\beta})|$ is transferred to the problem of calculating the values of y_1, y_2, $\ldots$, y_{2k} which minimize $f(y_1, y_2, \ldots, y_{2k})$ where y_1, y_2, $\ldots$, y_{2k} lie in the interval $[0, \frac{\pi}{2}]$. By reasonable starting values y_1^0, y_2^0, $\ldots$, y_{2k}^0, a numerical optimization routine, such as Newton-Raphson, is used for finding x_1^r, x_2^r, $\ldots$, x_k^r and w_1^r, w_2^r, $\ldots$, w_k^r via y_1^r , y_2^r, $\ldots$, y_{2k}^r at iteration r. Evaluating $\ln |M^{-1}(\xi; \boldsymbol{\beta})|$ at every step until convergence to final values y_1^f, y_2^f, $\ldots$, y_{2k}^f provides final values x_1^f, x_2^f, $\ldots$, x_k^f and w_1^f, w_2^f, $\ldots$, w_k^f which minimize $\ln |M^{-1}(\xi; \boldsymbol{\beta})|$.

The above algorithm for one explanatory variable, x, is easily extended for $\mathfrak{X} \subset \mathbb{R}^k$ where k is the number of explanatory variables in the generalized linear or nonlinear model. For each variable, say x_j for $j = 1, \ldots k$, the transformation (3.44) may be used for finding the value of y_j. Then, the back transformation (3.43) may used for finding the actual value of x_j. This method of constructing D-optimal designs will be used in numerical examples of this thesis with Gauss code adjusted from procedures written by Haines in unpublished research work.

Another algorithm for constructing optimal designs given in Atkinson and Donev (1992, p. 103) and in Atkinson et al. (2007, p. 129) can be described as follows.

Atkinson and Donev (1992), and Atkinson et al. (2007) consider a design with k support points on a cubic design space of the form $-1 \leq x_i \leq 1$ in the euclidian space $\mathbb{R}^k$ where $i = 1, 2, \ldots, k$) and k is an integer greater or equal to 1. Then, they take

$$x_i = \sin z_i \tag{3.45}$$

as a transformation from unconstrained z_i to constrained x_i, for $i = 1, 2, \ldots, k$. Atkinson and Donev (1992), and Atkinson et al. (2007) state that optimal designs can first be constructed on the unconstrained design space of the z_i, then the z_i are back transformed to the constrained space of the x_i provided that subspaces used are such that (3.45) is a one-to-one transformation). Furthermore, Atkinson and Donev (1992), and Atkinson et al. (2007) specify the weights at the k support points of the candidate optimal design by the following transformation

$$\begin{aligned}
w_1 &= \sin^2 z_1 \\
w_2 &= \sin^2 z_2 \cos^2 z_1 \\
&\vdots \\
w_i &= \sin^2 z_i \prod_{j=1}^{i-1} \cos^2 z_j \qquad (i = 2, 3, \ldots, k-1) \\
&\vdots \\
w_k &= \prod_{j=1}^{k-1} \cos^2 z_j.
\end{aligned} \tag{3.46}$$

For instance, if $k = 2$ that is a case of a design with two support points, then the transformation (3.46) is written as

$$w_1 = \sin^2 z_1 \text{ and } w_2 = \cos^2 z_2$$

so that

$$w_1 \geq 0, w_2 \geq 0 \text{ and } w_1 + w_2 = 1.$$

Atkinson et al. (2007, p. 130-131) use SAS IML to write a code for the numerical construction of D-optimal designs on the design space $[-1, 1]$.

3.9. Conclusions

This chapter has provided a brief review on optimal designs. The emphasis was on the definition of a design, the optimality criteria, the specific equivalence theorem for each optimality criterion, and on algorithms for calculating locally optimal designs.

It was emphasized that optimal designs are model based. In particular, it was indicated that generalized linear and nonlinear models constitute a large class of useful models in practical situations such as in biomedical studies. The general forms of these models were summarized and corresponding expressions of the information matrices were calculated. Some examples were given, namely (1) the normal linear model, (2) the linear binary logistic model, (3) the normal nonlinear model, and (4) the nonlinear binary logistic model. The expressions of the information matrix for the different models were given, as these will be used in subsequent chapters.

The designs were defined and classified as exact or approximate. Emphasis was on approximate designs since they are easy to calculate. A review of the alphabetic design criteria was given. These were broadly classified in three categories, namely (1) determinant based criteria, (2) linear based criteria and (3) minimax based criteria. The General Equivalence Theorem of Kiefer and Wolfowitz (1960) for constructing optimal designs for linear models was reviewed, then adjusted to each optimality criterion discussed in the text. Furthermore, a review on how the theory of optimal design for linear models can be extended to nonlinear models using the idea of locally optimal design introduced by Chernoff (1953) was done, and the referring Equivalence Theorem was introduced by White (1973). Finally, two algorithms for constructing locally D-optimal design were provided, one related to Haines' unpublished work and another on Atkinson et al. (2007)'s approach in numerically constructing optimal designs using respectively Gauss and SAS.

4

Review of D-optimal Designs for the Two-Variable Binary Logistic Model

4.1. Introduction

This chapter reviews some existing work on constructing D-optimal designs for the precise estimation of the parameters of the two-variable binary logistic model that adequately describes the effects of two drugs administered jointly. In Section 4.2, some theory regarding the general logistic model and associated designs is presented. In Section 4.3, the construction of the D-optimal design for a single explanatory variable is reviewed and a proof for D-optimality of the design is given. The two-variable binary logistic models without and with interaction are introduced in Section 4.4. A brief account of the existing work on the construction of D-optimal designs for the two-variable binary logistic model without interaction is given in Section 4.5. Specifically, the approaches reviewed are those of Sitter and Torsney (1995a), Atkinson and Haines (1996), and Jia and Myers (2001). Finally, in Section 4.6, a brief review of work reported on the construction of the D-optimal designs for the two-variable binary logistic model with interaction is presented. Specifically, the approaches reviewed are those of Kupchak (2000), and Jia and Myers (2001).

4.2. Binary logistic model and D-optimal designs

4.2.1 Model and information matrix

Consider a binary response variable Y_i at the i^{th} setting of explanatory variables represented by the vector $\boldsymbol{x}_i = (x_{i1}, x_{i2}, \ldots, x_{ik})^T$ for $i = 1, 2, \ldots, d \geq k + 1$. The two possible values y_i at $\boldsymbol{x}_i$ can be coded as 1 for success or positive response and 0 for a failure or a negative response. If the probability of success, say p_i, at $\boldsymbol{x}_i$ is described by the binary logistic model, then

$$p_i = E[Y_i|\boldsymbol{x}_i] = \frac{1}{1 + \exp\{-\widetilde{\boldsymbol{x}}_i^T \boldsymbol{\beta}\}} \tag{4.1}$$

or equivalently

$$u_i = \text{logit}(p_i) = \ln \frac{p_i}{1 - p_i} = \widetilde{\boldsymbol{x}}_i^T \boldsymbol{\beta} \tag{4.2}$$

where $\widetilde{\boldsymbol{x}}_i = \left[1, \boldsymbol{x}_i^T\right]^T$ and $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \ldots, \beta_k)^T$ is the vector of model parameters (see Dobson (2002, p. 115)). As in Example 3.2, the information matrix for $\boldsymbol{\beta}$ evaluated at $\boldsymbol{x}_i$ is

$$M(\boldsymbol{x}_i; \boldsymbol{\beta}) = p_i(1 - p_i)\widetilde{\boldsymbol{x}}_i \widetilde{\boldsymbol{x}}_i^T = p_i(1 - p_i) \begin{bmatrix} 1 \\ \boldsymbol{x}_i \end{bmatrix} \left[1, \boldsymbol{x}_i^T\right] \tag{4.3}$$

for $i = 1, 2, \ldots, d \geq k + 1$.

4.2.2 Design

An exact design with d distinct support points $\boldsymbol{x}_i$ is given by

$$\xi_N = \left\{ \begin{array}{cccc} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \ldots & \boldsymbol{x}_d \\ n_1 & n_2 & \ldots & n_d \end{array} \right\}$$

where n_i is the number of replications of $\boldsymbol{x}_i$ for $i = 1, 2, \ldots, d \geq k + 1$, and $N = \sum_{i=1}^{d} n_i$ is the total sample size. The exact design ξ_E on a per point basis is given by

$$\xi_E = \left\{ \begin{array}{cccc} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \ldots & \boldsymbol{x}_d \\ \frac{n_1}{N} & \frac{n_2}{N} & \ldots & \frac{n_d}{N} \end{array} \right\} \tag{4.4}$$

where $\sum_{i=1}^{d} \frac{n_i}{N} = 1$. The approximate counterpart of design (4.4) is defined as

$$\xi = \left\{ \begin{array}{cccc} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \ldots & \boldsymbol{x}_d \\ \lambda_1 & \lambda_2 & \ldots & \lambda_d \end{array} \right\} \tag{4.5}$$

where the design weights λ_i satisfy the conditions $0 < \lambda_i < 1$ and $\sum_{i=1}^{d} \lambda_i = 1$. Now, using the matrix (4.3), the information matrix for $\boldsymbol{\beta}$ evaluated at the approximate design (4.5) is given by

$$M(\xi;\boldsymbol{\beta}) = \sum_{i=1}^{d} \lambda_i M(\boldsymbol{x}_i;\boldsymbol{\beta}) = \sum_{i=1}^{d} \lambda_i p_i(1-p_i) \begin{bmatrix} 1 \\ \boldsymbol{x}_i \end{bmatrix} \begin{bmatrix} 1, \boldsymbol{x}_i^T \end{bmatrix} \tag{4.6}$$

or more compactly

$$M(\xi;\boldsymbol{\beta}) = \boldsymbol{X}^T \boldsymbol{W} \boldsymbol{X}$$

where $\boldsymbol{X}$ is a $d \times (k+1)$ matrix whose i^{th} row is $[1, \boldsymbol{x}_i^T]$ and $\boldsymbol{W}$ is a $d \times d$ diagonal matrix of weights whose $(i,i)^{th}$ element is

$$w_{ii} = \lambda_i p_i(1-p_i), \ \ i = 1, 2, \ldots, d \geq k+1. \tag{4.7}$$

Recall from Section 3.5.4 that if Ξ is the set of all designs on a design space $\mathfrak{X}$, then the D-optimal design is the design that maximizes the determinant or log-determinant or, equivalently, minimizes minus the log-determinant of the information matrix over all designs in Ξ. The construction of the exact D-optimal designs of the form (4.4) is a difficult combinatorial problem (see Atkinson and Donev (1992, Chapter 15), and Atkinson et al. (2007, Chapter 12)). The difficulty of finding exact D-optimal designs has been circumvented by approximating the design ξ_E in (4.4) with the design ξ in (4.5) which is much easier to construct (Atkinson et al. (2007, p. 119)).

The design problem may be further simplified by reducing the problem to an equivalent canonical form which is solvable independently of $\boldsymbol{\beta}$ (Ford et al. (1992), Sitter and Torsney (1995a), Sitter and Torsney (1995b), Atkinson and Haines (1996), and Torsney and Gunduz (2001)) as follows. Let B be a $(k+1) \times (k+1)$ nonsingular matrix of the form

$$B = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ \beta_0 & \beta_1 & \beta_2 & \ldots & \beta_k \\ b_{01} & b_{11} & b_{21} & \ldots & b_{k1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{0(k-1)} & b_{1(k-1)} & b_{2(k-1)} & \ldots & b_{k(k-1)} \end{bmatrix} \tag{4.8}$$

where b_{ij} ($i = 0, 1, \ldots, k$ and $j = 1, 2, \ldots, k-1$) are coefficients of $(k-1)$ linear constraints $u_j = b_{0j} + b_{1j}x_1 + b_{2j}x_2 + \ldots + b_{kj}x_k$. Then, let $B\widetilde{\boldsymbol{x}} = B \begin{bmatrix} 1 \\ \boldsymbol{x} \end{bmatrix} = \begin{bmatrix} 1 \\ \boldsymbol{u} \end{bmatrix} = \widetilde{\boldsymbol{u}}$ for all $\boldsymbol{x} \in \mathfrak{X}$, and let $\mathfrak{X}_u = \{\boldsymbol{u} : B\widetilde{\boldsymbol{x}} = \widetilde{\boldsymbol{u}}, \boldsymbol{x} \in \mathfrak{X}\}$ be the design space induced by the mapping $\boldsymbol{x}$ to $\boldsymbol{u}$. The mapping $\boldsymbol{x}$ to $\boldsymbol{u}$ applied to the support points of design (4.5) results in the induced design

$$\xi_u = \left\{ \begin{array}{cccc} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \dots & \boldsymbol{u}_d \\ \lambda_1 & \lambda_2 & \dots & \lambda_d \end{array} \right\}, \tag{4.9}$$

and hence $B\widetilde{\boldsymbol{x}}_i = B \begin{bmatrix} 1 \\ \boldsymbol{x}_i \end{bmatrix} = \begin{bmatrix} 1 \\ \boldsymbol{u}_i \end{bmatrix} = \widetilde{\boldsymbol{u}}_i$ for $i = 1, 2, \dots, d$. In terms of the design ξ_u, the information matrix (4.6) is

$$\begin{aligned} M(\xi; \boldsymbol{\beta}) &= \sum_{i=1}^{d} \lambda_i p_i (1 - p_i) \widetilde{\boldsymbol{x}}_i \widetilde{\boldsymbol{x}}_i^T \\ &= \sum_{i=1}^{d} \lambda_i p_i (1 - p_i) B^{-1} \widetilde{\boldsymbol{u}}_i \widetilde{\boldsymbol{u}}_i^T (B^{-1})^T \\ &= B^{-1} \left[\sum_{i=1}^{d} \lambda_i p_i (1 - p_i) \widetilde{\boldsymbol{u}}_i \widetilde{\boldsymbol{u}}_i^T \right] (B^{-1})^T \\ &= B^{-1} M(\xi_u) (B^{-1})^T. \end{aligned}$$

Consequently,

$$M(\xi_u) = B M(\xi; \beta) B^T \tag{4.10}$$

and $|M(\xi_u)| = |B|^2 |M(\xi; \boldsymbol{\beta})|$. This means that maximizing $|M(\xi_u)|$ by choice of ξ_u on $\mathfrak{X}_u$ is equivalent to maximizing $|M(\xi; \boldsymbol{\beta})|$ by choice of ξ on $\mathfrak{X}$. Hence the optimal design ξ may easily be found as the back transformation of the optimal design ξ_u using mapping $\boldsymbol{u}$ to $\boldsymbol{x}$.

4.3. D-optimal design for the one-variable binary logistic model

The one-variable binary logistic model is model (4.2) with $\boldsymbol{\beta} = (\beta_0, \beta_1)^T$ and $\widetilde{\boldsymbol{x}} = [1, x]^T$ where x is a scalar, i.e. $u = \beta_0 + \beta_1 x$.

The D-optimal design for the one-variable binary logistic model is a well known two-point design. The following are some reported studies of this design.

Abdelbasit and Plackett (1983) considered model $u = \text{logit}(p) = \beta_0 + \beta_1 x$ re-parameterized to $u = \text{logit}(p) = \beta_1(x - \mu)$ where $\mu = -\dfrac{\beta_0}{\beta_1}$ and $x \in \mathbb{R}$, and found that a 2-point D-optimal design has support points at probabilities of response $p = 0.824$ and $q = 1 - p = 0.176$, which respectively correspond to logits $u_1 = 1.5434$ and $u_2 = -1.5434$. Atkinson and Donev (1992, pp. 292-293), and Atkinson et al. (2007, pp. 399-400) considered model $u = \text{logit}(p) = \beta_0 + \beta_1 x$ where $\beta_0 = 0$, $\beta_1 = 1$ and the scalar x is an element of the unbounded design space $\mathfrak{X} = \mathbb{R}$. The authors found algebraically that the D-optimal design has two equally weighted support points ± 1.5434, but did not prove global optimality of the design. Silvey (1980, pp. 59-60) briefly reported the two-point design for the one-variable binary logistic model in a bounded design space, but no proof of D-optimality of the design was provided. A relatively more detailed geometrical/numerical construction of the two-point D-optimal designs on the design spaces $\mathfrak{X} = [a, b]$, $\mathfrak{X} = [a, \infty)$ and $\mathfrak{X} = (-\infty, a]$, where a and b are real numbers, was introduced by Ford et al. (1992). They discussed the construction of c- and D-optimal designs for nine one-variable nonlinear models that are functions of $u = \beta_0 + \beta_1 x$ among which is the one-variable binary logistic model. Ford et al. (1992) expressed the optimal designs in "canonical form", i.e. in terms of u, and presented a geometric argument based on the minimum ellipsoid enclosing the canonical design locus to prove D-optimality. A formal algebraic proof of D-optimality of the two-point D-optimal design for model $u = \text{logit}(p) = \beta_0 + \beta_1 x$ on the design spaces $\mathfrak{X} = [a, b]$, $\mathfrak{X} = [a, \infty)$ and $\mathfrak{X} = (-\infty, a]$ is given by Sebastiani and Settimi (1997). Sebastiani and Settimi (1997) report that White (1975) proved D-optimality of the D-optimal design in the case of the unbounded design space $\mathbb{R}$. However, this does not clearly appear to be the case in White's thesis (1975). Furthermore, the proof for the case of the design space $\mathbb{R}$ does not immediately follow from the proof of Sebastiani and Settimi (1997).

The aim of this thesis is to construct D-optimal designs for model (4.2) with two variables. Since the case of the one-variable model constitutes a building block for constructing such designs, a new approach of algebraical construction of the two-point D-optimal design for the one-variable binary logistic model in the unrestricted design space $\mathbb{R}$ is presented here. Theorem 4.1 summarizes the results.

Theorem 4.1. The D-optimal design for a one-variable binary logistic model on the design space $\mathbb{R}$ is an equally weighted two-point design with support points located on the logits $\pm u = 1.5434$.

Proof

The proof of this theorem has two components, namely (i) construction of the candidate D-optimal design and (ii) proof of D-optimality of the candidate D-optimal design.

Step (i): The one-variable binary logistic model is the binary logistic model (4.2) but with $\widetilde{\boldsymbol{x}} = [1, x]^T$, where x is a scalar and $\boldsymbol{\beta} = (\beta_0, \beta_1)^T$. Therefore, the logit for the model can be written as $u = \text{logit}(p) = \beta_0 + \beta_1 x$. The information matrix for $\boldsymbol{\beta}$ at the 2-point candidate

D-optimal design $\xi = \left\{ \begin{array}{cc} x_1 & x_2 \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\}$ is given by

$$M(\xi;\boldsymbol{\beta}) = \frac{1}{2}\sum_{i=1}^{2}\psi(u_i)\left[\begin{array}{c} 1 \\ x_i \end{array}\right]\left[\begin{array}{cc} 1 & x_i \end{array}\right] = \frac{1}{2}\sum_{i=1}^{2}\psi(u_i)\left[\begin{array}{cc} 1, & x_i \\ x_i & x_i^2 \end{array}\right] \tag{4.11}$$

where $\psi(u_i) = p_i(1-p_i) = \dfrac{e^{u_i}}{(1+e^{u_i})^2}$ and $u_i = \text{logit}(p_i) = \beta_0 + \beta_1 x_i$ for $i = 1, 2$. The transformation of this design problem to canonical form as in (4.8), (4.9) and (4.10) is accomplished with $B = \left[\begin{array}{cc} 1 & 0 \\ \beta_0 & \beta_1 \end{array}\right]$. The transformation results in

$$M(u) = BM(\boldsymbol{x};\boldsymbol{\beta})B = \psi(u)\left[\begin{array}{cc} 1 & u \\ u & u^2 \end{array}\right],$$

and hence

$$M(\xi_u) = BM(\xi;\beta)B^T = \frac{1}{2}\sum_{i=1}^{2}\psi(u_i)\left[\begin{array}{cc} 1 & u_i \\ u_i & u_i^2 \end{array}\right]$$

where $\xi_u = \left\{ \begin{array}{cc} u_1 & u_2 \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\}$. As in Atkinson and Donev (1992, p. 292) assume that the support points of the candidate D-optimal design are symmetric about $u = 0$. Then the candidate two-point D-optimal design is

$$\xi_u = \left\{ \begin{array}{cc} -u & u \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\} \tag{4.12}$$

and the corresponding information matrix is

$$M(\xi_u) = BM(\xi;\boldsymbol{\beta})B^T = \psi(u)\left[\begin{array}{cc} 1 & 0 \\ 0 & u^2 \end{array}\right]$$

for which the determinant is

$$D = u^2\left[\psi(u)\right]^2 = \frac{u^2 e^{2u}}{(1+e^u)^4}. \tag{4.13}$$

Differentiating (4.13) with respect to u and setting the result to zero gives

$$\frac{dD}{du} = \frac{2ue^{2u}(1+u+e^u-ue^u)}{(1+e^u)^5} = 0.$$

Therefore, the stationary points of D are $u = 0$ and the solutions for u to the equation

$$1 + u + e^u - ue^u = 0. \tag{4.14}$$

The equation (4.14) has a unique solution for $u \geq 0$ for the following reasons. Firstly, consider $f(u) = 1 + u + e^u - ue^u$. Clearly, $f(u)$ is a continuous function on $[0, \infty)$. In addition, $f(0) = 2 > 0$ and $\lim_{u\to\infty} f(u) = -\lim_{u\to\infty} ue^u(1 - \frac{1}{ue^u} - \frac{1}{e^u} - \frac{1}{u}) = -\infty$. Thus, the continuous function $f(u)$ changes sign from positive to negative at least once on $[0, \infty)$. Secondly, note that $\dfrac{d^2 f(u)}{du^2} = -(1+u)e^u < 0$ for $u \geq 0$. Thus, the function $f(u)$ is strictly concave on $[0, \infty)$, and therefore the curve $f(u)$ intersects the u-axis only once. In other words, the equation (4.14) possesses a unique solution in the interval $[0, \infty)$. Figure 4.1 (a) gives a graphical representation of $f(u)$ for $u \in [0, \infty)$. Solving the equation (4.14) numerically gives $u = 1.5434$ as the unique solution on the interval $[0, \infty)$. The second derivative of D with respect to u is

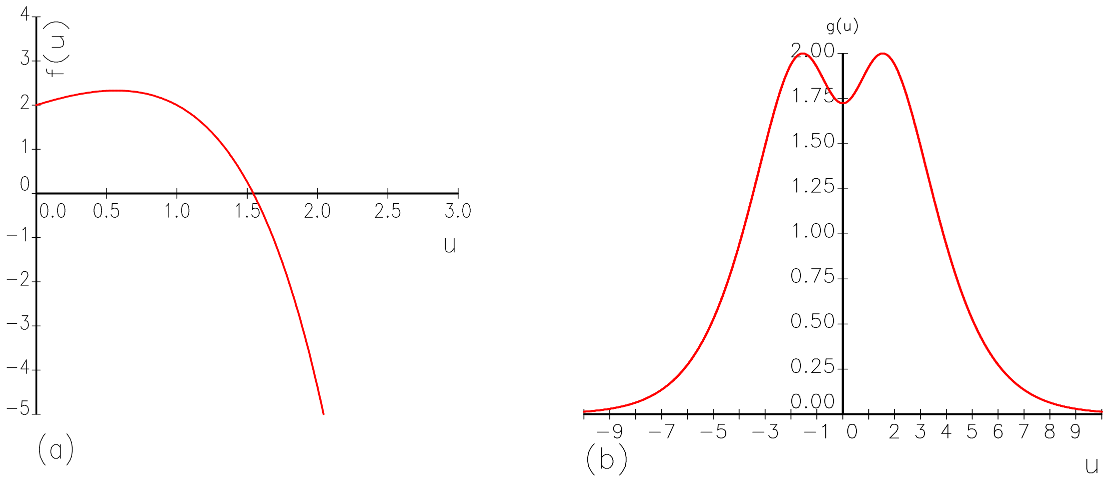


Figure 4.1: Plot of (a) the function $f(u) = 1 + u + e^u - ue^u$ for $u \geq 0$, and (b) the function $g(u) = \dfrac{e^{u-u^*}(1+e^{u^*})^2(u^{*2}+u^2)}{u^{*2}(1+e^u)^2}$ for $u \in \mathbb{R}$.

$$\frac{d^2 D}{du^2} = \frac{2e^{2u}\left[(1+4u+2u^2) + (2-6u^2)e^u + (1-4u+2u^2)e^{2u}\right]}{(1+e^u)^6}$$

so that $\left.\dfrac{d^2 D}{du^2}\right|_{u=0} = \dfrac{1}{8} > 0$ and $\left.\dfrac{d^2 D}{du^2}\right|_{u=1.5434} = -0.071 < 0$. Thus, D possesses a minimum at $u = 0$ and a maximum at $u = 1.5434$. It follows from this result that the two-point design (4.12) which maximizes determinant (4.13) is

$$\xi_u^* = \left\{ \begin{array}{cc} -1.5434 & 1.5434 \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\}. \tag{4.15}$$

Step (ii): It remains to prove that the design (4.15) is D-optimal. This is done by invoking the Equivalence Theorem for D-optimality, i.e. by showing that the directional derivative function, $\phi(u, \xi_u^*)$, at the design ξ_u^* in the direction of any point $u = \beta_0 + \beta_1 x \in \mathbb{R}$ is such that

$$\phi(u, \xi_u^*) = 2 - \mathrm{tr}[M^{-1}(\xi_u^*)M(u)] \geq 0$$

with equality holding at the support points of ξ_u^*. Equivalently, it has to be shown that the standardized variance function $d(u, \xi_u^*) = \mathrm{tr}[M^{-1}(\xi_u^*)M(u)]$ is such that

$$g(u) = d(u, \xi_u^*) = \mathrm{tr}[M^{-1}(\xi_u^*)M(u)] \leq 2$$

for all $u \in \mathbb{R}$ with equality holding at the support points of ξ_u^*.

In the present case, simple algebra gives

$$g(u) = \frac{e^{u-u^*}(1+e^{u^*})^2(u^{*^2}+u^2)}{u^{*^2}(1+e^u)^2} = \frac{\psi(u)}{\psi(u^*)}\left(1+\frac{u^2}{u^{*2}}\right) \quad \forall u \in \mathbb{R}, \tag{4.16}$$

where $u^* = 1.5434$. Note that the function $g(u)$ is positive and even in u since $g(-u) = g(u) > 0$, $\forall u \in \mathbb{R}$. The fact that $g(u)$ is even implies that $g(u)$ is symmetric about $u = 0$. Therefore, any conclusion drawn for $u \geq 0$ is applicable to $u \leq 0$.

The derivative of $g(u)$ with respect to u is

$$g'(u) = \frac{e^{u-u^*}(1+e^{u^*})^2[(u^{*^2}+u^2+2u)-(u^{*^2}+u^2-2u)e^u]}{u^{*2}(1+e^u)^3} = \frac{e^{u-u^*}(1+e^{u^*})^2}{u^{*2}(1+e^u)^3}h(u)$$

where

$$h(u) = u^{*^2} + u^2 + 2u - (u^{*^2} + u^2 - 2u)e^u. \tag{4.17}$$

Now, $g'(0) = 0$ and $g'(u^*) = \dfrac{2(1+u^*+e^{u^*}-u^*e^{u^*})}{u^*(1+e^{u^*})} = 0$ since u^* is a solution of equation (4.14). Therefore $u = 0$ and $u = u^*$ are stationary points of the function $g(u)$.

The second derivative of $g(u)$ is

$$g''(u) = \frac{e^{u-u^*}(1+e^{u^*})^2[(2+u^{*^2}+u^2)(1+e^{2u})+4u(1-e^{2u})+4(1-u^{*2}-u^2)e^u]}{u^{*2}(1+e^u)^4}.$$

Here, $g''(0) = \dfrac{e^{-u^*}(1+e^{u^*})^2(4-u^{*2})}{8u^{*2}} > 0$ since $u^* = 1.5434 < 2$ implies that $4 - u^{*2} > 0$. Therefore $u = 0$ is a minimum of $g(u)$ with $g(0) = \dfrac{e^{-u^*}(1+e^{u^*})^2}{4} \simeq 1.72 < 2$. Also,

$$g''(u^*) = \frac{2(1+2e^{u^*}+e^{2u^*}+2u^*-2u^*e^{2u^*}+u^{*2}-4u^{*2}e^{u^*}+u^{*2}e^{2u^*})}{u^{*2}(1+e^{u^*})^2} = -\frac{4e^{u^*}}{(1+e^{u^*})^2} < 0$$

since $1 + u^* + e^{u^*} - u^*e^{u^*} = 0$ implies that

$$1+2e^{u^*}+e^{2u^*}+2u^*-2u^*e^{2u^*}+u^{*2}-4u^{*2}e^{u^*}+u^{*2}e^{2u^*} = (1+u^*+e^{u^*}-u^*e^{u^*})^2-2u^{*2}e^{u^*} = -2u^{*2}e^{u^*}.$$

Hence, $u = u^*$ corresponds to a maximum of $g(u)$ with $g(u^*) = 2$ as required. Figure 4.1 (b) gives a graphical representation of $g(u)$.

To prove that $u = u^*$ is the global maximum of $g(u)$ for $u \geq 0$, consider evaluating the sign of the slope of $g(u)$ which is determined by the sign of $h(u)$ in the expression (4.17). The purpose of the derivations below is to show that $h(u) \geq 0$ for $u \in [0, u^*]$ and $h(u) \leq 0$ for $u \in [u^*, \infty)$.

It follows from (4.17) that $h(0) = 0$ and $h(u^*) = 2u^*(1+u^*+e^{u^*}-u^*e^{u^*}) = 0$ since u^* is solution of the equation (4.14). Since $h(u)$ is continuous on $[0, \infty)$, it has at least one stationary point in $(0, u^*)$. Also, $\lim_{u\to\infty} h(u) = -\lim_{u\to\infty} u^2e^u = -\infty$. Figure 4.2 (a) gives a graphical representation of $h(u)$. To prove that $h(u)$ possesses only one stationary point on $[0, \infty)$, let

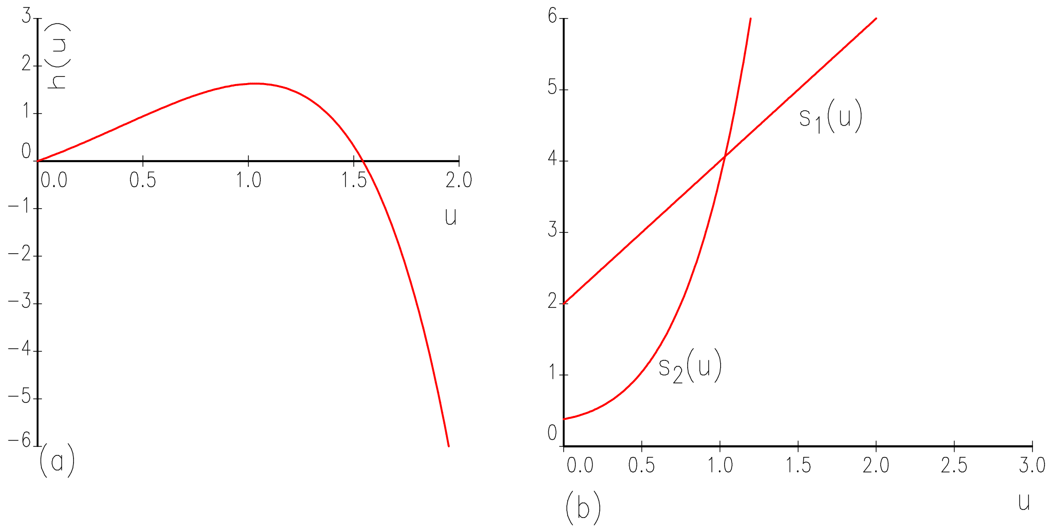


Figure 4.2: Plot of (a) the function $h(u) = u^{*^2} + u^2 + 2u - (u^{*^2} + u^2 - 2u)e^u$ for $u \geq 0$, and (b) the functions $s_1(u) = 2(1 + u)$ and $s_2(u) = (u^2 + u^{*2} - 2)e^u$ for $u \geq 0$.

$$s(u) = h'(u) = s_1(u) - s_2(u) = 2(1 + u) - (u^2 + u^{*2} - 2)e^u$$

where $s_1(u) = 2(1+u)$ and $s_2(u) = (u^2+u^{*2}-2)e^u$. Figure 4.2 (b) gives a graphical illustration of $s_1(u)$ and $s_2(u)$. The number of solutions of the equation $s(u) = h'(u) = 0$ on $[0, \infty)$ is equal to the number of points of intersection of the curves $s_1(u)$ and $s_2(u)$. Clearly, $s_1(u)$ is a straight line with intercept 2 and slope 2. Furthermore $s_2(u)$ is a convex function for all $u > 0$ since

$$s_2''(u) = (u^2 + 4u + u^{*2})e^u = (u^2 + 4u + 2.382)e^u > 0$$

and, hence there is only one point of intersection of the curves of $s_1(u)$ and $s_2(u)$. In other words the equation $h'(u) = 0$ has a unique solution u^{**}, or equivalently $h(u)$ has only one stationary point on $[0, \infty)$. Solving numerically gives $u^{**} = 1.032$ which is a maximum since $h''(u) = 2 - (u^2 + u^{*2} + 2u - 2)e^u$ implies that $h''(u^{**}) = 2 - (u^{**2} + u^{*2} + 2u^{**} - 2)e^{u^{**}} = -7.855 < 0$. Thus, the sign of $h(u)$ and hence the sign of the slope of $g(u)$ changes from positive to negative only once at $u^* = 1.5434$ on $[0, \infty)$. It can then be concluded that $u^* = 1.5434$ is the global maximum of $g(u)$ on $[0, \infty)$. By the symmetry of $g(u)$ about $u = 0$ it can also be concluded that $-u^*$ is the global maximum of $g(u)$ on $(-\infty, 0]$. ■

4.4. Two-variable binary logistic models

There are two cases of the two-variable binary logistic model that are considered in this thesis. These are the two-variable binary logistic model without interaction and the two-variable binary logistic model with interaction. Both are used to describe the effects, on the binary response, of two drugs administered simultaneously to subjects at various doses. Little research appears to have been done on the construction of optimal designs for the two-variable binary logistic model. Where research has been done, only the two-variable binary logistic model without interaction was considered in most of the cases. Well-known studies on this design problem are those by Sitter and Torsney (1995a), Atkinson and Haines (1996), and Jia and Myers (2001). Sitter and Torsney (1995b), and Torsney and Gunduz (2001) extended the work of Sitter and Torsney (1995a) to more than two explanatory variables. The two-variable binary logistic model with interaction was considered by Kupchak (2000) in his PhD thesis and by Jia and Myers (2001) in an unpublished report. Kupchak (2000) was mainly interested in finding optimal designs for the precise estimation of the interaction parameter and, to a very small extent, for the precise estimation of all model parameters, while Jia and Myers (2001) were interested in finding D-optimal designs for the precise estimation of all the model parameters.

4.4.1 Two-variable binary logistic model without interaction

The two-variable binary logistic model without interaction, which is also called the two-variable first-order or main effects binary logistic model, or the parallel line two-variable binary logistic model (Jia and Myers (2001)), is the binary logistic model described in Section 4.2 but with $\boldsymbol{x} = (x_1, x_2)^T$ and $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^T$, i.e.

$$u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2. \tag{4.18}$$

Model (4.18) assumes that two drugs, say A and B, at doses or log-doses x_1 and x_2 affect the binary response additively, i.e. the model assumes that there are no drug interaction effects on the binary response. The contours of constant logits specified by equation (4.18) are parallel straight lines with slope $-\dfrac{\beta_1}{\beta_2}$ in the (x_1, x_2)-space (Jia and Myers (2001)).

The information matrix for the parameter vector $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^T$ evaluated at the approximate design ξ given by (4.5), with $\boldsymbol{x}_i = (x_{i1}, x_{i2})^T$, is

$$M(\xi; \boldsymbol{\beta}) = \sum_{i=1}^{d} \lambda_i \psi(u_i) \widetilde{\boldsymbol{x}}_i \widetilde{\boldsymbol{x}}_i^T \tag{4.19}$$

where $\widetilde{\boldsymbol{x}}_i = [1, \boldsymbol{x}_i^T]^T = (1, x_{i1}, x_{i2})^T$ and $\psi(u_i) = \dfrac{e^{u_i}}{(1+e^{u_i})^2}$ for $u_i = \text{logit}(p_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}$ and $i = 1, 2, \ldots, d \geq 3$.

4.4.2 Two-variable binary logistic model with interaction

The two-variable binary logistic model with interaction is the model described in Section 4.2 but with $\boldsymbol{x} = (x_1, x_2, x_1 x_2)^T$ and $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$, i.e.

$$u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 \tag{4.20}$$

where the unknown parameter $\beta_{12} \neq 0$ is the interaction effect between two drugs, say A and B. In other words, model (4.20) assumes that the two drugs affect the response both additively and interactively, i.e. the model assumes that there are also drug interaction effects on the binary response. The contours of constant logits specified by equation (4.20) are pairs of hyperbolae in the (x_1, x_2)-space (Jia and Myers (2001)).

For model (4.20), the information matrix for the parameter vector $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$ evaluated at the approximate design ξ given by (4.5), with $\boldsymbol{x}_i = (x_{i1}, x_{i2}, x_{i1}x_{i2})^T$, is

$$M(\xi; \boldsymbol{\beta}) = \sum_{i=1}^{d} \lambda_i \psi(u_i) \widetilde{\boldsymbol{x}}_i \widetilde{\boldsymbol{x}}_i^T \tag{4.21}$$

where $\widetilde{\boldsymbol{x}}_i = [1, \boldsymbol{x}_i^T]^T = (1, x_{i1}, x_{i2}, x_{i1}x_{i2})^T$ and $\psi(u_i) = \dfrac{e^{u_i}}{(1+e^{u_i})^2}$ for $u_i = \text{logit}(p_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_{12} x_{i1} x_{i2}$ and $i = 1, 2, \ldots, d \geq 4$.

4.4.3 Practical examples

Example 4.1. Greco and Lawrence (1988) analyzed data originally reported by Martin (1942) and presented in Table A.1 in Appendix A. Interest focussed on investigating the effect of jointly applying the insecticides rotenone and deguelin on the number of chrysanthemum aphides that died.

SAS, PROC LOGISTIC, was used here to fit the two-variable binary logistic model with interaction to the data in Table A.1 and the results are given in Table 4.1. Clearly, $\beta_1 > 0$ and

Table 4.1: Parameter estimates together with their standard errors (s.e) and p-values for the data reported in Martin (1942).

Parameters	Estimates	s.e.	p-value
β_0	-1.9864	0.2071	< 0.0001
β_1	0.4109	0.0429	< 0.0001
β_2	0.1473	0.0141	< 0.0001
β_{12}	0.0074	0.0061	0.2237

$\beta_2 > 0$, but the interaction parameter β_{12} is not significantly different from zero. Therefore the two insecticides rotenone and deguelin can be assumed to affect the mortality of aphides additively. This example is used in Chapter 5 to illustrate the construction of D-optimal designs for the two-variable binary logistic model without interaction.

Example 4.2. Greco and Lawrence (1988) and Kupchak (2000) analyzed data originally reported by LePelly and Sullivan (1936) and presented in Table A.2 in Appendix A. Interest focussed on investigating the effect of jointly applying the insecticides rotenone and pyrethrin on the number of houseflies that died. The total number of insects exposed to each drug combination was 1000.

SAS, PROC LOGISTIC, was used to fit the two-variable binary logistic model with interaction to the data in Table A.2 and the results are given in Table 4.2. Clearly, $\beta_1 > 0$, $\beta_2 > 0$ and the interaction parameter β_{12} is significantly greater than zero. Therefore rotenone and pyrethrin affect the mortality of houseflies both additively and interactively. This is a case of the synergy action of the two insecticides since $\beta_{12} > 0$. This example is used in Chapter 6 to illustrate the construction of D-optimal designs for the two-variable binary logistic model with interaction.

Table 4.2: Parameter estimates together with their standard errors (s.e.) and p-values for the data of reported in LePelly and Sullivan (1936).

Parameter	Estimate	s.e.	p-value
β_0	-2.2054	0.0510	< 0.0001
β_1	13.5803	0.3035	< 0.0001
β_2	2.2547	0.0527	< 0.0001
β_{12}	1.6261	0.5874	< 0.0056

4.5. Review of D-optimal designs for the two-variable binary logistic model without interaction

4.5.1 Sitter and Torsney approach

Sitter and Torsney (1995a) considered constructing optimal designs for generalized linear models with two explanatory variables without interaction. For the special case of the binary logistic model (4.18), their approach to constructing D-optimal designs can be summarized as follows.

In Sitter and Torsney (1995a), the matrix (4.8) is set to $B = \begin{bmatrix} 1 & 0 & 0 \\ \beta_0 & \beta_1 & \beta_2 \\ b_0 & b_1 & b_2 \end{bmatrix}$ where b_0, b_1 and b_2 are real numbers with $\dfrac{\beta_1}{\beta_2} \neq \dfrac{b_1}{b_2}$. Hence, the information matrix (4.10) is

$$M(\xi_u) = BM(\xi;\boldsymbol{\beta})B^T = \sum_{i=1}^{d} \lambda_i \psi(u_{i1}) \begin{bmatrix} 1 \\ u_{i1} \\ u_{i2} \end{bmatrix} [1, u_{i1}, u_{i2}] = \sum_{i=1}^{d} \lambda_i \boldsymbol{g}_i \boldsymbol{g}_i^T \tag{4.22}$$

where $u_{i1} = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}$, $u_{i2} = b_0 + b_1 x_{i1} + b_2 x_{i2}$ and $\boldsymbol{g}_i = \sqrt{\psi(u_{i1})}\,[1, u_{i1}, u_{i2}]^T$ for $i = 1, 2, \ldots, d \geq 3$. Thus, for given $\boldsymbol{\beta}$, maximizing $|M(\xi_u)|$ by choice of design ξ_u is the same as maximizing $|M(\xi;\boldsymbol{\beta})|$ by choice of design ξ.

Sitter and Torsney (1995a) first considered the case of u_1 unbounded in $\mathbb{R}$, and secondly the case of u_1 bounded as $a \leq u_1 \leq b$ with $-\infty < a < 0$ and $0 < b < \infty$. In both cases u_2 was assumed to be bounded as $-1 \leq u_2 \leq 1$. The design loci generated in the two cases, described in Sitter and Torsney (1995a) as 3 dimensional "signet rings", are given by the sets

$$G_w = \left\{ \boldsymbol{g} \in \mathbb{R}^3 : \boldsymbol{g} = \sqrt{\psi(u_1)}[1, u_1, u_2]^T,\ u_1 \in \mathbb{R},\ \ -1 \leq u_2 \leq 1 \right\} \tag{4.23}$$

and

$$G = \left\{ \boldsymbol{g} \in \mathbb{R}^3 : \boldsymbol{g} = \sqrt{\psi(u_1)}[1, u_1, u_2]^T,\ a \leq u_1 \leq b,\ \ -1 \leq u_2 \leq 1 \right\} \tag{4.24}$$

respectively. Clearly $G \subset G_w$.

Sitter and Torsney (1995a) invoked Elfving (1952), Chernoff (1979) and Silvey (1980, p. 41) in stating that the support points of the D-optimal design are the points of contact between the ridges of G_w and the minimum ellipsoid centered on the origin and enclosing G_w. Sitter and Torsney (1995b) invoked Sibson (1972), Silvey and Titterington (1972), and Silvey (1980, Chapter 5) to assume the same arguments that the support points of the D-optimal design are the points of contact between the ridges of G_w and the minimum ellipsoid centered on the origin and enclosing G_w. Then, by arguments of symmetry of G_w about $u_2 = 0$, and the symmetry of $\psi(u_1)$ about $u_1 = 0$, Sitter and Torsney (1995a) conjectured that the D-optimal design for the two-variable binary logistic model without interaction is an equally weighted design with support points at $(-u_1, -1)$, $(u_1, -1)$, $(-u_1, 1)$ and $(u_1, 1)$ for some $u_1 > 0$. Thus, their approximate design (in terms of the logits) is

$$\xi_u = \left\{ \begin{array}{cccc} (-u_1, -1) & (u_1, -1) & (-u_1, 1) & (u_1, 1) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array} \right\} \tag{4.25}$$

and for this design the determinant of the matrix (4.22) is proportional to

$$D = \frac{u_1^2 e^{3u_1}}{(1 + e^{u_1})^6}. \tag{4.26}$$

The first and second derivatives of D with respect to u_1 are given by

$$\frac{dD}{du_1} = \frac{u_1 e^{u_1}(2 + 3u_1 + 2e^{u_1} - 3u_1 e^{u_1})}{(1 + e^{u_1})^7}$$

and

$$\frac{d^2 D}{du_1^2} = \frac{e^{u_1}(2 + 4e^{u_1} + 2e^{2u_1} + 12u_1 - 12u_1 e^{2u_1} + 9u_1^2 - 24u_1^2 e^{u_1} + 9u_1^2 e^{2u_1})}{(1 + e^{u_1})^8}.$$

Thus, $\dfrac{dD}{du_1} = 0$ implies that $u_1 = 0$ or

$$2 + 3u_1 + 2e^{u_1} - 3u_1 e^{u_1} = 0. \tag{4.27}$$

Solving equation (4.27) numerically for $u_1 \geq 0$ gives $u_1 = 1.22291$. Since $\left.\dfrac{d^2 D}{du_1^2}\right|_{u=0} = \dfrac{1}{32} > 0$ and $\left.\dfrac{d^2 D}{du_1^2}\right|_{u=1.22291} = -0.0017 < 0$, then the respective minimum and maximum of determinant

D are at $u_1 = 0$ and $u_1 = 1.22291$. Thus, design (4.25) which maximizes (4.26), in the (u_1, u_2)-space, is

$$\xi_u^* = \left\{ \begin{array}{cccc} (-1.22291, -1) & (1.22291, -1) & (-1.22291, 1) & (1.22291, 1) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array} \right\}. \tag{4.28}$$

Sitter and Torsney (1995a) did not prove the D-optimality of design (4.28), but simply argued D-optimality of the design using intuitive geometric arguments. The proof of D-optimality of the design is given in Chapter 5 of this thesis.

In the case of the design locus G given by (4.24), Sitter and Torsney (1995a) report that the support points of the D-optimal design are located on the ridge of G, and that half of the weights are put on the support points on $u_2 = -1$ and the other half on the support points on $u_2 = 1$. Furthermore, Sitter and Torsney (1995a) were interested in the number and the patterns of the support points of the D-optimal design. For the two-variable binary logistic model with no interaction, they report that the D-optimal design has 4 support points and thus is a design of the form

$$\xi_u = \left\{ \begin{array}{cccc} (a^*, -1) & (b^*, -1) & (a^*, 1) & (b^*, 1) \\ \frac{\lambda}{2} & \frac{1-\lambda}{2} & \frac{\lambda}{2} & \frac{1-\lambda}{2} \end{array} \right\} \tag{4.29}$$

for $a^*, b^* \in [a, b]$. Sitter and Torsney (1995a) present the following cases of design ξ_u.

(1) If $(-1.22291, 1.22291) \subset [a, b]$, then the D-optimal design is design (4.28), the design for the case of design locus G_w given in (4.23).

(2) If $-1.22291 < a$, then the candidate D-optimal design is (4.29) with $a^* = a$ and $b^* = \min(b, u_1^*)$ where u_1^* maximizes the determinant of the information matrix $M(\xi_u)$ in (4.22) on $[a, b]$.

(3) If $1.22291 > b$, then the D-optimal design is (4.29) with $b^* = b$ and $a^* = \max(a, u_1^*)$ where u_1^* maximizes the determinant of (4.22) on $[a, b]$.

(4) If $[a, b] \subset (-1.22291, 1.22291)$, then the D-optimal design is (4.29) with $a^* = a$ and $b^* = b$.

4.5.2 Atkinson and Haines approach

Atkinson and Haines (1996) considered constructing optimal designs for nonlinear and generalized linear models with the two-variable binary logistic model without interaction (4.18) as one of the special cases. In model (4.18), the explanatory variables were taken to be $-1 \leq x_1 \leq 1$ and $-1 \leq x_2 \leq 1$, and the D-optimal designs were found numerically. Their results indicated that the designs are highly dependent on the values adopted for the parameters. In particular, the number of support points of the D-optimal designs were found to be 3, 4, or 6 depending on the values of the parameters β_0, β_1 and β_2. Table 4.3 gives the D-optimal designs for selected values of the parameters. Clearly, the number of support points, the weights and the optimal logits are parameter dependent. The optimal logits in Table 4.3 are the same as those found in Sitter and Torsney (1995a) and in Jia and Myers (2001) only for the design ξ_6^* for which $\boldsymbol{\beta} = (0,3,1)^T$. Similar conclusions concerning parameter dependency of the

Table 4.3: D-optimal logits and the number of support points of the D-optimal designs of Atkinson and Haines (1996) for selected values of the parameters of the two-variable binary logistic model without interaction on the design space $\mathfrak{X} = [-1,1] \times [-1,1]$.

$\boldsymbol{\beta}^T$	D-optimal designs
$(0,1,1)$	$\xi_1^* = \left\{ \begin{array}{cccc} (-1,-1) & (-1,1) & (1,1) & (1,-1) \\ 0.2041 & 0.2959 & 0.2041 & 0.2959 \\ -2 & 0 & 2 & 0 \end{array} \right\}$
$(0,2,2)$	$\xi_2^* = \left\{ \begin{array}{cccccc} (0.118,-1) & (-1,0.118) & (-1,1) & (-0.118,1) & (1,-0.118) & (1,-1) \\ 0.1306 & 0.1098 & 0.2654 & 0.1098 & 0.1306 & 0.2538 \\ -1.764 & -1.764 & 0 & 1.764 & 1.764 & 0 \end{array} \right\}$
$(2,2,2)$	$\xi_3^* = \left\{ \begin{array}{cccc} (-0.737,-1) & (-1,-0.737) & (-1,0.737) & (0.737,-1) \\ 0.1686 & 0.1686 & 0.3314 & 0.3314 \\ -1.474 & -1.474 & 1.474 & 1.474 \end{array} \right\}$
$(0,\frac{3}{2},\frac{1}{2})$	$\xi_4^* = \left\{ \begin{array}{cccc} (-0.575,-1) & (-1,1) & (0.575,1) & (1,-1) \\ 0.2381 & 0.2619 & 0.2381 & 0.2619 \\ -1.363 & -1 & 1.363 & 1 \end{array} \right\}$
$(3,3,1)$	$\xi_5^* = \left\{ \begin{array}{ccc} (-1,-1) & (-1,1) & (-0.068,-1) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ -1 & 1 & 1.796 \end{array} \right\}$
$(0,3,1)$	$\xi_6^* = \left\{ \begin{array}{cccc} (-0.074,-1) & (-0.741,1) & (0.074,1) & (0.741,-1) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ -1.22291 & -1.22291 & 1.22291 & 1.22291 \end{array} \right\}$

D-optimal designs for the two-variable binary logistic model without interaction on the design space $[-1,1] \times [-1,1]$ are reported in Chipman and Welch (1996).

The design patterns reported in Atkinson and Haines (1996) are very different to those reported in Sitter and Torsney (1995a) because they are not restricted to 4-point designs. Clearly the observed relationships between the design patterns and the parameter values need to be explained.

4.5.3 Jia and Myers approach

The approach of Jia and Myers (2001) to constructing D-optimal designs for the two-variable binary logistic model without interaction (4.18) is as follows. The authors assumed that the support points of the D-optimal design lie on lines of constant logit in the (x_1, x_2)-space. The information matrix (4.19) depends on u through $\psi(u)$. Since $\psi(-u) = \psi(u)$, it appears that optimal logits are of the form $\pm u$. Jia and Myers (2001) imposed restrictions on the design space of the form $b_{01} + b_1x_1 + b_2x_2 = 0$ and $b_{02} + b_1x_1 + b_2x_2 = 0$ with $b_{01} \neq b_{02}$ and $\dfrac{b_1}{b_2} \neq \dfrac{\beta_1}{\beta_2}$. Figure 4.3 displays the two logit lines $\pm u$ and the two parallel lines that restrict the design space. The vertices of the parallelogram, indicated by circles, were taken as candidate support points of the D-optimal design. Support points on the same logit line were assumed to have

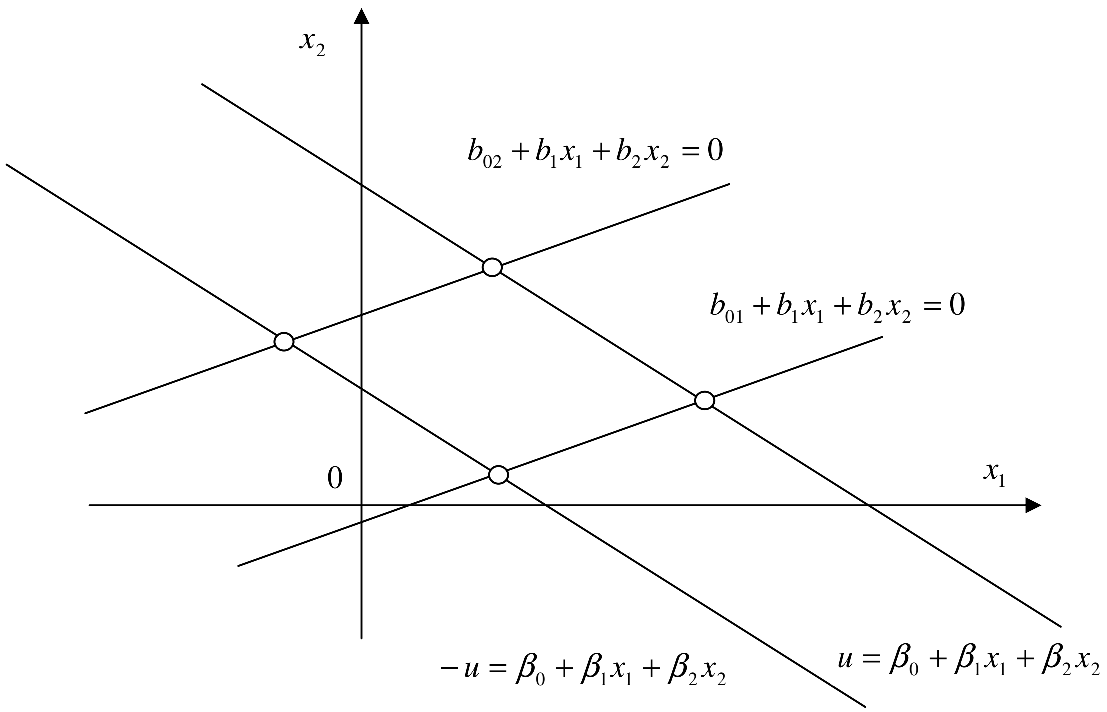


Figure 4.3: The design for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1x_1 + \beta_2x_2$: Jia and Myers (2001) approach. The support points of the candidate D-optimal design are represented by circles.

equal weights, denoted by $\dfrac{1-\lambda}{2}$ and $\dfrac{\lambda}{2}$ for each candidate support point located on the logits $-u$ and u respectively. Simple algebra shows that the coordinates of the support points are functions of u, $\mathbf{b} = (b_{01}, b_{02}, b_1, b_2)^T$ and $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^T$. The determinant of the information matrix (4.19) evaluated at the candidate design of Jia and Myers (2001) is given by

$$D = \frac{\lambda(1-\lambda)(b_{02}-b_{01})^2 u^2 e^{3u}}{(b_2\beta_1 - b_1\beta_2)^2(1+e^u)^6}. \tag{4.30}$$

Clearly the determinant (4.30) is maximized by $\lambda = \frac{1}{2}$. In addition, the determinant (4.30) is proportional to the determinant (4.26) if u_1 in (4.26) is replaced by u, and hence the optimal values of u here are also $u^* = \pm 1.22291$. Thus, the approach of Jia and Myers (2001) gives the same design as that of Sitter and Torsney (1995a) in the case of the design locus (4.23). The approach of Jia and Myers (2001) for the two-variable binary logistic model without interaction is also briefly discussed in Myers et al. (2002, pp. 240-241).

4.5.4 Some summary comments

There are some similarities, but also differences, among the results of Sitter and Torsney (1995a), Atkinson and Haines (1996) and Jia and Myers (2001) on the D-optimal designs for the two-variable binary logistic model without interaction.

1) Jia and Myers (2001) considered a design space restricted to lie inside the region defined by two parallel lines $b_{01} + b_1x_1 + b_2x_2 = 0$ and $b_{02} + b_1x_1 + b_2x_2 = 0$ with $b_{01} \neq b_{02}$ in (x_1, x_2)-space. The candidate D-optimal design derived using the model with $u_1 = \text{logit}(p) = \beta_0 + \beta_1x_1 + \beta_2x_2$ was found to be an equally weighted 4-point design with support points located in pairs on each of the logit lines $\pm u_1 = 1.22291$.

2) Sitter and Torsney (1995a) used a canonical transformation from the (x_1, x_2)-space to the (u_1, u_2)-space where $u_1 = \text{logit}(p) = \beta_0 + \beta_1x_1 + \beta_2x_2$ and $u_2 = b_0 + b_1x_1 + b_2x_2$. Then they restricted u_2 as $-1 \leq u_2 \leq 1$ and considered two design cases, namely the case of $u_1 \in \mathbb{R}$ and the case of $u_1 \in [a, b]$ with $a < 0$ and $b > 0$. The D-optimal design in the case of $u_1 \in \mathbb{R}$ (design locus (4.23)) is an equally weighted 4-point design with support points located in pairs on each of the logit lines $\pm u_1 = 1.22291$. The similarity of the results of Jia and Myers (2001) and Sitter and Torsney (1995a) (design locus (4.23)) is due to the fact that for $u_1 \in \mathbb{R}$, the design spaces of Jia and Myers (2001) and Sitter and Torsney (1995a) are equivalent. In fact, the restriction $-1 \leq u_2 \leq 1$ in the design locus (4.23) is equivalent to the restrictions $b_{01} + b_1x_1 + b_2x_2 \geq 0$ and $b_{02} + b_1x_1 + b_2x_2 \leq 0$ where $b_{01} = b_0 + 1$ and $b_{02} = b_0 - 1$. The above

inequalities jointly define the set of points within the parallel lines $b_{01} + b_1 x_1 + b_2 x_2 = 0$ and $b_{02} + b_1 x_1 + b_2 x_2 = 0$ in Figure 4.3 with u replaced by u_1.

In the case of $u_1 \in [a, b]$ (design locus (4.24)), Sitter and Torsney (1995a) present four design scenarios. Figure 4.4 indicates the restriction $a \le u_1 \le b$ as the set of points within the dashed lines $u_1 = a$ and $u_1 = b$ that are both parallel to the logit lines $\pm u_1$. The constraint $u_1 \in [a, b]$ and the restriction $-1 \le u_2 \le 1$ jointly define a parallelogram design space in the (x_1, x_2)-space. This case is not discussed by Jia and Myers (2001).

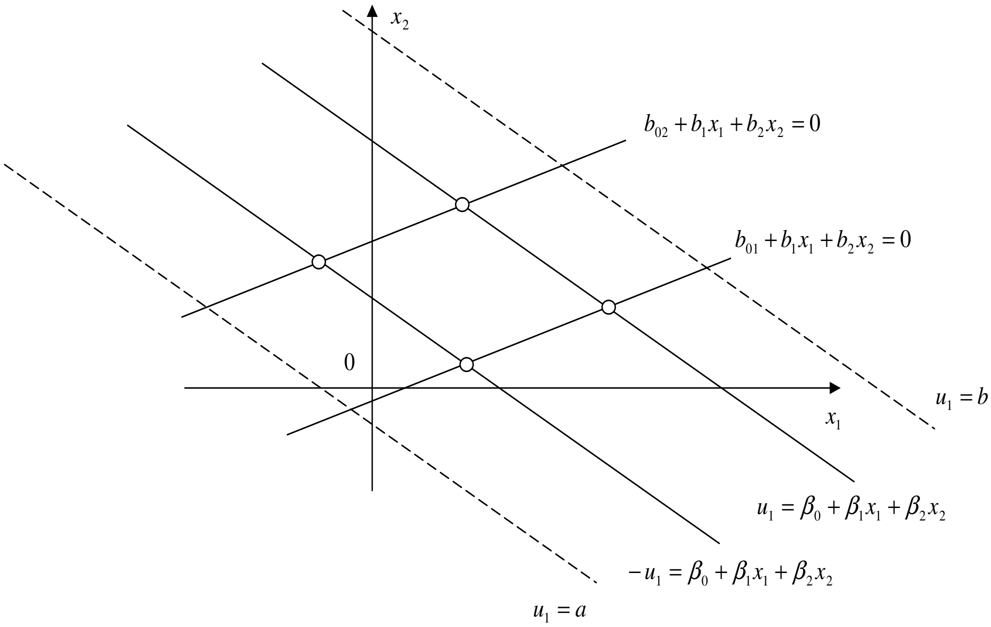


Figure 4.4: The design for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$: Sitter and Torsney (1995a) design locus (4.24) approach. The support points of the candidate D-optimal design are represented by circles.

3) Atkinson and Haines (1996) considered designs on the design space $[-1, 1] \times [-1, 1]$ in the (x_1, x_2)-space and found that the number, which varies from 3 to 6, and patterns of the support points and associated weights of the D-optimal designs are parameter dependent. Let $u_1 = \beta_0 + \beta_1 x_1 + \beta_2 x_2$. Clearly, $(x_1, x_2) \in [-1, 1] \times [-1, 1]$ implies that $a \le u_1 \le b$ where $a = \beta_0 - \beta_1 - \beta_2$ and $b = \beta_0 + \beta_1 + \beta_2$ are the values of the logit u_1 at the vertices $A(-1, -1)$ and $C(1, 1)$ in the design space $[-1, 1] \times [-1, 1]$ represented by the square $ABCD$ in Figures 4.5 (a) and (b). The restriction $a \le u_1 \le b$ is similar to that imposed by Sitter and Torsney (1995a) on $u_1 = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ in design locus (4.24). However, in Atkinson and Haines

(1996), there are two restrictions on $u_2 = b_0 + b_1x_1 + b_2x_2$ instead of only one in Sitter and Torsney (1995a). These two restrictions are $-1 \leq u_2 = x_1 \leq 1$ (i.e. $b_0 = b_2 = 0$) and $-1 \leq u_2 = x_2 \leq 1$ (i.e. $b_0 = b_1 = 0$).

The difference between the designs of Atkinson and Haines (1996) and those of Sitter and Torsney (1995a) and Jia and Myers (2001) can be attributed to the differences in the design space restrictions. For example, consider Figure 4.5 (a) in the context of the design space of Sitter and Torsney (1995a), and thus design locus (4.23). If $u_1 \in \mathbb{R}$ and $-1 \leq u_2 = x_2 \leq 1$,

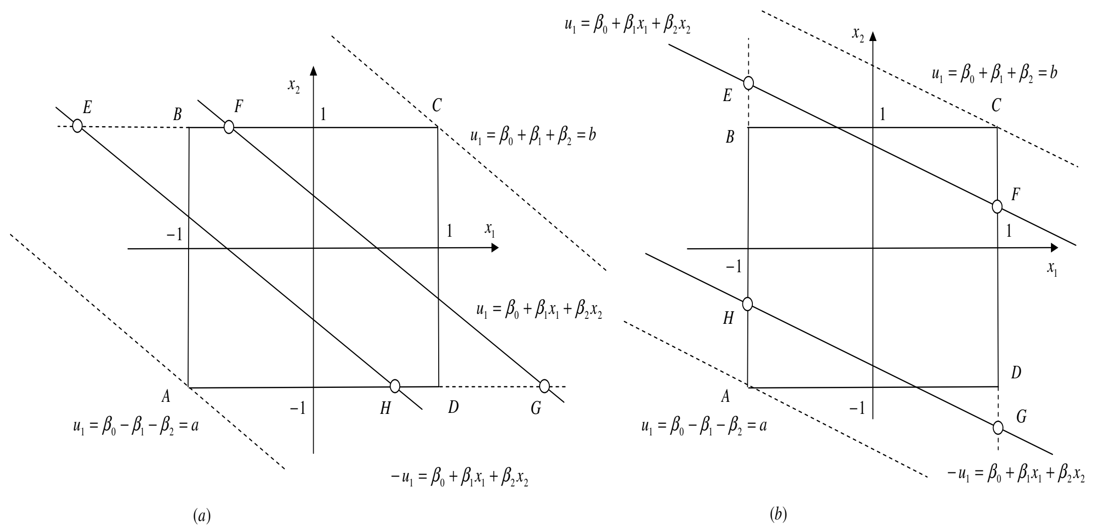


Figure 4.5: The design of the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1x_1 + \beta_2x_2$ using $[-1, 1] \times [-1, 1]$ as the design space of reference.

it follows that $x_1 \in \mathbb{R}$. Furthermore, if the optimal logits are the lines HE and GF, then the support points of the D-optimal designs of Sitter and Torsney (1995a) are H, E, F and G. Note that the horizontal dashed lines in Figure 4.5 (a) extending the continuous lines serve to indicate the intersections of the logits $\pm u_1$ and the parallel line restrictions in case the restriction $-1 \leq x_1 \leq 1$ was not imposed on the design space. However, if the restriction on $-1 \leq x_1 \leq 1$ is taken into account the points E and G are outside the square $ABCD$. Hence, the points E and G would not satisfy the requirements to be support points of the D-optimal designs on the design space of Atkinson and Haines (1996). Similarly, considering Figure 4.5 (b) and assuming that $x_2 \in \mathbb{R}$ with $-1 \leq u_2 = x_1 \leq 1$, leads to a D-optimal design with support points at E, F, G and H. Here the vertical dashed lines in Figure 4.5 (b) extending the continuous lines serve to indicate the intersections of the logits $\pm u_1$ and the parallel line restrictions in case the restriction $-1 \leq x_2 \leq 1$ was not imposed on the design space. However, points E and G are outside the square $ABCD$, and hence do not satisfy the conditions to be support points of the D-optimal designs on the design space of Atkinson and Haines (1996).

Despite the difference in design space restrictions, there are cases where the designs of Atkinson and Haines (1996) are similar to those of Sitter and Torsney (1995a), and Jia and Myers (2001). For example, the designs are similar when the two logit lines $-u_1 = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ and $u_1 = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ intersect two parallel sides of the design space $[-1, 1] \times [-1, 1]$. In such cases, $0 < u_1 \leq 1.22291 \leq \min\{\beta_1 - \beta_2 + \beta_0, \beta_1 - \beta_2 - \beta_0\}$ as illustrated in Figure 4.6 (a), or $0 < u_1 \leq 1.22291 \leq \min\{\beta_2 - \beta_1 + \beta_0, \beta_2 - \beta_1 - \beta_0\}$ as illustrated in Figure 4.6 (b). The two conditions are jointly satisfied by $0 < u_1 \leq 1.22291 \leq \min\{|\beta_1 - \beta_2| + \beta_0, |\beta_1 - \beta_2| - \beta_0\}$.

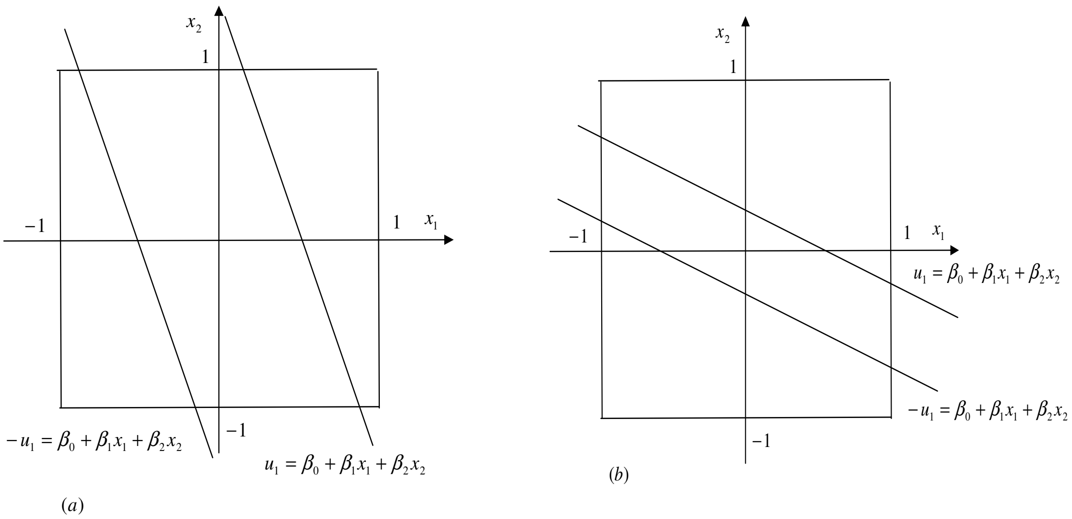


Figure 4.6: The design for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$: Atkinson and Haines (1996) approach. (a) $0 < u_1 \leq 1.22291 \leq \min\{\beta_1 - \beta_2 + \beta_0, \beta_1 - \beta_2 - \beta_0\}$ (b) $0 < u_1 \leq 1.22291 \leq \min\{\beta_2 - \beta_1 + \beta_0, \beta_2 - \beta_1 - \beta_0\}$.

In the three papers, namely those by Sitter and Torsney (1995a), Atkinson and Haines (1996) and Jia and Myers (2001), the authors constructed D-optimal designs but did not provide proofs of the D-optimality of their designs. In addition, the conditions for the D-optimality of the 3-point, 4-point and 6-point designs of Atkinson and Haines (1996) need to be fully investigated since rectangular design spaces are generally re-parameterized as in Atkinson and Haines (1996) (see for example Atkinson and Donev (1992), Myers and Montgomery (1995), and Atkinson et al. (2007)). These two issues will be addressed in this thesis, and the geometry of Jia and Myers (2001) will be used since it is easier to visualize and to generalize than the approach of Sitter and Torsney (1995a).

4.6. Review of D-optimal design for the two-variable binary logistic model with interaction

4.6.1 Kupchak approach

Kupchak (2000) investigated c-, G- and D-optimal designs for the two-variable binary logistic model with interaction

$$u = \text{logit}(p) = \beta_0 + \beta_1 d_1 + \beta_2 d_2 + \beta_{12} d_1 d_2 \tag{4.31}$$

as specified in (4.20), where $d_1 \geq 0$ and $d_2 \geq 0$ represent the doses of two drugs, say A and B. Kupchak (2000) first considered the case of designs for detecting "local synergy" which are the c- and the G-optimal designs for the precise estimation of the interaction parameter β_{12} in the model with constant probability of response p. Next, he considered the case of the designs for detecting "global synergy" which are the c-optimal designs for the precise estimation of the interaction parameter β_{12}, and the D-optimal designs for the precise estimation of all the parameters under the assumption that $\beta_{12} = 0$, in the model with no constraint on the probability of response p. This section briefly describes the Kupchak's approach to constructing optimal designs for detecting "local" and "global" synergy with emphasis on the construction of D-optimal designs which are the main subject of this thesis.

Optimal designs for detecting local synergy

Kupchak (2000) considered model (4.31), with p constant, on the restricted design space

$$\mathcal{D}_p = \left\{(d_1, d_2) : \text{logit}(p) = \beta_0 + \beta_1 d_1 + \beta_2 d_2\right\}.$$

The information matrix $M(\xi; \boldsymbol{\beta})$ for model (4.31), as given by (4.21), is singular because the support points of the design lie on the straight line $\mathcal{D}_p$. Therefore, the D-optimality criterion is not applicable in this case. However, the c- or the G- optimality criterion can be used.

(1) c-optimality criterion

For the precise estimation of the parameter β_{12}, Kupchak (2000) proposed minimizing the criterion

$$\text{Var}\left(\widehat{\beta}_{12}\right) = \boldsymbol{a}^T M^{-}(\xi; \boldsymbol{\beta}) \boldsymbol{a} \tag{4.32}$$

where $M^{-}(\xi;\boldsymbol{\beta})$ is a generalized inverse of $M(\xi;\boldsymbol{\beta})$ for all $M(\xi;\boldsymbol{\beta})$ such that $\boldsymbol{a} = [0\ 0\ 0\ 1]^T = M(\xi;\boldsymbol{\beta})\boldsymbol{q}$, for some vector $\boldsymbol{q} = [q_1, q_2, q_3, q_4]$, i.e. such that $\boldsymbol{q}$ is in the column space of $M(\xi;\boldsymbol{\beta})$. As the design ξ that minimizes criterion (4.32), Kupchak (2000) suggested, with no explanation, the 3-point design

$$\xi^* = \left\{ \begin{array}{ccc} (ED_{100p,1}, 0) & (\frac{1}{2}ED_{100p,1}, \frac{1}{2}ED_{100p,2}) & (0, ED_{100p,2}) \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{array} \right\} \tag{4.33}$$

where $ED_{100p,1} = \dfrac{\text{logit}(p) - \beta_0}{\beta_1}$ and $ED_{100p,2} = \dfrac{\text{logit}(p) - \beta_0}{\beta_2}$ are the respective doses of A and B associated with the fixed probability of response p. Kupchak (2000) used the General Equivalence Theorem for c-optimality to prove that the design (4.33) is globally optimal.

(2) G-optimality criterion

On the design space $\mathcal{D}_p$, the information matrix $M(\xi;\boldsymbol{\beta})$ is singular because the support points of the design ξ lie on the straight line defined by D_p and hence the D-optimality criterion is not applicable. Kupchak (2000) invoked Atkinson and Donev (1992, p. 114) in suggesting a design ξ that minimizes the G-optimality criterion $\Psi_p(\xi) = \max\limits_{c\in[0,1]} \boldsymbol{c}^T M^{-}(\xi;\boldsymbol{\beta})\boldsymbol{c}$ where $\boldsymbol{c} = \sqrt{\psi(u)}[1, c, (1-c), c(1-c)]^T$ for some vector $\boldsymbol{q}$ such that $\boldsymbol{c} = M(\xi;\boldsymbol{\beta})\boldsymbol{q}$. Clearly, $\boldsymbol{c}^T M^{-}(\xi;\boldsymbol{\beta})\boldsymbol{c} = \boldsymbol{q}^T\boldsymbol{c}$ and hence, the criterion $\Psi_p(\xi)$ is invariant to the choice of a generalized inverse $M^{-}(\xi;\boldsymbol{\beta})$. Kupchak (2000) proposed, without explanation, the 3-point G-optimal design

$$\xi_G^* = \left\{ \begin{array}{ccc} (ED_{100p,1}, 0) & (\frac{1}{2}ED_{100p,1}, \frac{1}{2}ED_{100p,2}) & (0, ED_{100p,2}) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array} \right\}$$

and showed that the design is G-optimal under the null hypothesis that $\beta_{12} = 0$.

Optimal designs for detecting "global synergy"

Kupchak (2000) also considered model (4.31) with no constraint on the probability of response p on the design space given by

$$\mathcal{D} = \{(d_1, d_2) : 0 \le d_1 \le a, 0 \le d_2 \le b\}$$

where a and b are positive real numbers. For this case, Kupchak (2000) discussed the construction of c-optimal designs for minimizing $\text{Var}(\widehat{\beta}_{12})$, and D-optimal designs for minimizing the determinant of the asymptotic variance-covariance matrix of $\widehat{\boldsymbol{\beta}}$. The two criteria are briefly described below following Kupchak's approach.

(1) c-optimality criterion

Kupchak (2000) argued that fixing the vector of parameters $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$ at a known value implies that the median effective doses $ED_{50,1} = -\dfrac{\beta_0}{\beta_1}$ and $ED_{50,2} = -\dfrac{\beta_0}{\beta_2}$ of the drugs A and B respectively are also known. This reduces the 4 parameter model (4.31) to the 2 parameter model

$$u = \text{logit}(p) = \beta_0(1 - d_1^* - d_2^*) + \beta_I d_1^* d_2^* \tag{4.34}$$

where $d_1^* = \dfrac{d_1}{ED_{50,1}}$, $d_2^* = \dfrac{d_2}{ED_{50,2}}$ and $\beta_I = \beta_{12} ED_{50,1} ED_{50,2}$. Thus, the information matrix for the parameters of model (4.34) evaluated at a single point $\boldsymbol{x}^* = (d_1^*, d_2^*)$ is

$$M(\boldsymbol{x}^*; \beta_0, \beta_I) = \psi(u) \begin{bmatrix} (1 - d_1^* - d_2^*)^2 & (1 - d_1^* - d_2^*)d_1^* d_2^* \\ (1 - d_1^* - d_2^*)d_1^* d_2^* & d_1^{*2} d_2^{*2} \end{bmatrix}$$

and the information matrix for the parameters of model (4.34) evaluated at the approximate design of the form

$$\xi = \left\{ \begin{array}{cccc} (d_{11}^*, d_{12}^*) & (d_{21}^*, d_{22}^*) & \ldots & (d_{d1}^*, d_{d2}^*) \\ \lambda_1 & \lambda_2 & \ldots & \lambda_d \end{array} \right\} \tag{4.35}$$

is the 2×2 matrix

$$M(\xi; \beta_0, \beta_I) = \sum_{i=1}^{d} \lambda_i \psi(u_i) \begin{bmatrix} (1 - d_{i1}^* - d_{i2}^*)^2 & (1 - d_{i1}^* - d_{i2}^*)d_{i1}^* d_{i2}^* \\ (1 - d_{i1}^* - d_{i2}^*)d_{i1}^* d_{i2}^* & d_{i1}^{*2} d_{i2}^{*2} \end{bmatrix}. \tag{4.36}$$

The c-optimality criterion proposed by Kupchak (2000) for the precise estimation of β_I, and hence for the precise estimation of β_{12}, is the asymptotic variance of $\widehat{\beta}_I$ given by

$$\text{Var}(\widehat{\beta}_I) = \left[M^{-1}(\xi; \beta_0, \beta_I)\right]_{22} \tag{4.37}$$

where $[M^{-1}(\xi; \beta_0, \beta_I)]_{22}$ is the $(2,2)$ element of $M^{-1}(\xi; \beta_0, \beta_I)$. Kupchak (2000) showed that if in model (4.34) $\beta_0 < -2\sqrt{2}$ and $\beta_I = 0$, then the optimal design ξ which minimizes criterion (4.37) is

$$\xi_I^* = \left\{ \begin{array}{c} (\frac{1}{2}, \frac{1}{2}) \\ 1 \end{array} \right\}.$$

Kupchak (2000) also showed that for some δ where $0 < \delta < 1$, if in model (4.34) $\beta_I = 0$ and $-2\sqrt{2} < \beta_0 < -2.39936$ where -2.39936 is a numerical solution of the equation $\beta_0(1 - e^{-\beta_0}) = 2(1 + e^{-\beta_0})$, then the optimal design ξ which minimizes criterion (4.37) is

$$\xi_I^* = \left\{ \begin{array}{cc} \frac{1}{2}(1 - \delta, 1 - \delta) & \frac{1}{2}(1 + \delta, 1 + \delta) \\ \frac{1}{2} & \frac{1}{2} \end{array} \right\}.$$

Finally, Kupchak (2000) reported that if $\beta_0 \geq -2.39936$, then the c-optimal design for estimating β_I is an unequally weighted two-point design with support points at $(0, 0)$ and at a point (d_1^*, d_2^*) which can only be determined numerically.

(2) D-optimality criterion

Kupchak (2000) also considered constructing designs for the precise estimation of all the parameters of model (4.31) with p not fixed, i.e. $0 < p < 1$. Thus, he used the D-optimality criterion

$$\Psi(\xi;\boldsymbol{\beta}) = \log|M^{-1}(\xi;\boldsymbol{\beta})| = -\log|M(\xi;\boldsymbol{\beta})|.$$

As in (4.21), the information matrix $M(\xi;\boldsymbol{\beta})$ at the approximate design of the form

$$\xi_d = \left\{ \begin{array}{cccc} (d_{11}, d_{12}) & (d_{21}, d_{22}) & \dots & (d_{d1}, d_{d2}) \\ \lambda_1 & \lambda_2 & \dots & \lambda_d \end{array} \right\} \tag{4.38}$$

is the 4×4 matrix

$$M(\xi_d;\boldsymbol{\beta}) = \sum_{i=1}^{d} \lambda_i \psi(u_i) \begin{bmatrix} 1 & d_{i1} & d_{i2} & d_{i1}d_{i2} \\ d_{i1} & d_{i1}^2 & d_{i1}d_{i2} & d_{i1}^2 d_{i2} \\ d_{i2} & d_{i1}d_{i2} & d_{i2}^2 & d_{i1}d_{i2}^2 \\ d_{i1}d_{i2} & d_{i1}^2 d_{i2} & d_{i1}d_{i2}^2 & d_{i1}^2 d_{i2}^2 \end{bmatrix} \tag{4.39}$$

where $\psi(u_i) = \dfrac{e^{u_i}}{(1+e^{u_i})^2}$ for $i = 1, 2, \dots, d \geq 4$. Kupchak (2000) used an example for numerically constructing D-optimal designs. Then, he proved that, under the hypothesis that $\beta_{12} = 0$, any candidate D-optimal design is independent of the choice of the parameters β_1 and β_2 as follows. Kupchak re-parameterized model (4.31) with $d_1^* = \dfrac{d_1}{ED_{100p,1}}$ and $d_2^* = \dfrac{d_2}{ED_{100p,2}}$ to obtain

$$u = \text{logit}(p) = \beta_0 + \beta_1^* d_1^* + \beta_2^* d_2^* + \beta_{12}^* d_1^* d_2^*$$

where $\beta_1^* = \beta_1 ED_{100p,1}$, $\beta_2^* = \beta_2 ED_{100p,2}$ and $\beta_{12}^* = \beta_{12} ED_{100p,1} ED_{100p,2}$. In terms of $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$ and the alternative parameter vector $\boldsymbol{\beta}^* = (\beta_0, \beta_1^*, \beta_2^*, \beta_{12}^*)^T$, d_1^* and d_2^* can be expressed as

$$d_1^* = \frac{\beta_1 d_1}{\beta_1 ED_{100p,1}} = \frac{\beta_1}{\beta_1^*} d_1 \quad \text{and} \quad d_2^* = \frac{\beta_2 d_2}{\beta_2 ED_{100p,2}} = \frac{\beta_2}{\beta_2^*} d_2.$$

Hence, the vector $\widetilde{\boldsymbol{x}} = (1, d_1, d_2, d_1 d_2)^T$ is transformed to $\widetilde{\boldsymbol{x}}^* = \left[1, \dfrac{\beta_1}{\beta_1^*} d_1, \dfrac{\beta_2}{\beta_2^*} d_2, \dfrac{\beta_1}{\beta_1^*}\dfrac{\beta_2}{\beta_2^*} d_1 d_2\right]^T = B\widetilde{\boldsymbol{x}}$ where

$$B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{\beta_1}{\beta_1^*} & 0 & 0 \\ 0 & 0 & \frac{\beta_2}{\beta_2^*} & 0 \\ 0 & 0 & 0 & \frac{\beta_1}{\beta_1^*}\frac{\beta_2}{\beta_2^*} \end{bmatrix},$$

and the design (4.38) is transformed to the design

$$\xi_{d^*} = \left\{ \begin{array}{cccc} (d_{11}^*, d_{12}^*) & (d_{21}^*, d_{22}^*) & \dots & (d_{d1}^*, d_{d2}^*) \\ \lambda_1 & \lambda_2 & \dots & \lambda_d \end{array} \right\}. \tag{4.40}$$

As a consequence, the information matrix for the parameter vector $\boldsymbol{\beta}^*$ evaluated at the design (4.40) is $M(\xi_{d^*}; \boldsymbol{\beta}^*) = BM(\xi_d; \boldsymbol{\beta})B$ where $M(\xi_d; \boldsymbol{\beta})$ is given by (4.39). Therefore, $|M(\xi_{d^*}; \boldsymbol{\beta}^*)| = |B|^2 |M(\xi_d; \boldsymbol{\beta})|$ and thus minimizing $\ln |M^{-1}(\xi_{d^*}; \boldsymbol{\beta}^*)|$ by choice of ξ_{d^*} is equivalent to minimizing $\ln |M^{-1}(\xi_d; \boldsymbol{\beta})|$ by choice of ξ_d. Furthermore, the directional derivative function is invariant to the choice of the parameters β_1 and β_2. In fact, $M(\xi_{d^*}; \boldsymbol{\beta}^*) = BM(\xi_d; \boldsymbol{\beta})B$ and $\psi(u) = p(1-p) = \dfrac{e^u}{(1+e^u)^2}$ imply that

$$M^{-1}(\xi_{d^*}; \boldsymbol{\beta}^*) = B^{-1} M^{-1}(\xi_d; \boldsymbol{\beta}) B^{-1},$$

and

$$\psi(u) \widetilde{\boldsymbol{x}}^{*T} M^{-1}(\xi_{d^*}, \boldsymbol{\beta}^*) \widetilde{\boldsymbol{x}}^* = \psi(u) \widetilde{\boldsymbol{x}}^T B \left(B^{-1} M^{-1}(\xi_d; \boldsymbol{\beta}) B^{-1} \right) B \widetilde{\boldsymbol{x}} = \psi(u) \widetilde{\boldsymbol{x}}^T M^{-1}(\xi_d; \boldsymbol{\beta}) \widetilde{\boldsymbol{x}}.$$

Consequently, the directional derivation functions of the designs (4.38) and (4.40) are equal because

$$\phi(\boldsymbol{x}^*, \xi_{d^*}; \boldsymbol{\beta}^*) = 4 - \psi(u) \widetilde{\boldsymbol{x}}^{*T} M^{-1}(\xi_{d^*}; \boldsymbol{\beta}^*) \widetilde{\boldsymbol{x}}^* = 4 - \psi(u) \widetilde{\boldsymbol{x}}^T M^{-1}(\xi_d; \boldsymbol{\beta}) \widetilde{\boldsymbol{x}} = \phi(\boldsymbol{x}, \xi_d; \boldsymbol{\beta}).$$

In addition to the discussion of D-optimal designs for the precise estimation of all the four parameters of model (4.31), Kupchak (2000) also considered D-optimal designs for the precise estimation of the parameters β_0 and β_I of model (4.34) assuming that both $ED_{50,1}$ and $ED_{50,2}$ are known. Finally, Kupchak (2000) briefly discussed robust designs to misspecification of the prior values of the parameter vector $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$. These were Bayesian D-optimal designs (see Chaloner and Lantz (1989), and Chaloner and Verdinelli (1995)) and minimax D-optimal designs (see Sitter (1992), and King and Wong (2000)). For the Bayesian approach, the uniform and the Gaussian distributions were used as priors for the parameters β_0, β_1, β_2 and β_{12}, and the number of support points increased as the prior became more diffuse.

4.6.2 Jia and Myers approach

Jia and Myers (2001) considered the construction of D-optimal designs for the two-variable binary logistic model with interaction as specified in (4.20) following similar arguments to those for the no interaction case. The following is a summary of their ideas.

For fixed logit u, the equation

$$u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 \tag{4.41}$$

with $\beta_{12} \neq 0$ is that of a hyperbola with orthogonal asymptotes $x_1 = -\frac{\beta_2}{\beta_{12}}$ and $x_2 = -\frac{\beta_1}{\beta_{12}}$. The variables x_1 and x_2 can be taken as doses or log-doses of two drugs, say A and B, respectively. As for the case of no interaction, it is assumed that $\beta_0 < 0$, $\beta_1 > 0$ and $\beta_2 > 0$. For each value of u the graph defined by equation (4.41) has two branches which are symmetric with respect to the point $(x_{01}, x_{02}) = \left(-\frac{\beta_2}{\beta_{12}}, -\frac{\beta_1}{\beta_{12}}\right)$ in the (x_1, x_2)-space. Using the transformation

$$\begin{cases} z_1 &= x_1 - x_{01} \\ z_2 &= x_2 - x_{02}, \end{cases} \tag{4.42}$$

the equation (4.41) can be rewritten as

$$u = \beta_0^* + \beta_{12} z_1 z_2 \tag{4.43}$$

or equivalently as

$$z_2 = \frac{u - \beta_0^*}{\beta_{12} z_1} \tag{4.44}$$

where $\beta_0^* = \beta_0 - \dfrac{\beta_1 \beta_2}{\beta_{12}}$. Thus for a given value of u, the hyperbola (4.44) has orthogonal asymptotes $z_1 = 0$ and $z_2 = 0$ in the (z_1, z_2)-space, and hence its graph has two symmetric branches with respect to the point $(0, 0)$.

If $\beta_0^* = u$ the hyperbola (4.44) degenerates into the asymptotes $z_1 = 0$ and $z_2 = 0$ or, equivalently the hyperbola (4.41) degenerates into the asymptotes $x_1 = -\frac{\beta_2}{\beta_{12}}$ and $x_2 = -\frac{\beta_1}{\beta_{12}}$. For $u \neq \beta_0^*$, Jia and Myers (2001) suggested that the D-optimal design for model (4.41) or equivalently for model (4.43) has 4 support points located on the lines of two constant response probability levels p_1 and p_2 with $p_1 < p_2$ or equivalently logits $u_1 < u_2$ where $u_i = \text{logit}(p_i)$ for $i = 1, 2$. Specifically, one support point lies on each of the two branches of the constant u_1 hyperbola, and likewise on each branch of the constant u_2 hyperbola. In addition, Jia and Myers (2001) chose the 4 support points of the candidate D-optimal design as vertices of a parallelogram, and proposed the following three design cases.

1. $u_1 < \beta_0^* < u_2$: In this case, the 4 branches of the u_1 and u_2 hyperbolae are in 4 different quadrants of the (z_1, z_2)-space. In other words, the support points of the candidate D-optimal design are spread over the four quadrants of the (z_1, z_2)-space. Figure 4.7 (a) illustrates the case when $\beta_{12} < 0$, i.e. $x_{01} = -\frac{\beta_2}{\beta_{12}} > 0$ and $x_{02} = -\frac{\beta_1}{\beta_{12}} > 0$, and Figure 4.7 (b) illustrates the case when $\beta_{12} > 0$, i.e. $x_{01} = -\frac{\beta_2}{\beta_{12}} < 0$ and $x_{02} = -\frac{\beta_1}{\beta_{12}} < 0$.

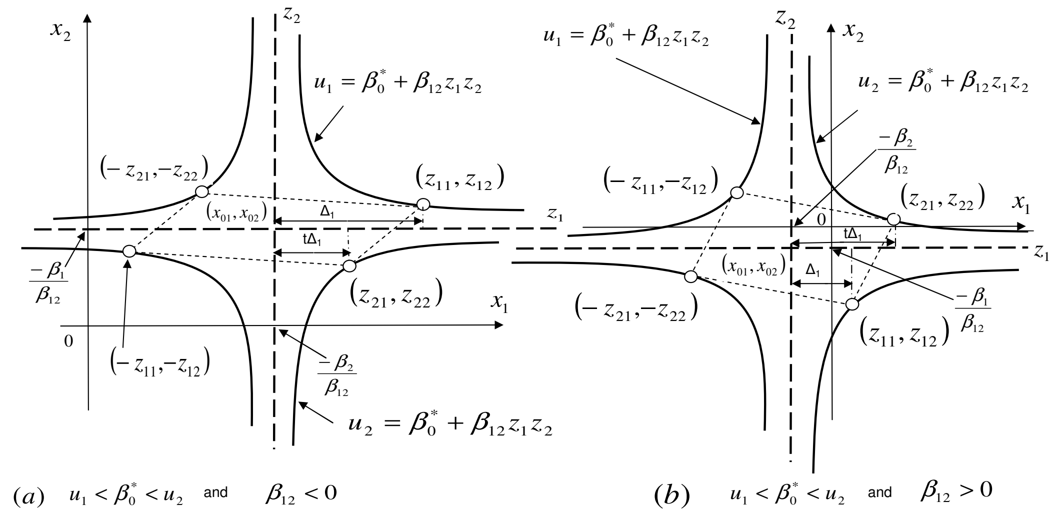


Figure 4.7: The hyperbola-based design for the two-variable binary logistic model with interaction $u = \beta_0^* + \beta_{12} z_1 z_2$ for $u_1 < \beta_0^* < u_2$ and (a) $\beta_{12} < 0$ and (b) $\beta_{12} > 0$: the candidate support points (z_{11}, z_{12}), (z_{21}, z_{22}), $(-z_{11}, -z_{12})$ and $(-z_{21}, -z_{22})$ are represented by circles.

2. $\beta_0^* < u_1 < u_2$: In this case, the 4 branches of the u_1 and u_2 hyperbolae are located in only two quadrants of the (z_1, z_2)-space. If $\beta_{12} < 0$, the branches of the u_1 and u_2 hyperbolae, and hence the support points of the candidate D-optimal design, are in the second and the fourth quadrants of the (z_1, z_2)-space as indicated in Figure 4.8 (a). If $\beta_{12} > 0$ the branches of the u_1 and u_2 hyperbolae, and hence the support points of the candidate D-optimal design are in the first and the third quadrants of the (z_1, z_2)-space as shown in Figure 4.8 (b).

3. $u_1 < u_2 < \beta_0^*$: As in case 2 the 4 branches of the u_1 and u_2 hyperbolae, and hence the support points of the candidate D-optimal design, are in only two quadrants of the (z_1, z_2)-space, but the order is reversed. That is if $\beta_{12} < 0$ the branches of the u_1 and u_2 hyperbolae, and hence the support points of the candidate D-optimal design are in the first and the third quadrants of the (z_1, z_2)-space as shown in Figure 4.9 (a). If $\beta_{12} > 0$, the branches of the u_1 and u_2 hyperbolae, and hence the support points of the candidate D-optimal design are in the second and the fourth quadrants of the (z_1, z_2)-space as shown in Figure 4.9 (b).

In constructing the D-optimal designs, Jia and Myers (2001) assumed equal weights for support points located on branches of the same hyperbola (u_1 or u_2) in each of the Figures 4.7 to 4.9. Now, letting $z_{11} = \Delta_1 > 0$ and $z_{21} = t\Delta_1$ for some $t \in (0, \infty)$, it follows from (4.44) that

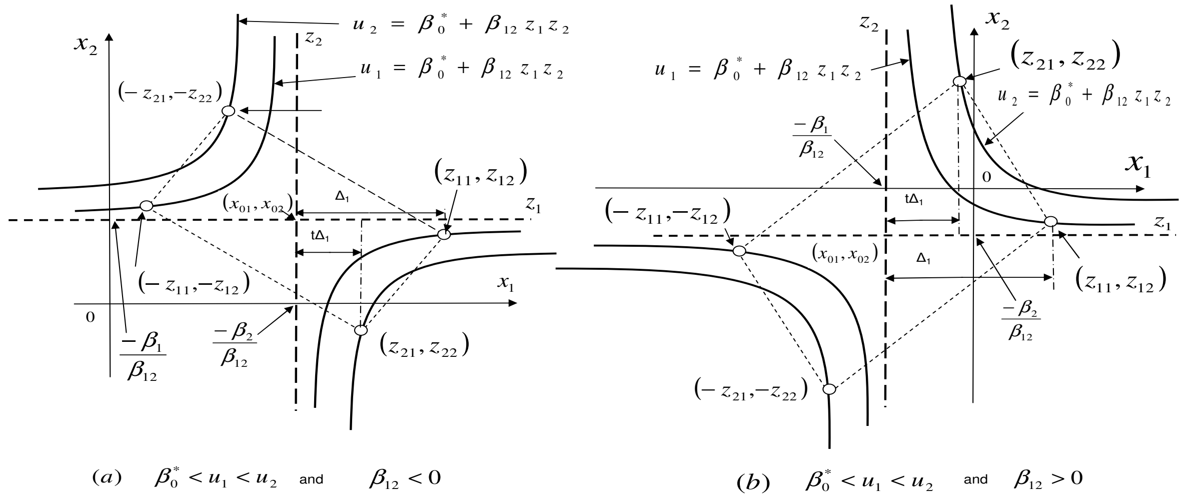


Figure 4.8: The hyperbola-based design for the two-variable binary logistic model with interaction $u = \beta_0^* + \beta_{12}z_1z_2$ for $\beta_0^* < u_1 < u_2$ and (a) $\beta_{12} < 0$ and (b) $\beta_{12} > 0$: the candidate support points (z_{11}, z_{12}), (z_{21}, z_{22}), $(-z_{11}, -z_{12})$ and $(-z_{21}, -z_{22})$ are represented by circles.

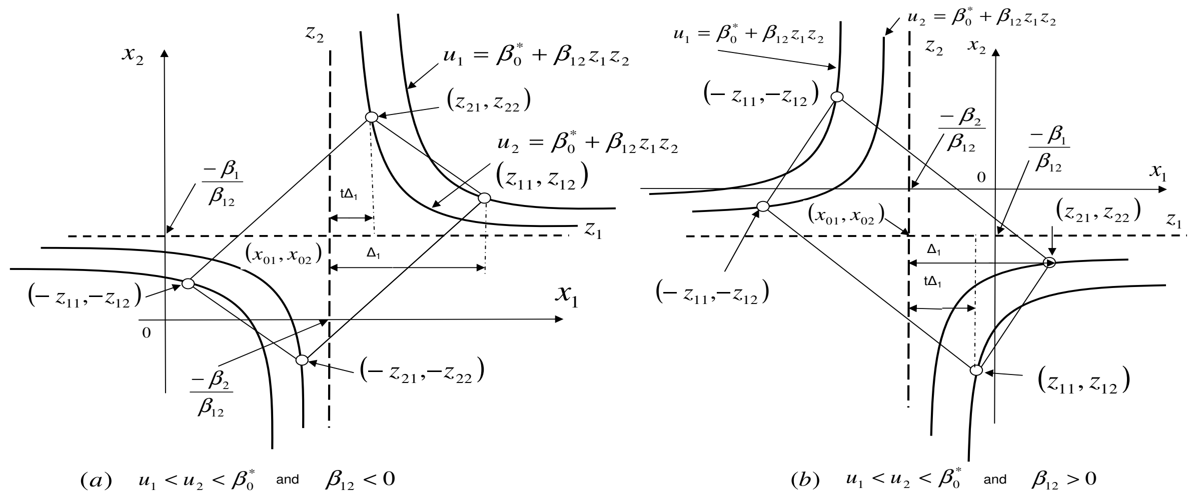


Figure 4.9: The hyperbola-based design for the two-variable binary logistic model with interaction $u = \beta_0^* + \beta_{12}z_1z_2$ for $u_1 < u_2 < \beta_0^*$ and (a) $\beta_{12} < 0$ and (b) $\beta_{12} > 0$: the candidate support points (z_{11}, z_{12}), (z_{21}, z_{22}), $(-z_{11}, -z_{12})$ and $(-z_{21}, -z_{22})$ are represented by circles.

$z_{12} = \frac{u_1-\beta_0^*}{\beta_{12}\Delta_1}$ and $z_{22} = \frac{u_2-\beta_0^*}{t\beta_{12}\Delta_1}$, and that $(-z_{11}, -z_{12}) = (-\Delta_1, -\frac{u_1-\beta_0^*}{\beta_{12}\Delta_1})$ and $(-z_{21}, -z_{22}) =$

$(-t\Delta_1, -\frac{u_2-\beta_0^*}{t\beta_{12}\Delta_1})$. Hence, the approximate design corresponding to Figures 4.7 to 4.9 is

$$\xi = \left\{ \begin{array}{cccc} \left(\Delta_1, \dfrac{u_1-\beta_0^*}{\beta_{12}\Delta_1}\right) & \left(t\Delta_1, \dfrac{u_2-\beta_0^*}{t\beta_{12}\Delta_1}\right) & \left(-\Delta_1, -\dfrac{u_1-\beta_0^*}{\beta_{12}\Delta_1}\right) & \left(-t\Delta_1, -\dfrac{u_2-\beta_0^*}{t\beta_{12}\Delta_1}\right) \\ \frac{\lambda}{2} & \frac{1-\lambda}{2} & \frac{\lambda}{2} & \frac{1-\lambda}{2} \end{array} \right\} \tag{4.45}$$

on the (z_1, z_2)-space. Jia and Myers (2001) used the transformation (4.42) to back transform the coordinates of the support points of the above design from the (z_1, z_2)-space to the (x_1, x_2)-space.

The information matrix associated with the design obtained by back transforming design (4.45) to the (x_1, x_2)-space is intricate and lengthy and is not reported here. However, its determinant is given compactly by

$$D_t = \frac{\lambda^2(1-\lambda)^2 e^{2u_1+2u_2}(u_2-u_1)^2[t(u_2-\beta_0^*) - \frac{1}{t}(u_1-\beta_0^*)]^2}{\beta_{12}^4(1+e^{u_1})^4(1+e^{u_2})^4}. \tag{4.46}$$

Clearly, as highlighted by Jia and Myers (2001), D_t approaches infinity as $t \to 0^+$ or $t \to \infty$. Therefore, D_t has a finite value only if $0 < t < \infty$. The magnitude of t determines the size of the design space and hence the magnitude of the determinant D_t. Giving a value to t can be interpreted as imposing a restriction on the design space. Jia and Myers (2001) state that in many practical situations t can be taken as 1, and for $t = 1$ the determinant (4.46) becomes

$$D_1 = \frac{\lambda^2(1-\lambda)^2 e^{2u_1+2u_2}(u_2-u_1)^4}{\beta_{12}^4(1+e^{u_1})^4(1+e^{u_2})^4}. \tag{4.47}$$

which is maximized by $\lambda = \frac{1}{2}$, and is symmetric in u_1 and u_2. Symmetry arguments imply that $u_1 = -u_2 = u$, and hence that the determinant (4.47) is proportional to

$$D = \frac{u^4 e^{4u}}{(1+e^u)^8} \tag{4.48}$$

which is the square of the determinant (4.13) in Section 4.3. This means that the determinant (4.48) also has maxima at $u = \pm 1.5434$.

Comments

The approach of Jia and Myers (2001) for constructing D-optimal designs for the two-variable binary logistic model with interaction is based on relatively easy calculations since the candidate D-optimal designs have 4 support points, and are therefore equally weighted 4-point designs (see Silvey (1980, p. 42). In addition, the fact that the support points of the candidate

D-optimal designs are located on the two constant logits $\pm u$ further simplifies the calculations of the D-optimal designs for a given best guess of the parameters β_0, β_1, β_2 and β_{12}. However, the approach of Jia and Myers (2001) has the following drawbacks.

1. The reason for using the special case of $t = 1$ in the expression (4.46) has the effect of removing the dependency of the determinant D_t on β_0^*, and hence results in "loss of generality". With such a restriction, it appears that Jia and Myers (2001) searched for D-optimal designs in a limited class of designs. In addition, the dependency of the location of the support points of the candidate D-optimal design on the value of Δ_1 increases the degree of uncertainty about the design since in addition to guessing the values of the parameters, the value of Δ_1 has also to be provided.

2. Jia and Myers (2001) do not prove the D-optimality of their design.

The geometric approach of Jia and Myers (2001) for the two-variable binary logistic model with interaction will be used in a more elaborate way in Chapter 7 for constructing D-optimal designs in a restricted design space.

5

D-optimal Designs for the Two-Variable Binary Logistic Model without Interaction: Theoretical Results

5.1. Introduction

This chapter introduces a new design approach for the construction of D-optimal designs for the two-variable binary logistic model without interaction and draws together the approaches of Sitter and Torsney (1995a), Atkinson and Haines (1996) and Jia and Myers (2001). Specifically, the approach takes into account the fact that for generalized linear models, for which the binary logistic model is a special case, the D-optimal design depends on the parameter vector and the design space. The chapter contains the following sections. Section 5.2 gives a proof of D-optimality of the equally weighted 4-point design introduced by Sitter and Torsney (1995a) for the design locus G_w given in (4.23) and by Jia and Myers (2001). In the proof, the design space is assumed to be the log-dose space. Section 5.3 discusses the construction of new D-optimal designs, defined on the actual dose space, termed trapezium designs. The designs obtained are 4-point or 3-point D-optimal designs depending on the value of the intercept parameter β_0 in the logit $u = \beta_0 + \beta_1 x_1 + \beta_2 x_2$. Section 5.4 contains the conclusions for this chapter.

5.2. D-optimal designs with parallel line restrictions.

Section 4.5 described the approaches of Sitter and Torsney (1995a), Atkinson and Haines (1996), and Jia and Myers (2001) to constructing D-optimal designs for the two-variable binary logistic model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ defined in Section 4.4.1. To obtain a restriction on the design space, Sitter and Torsney (1995a) used the constraint $-1 \leq u_2 \leq 1$ where $u_2 = b_0 + b_1 x_1 + b_2 x_2$ is a straight line in (x_1, x_2)-space for a given value of u_2. Setting $u_2 = \pm 1$ gives the same constraints as the two parallel line restrictions $b_{01} + b_1 x_1 + b_2 x_2 = 0$ and $b_{02} + b_1 x_1 + b_2 x_2 = 0$ implicitly used by Jia and Myers (2001). Note that the restriction $-1 \leq u_2 \leq 1$ can be generalized to $-b \leq u_2 \leq b$. For logit $u_1 = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ unrestricted, the candidate D-optimal design in both approaches was an equally weighted 4-point design with support points located on the intersection of $u_2 = \pm b$ and two logit lines which are symmetric about the logit line $u_1 = 0$, and thus the candidate support points are of the form $(-u_1, -b)$, $(u_1, -b)$, $(-u_1, b)$ and (u_1, b) in (u_1, u_2)-space. Figure 5.1, where circles represent the support points, is a graphical representation of the candidate 4-point D-optimal design of Sitter and Torsney (1995a). The approach of Atkinson and Haines (1996) leads to the same

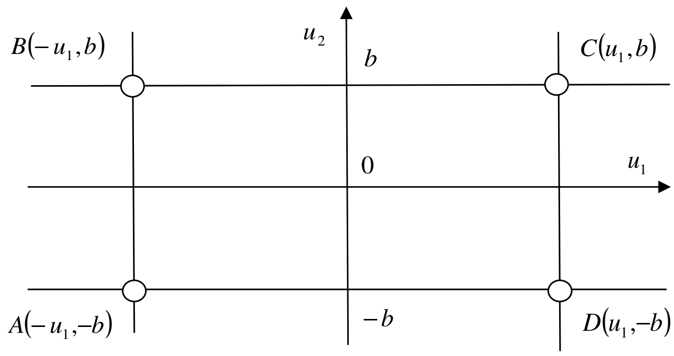


Figure 5.1: Candidate 4-point D-optimal design of Sitter and Torsney (1995a) in the (u_1, u_2)-design space $\mathbb{R} \times [-b, b]$ where b is a fixed real number. Circles are support points.

conclusion for certain values of the parameters β_0, β_1 and β_2 (see Section 4.5.2, Table 4.3 for $\boldsymbol{\beta} = (0, 3, 1)^T$, and item (3) in Section 4.5.4). The information matrix on a per point basis evaluated at the 4-point design

$$\xi_u = \left\{ \begin{array}{cccc} \text{A} & \text{D} & \text{B} & \text{C} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array} \right\} = \left\{ \begin{array}{cccc} (-u_1, -b) & (u_1, -b) & (-u_1, b) & (u_1, b) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array} \right\} \tag{5.1}$$

is

$$M(\xi_u) = \psi(u_1)\begin{bmatrix} 1 & 0 & 0 \\ 0 & u_1^2 & 0 \\ 0 & 0 & b^2 \end{bmatrix},$$

and its determinant is

$$D = [\psi(u_1)]^3\, u_1^2 b^2,$$

where $\psi(u_1) = \dfrac{e^{u_1}}{(1+e^{u_1})^2}$. The determinant D is maximized by $\pm u^*$ where $u^* = 1.22291$ is the positive solution for u_1 to the equation

$$2 + 2e^{u_1} + 3u_1 - 3u_1 e^{u_1} = 0. \tag{5.2}$$

Thus, the design (5.1) is

$$\xi_u^* = \left\{ \begin{array}{cccc} (-1.22291,-b) & (1.22291,-b) & (-1.22291,b) & (1.22291,b) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array} \right\}$$

and can be taken as a candidate D-optimal design. Hence, the following theorem.

Theorem 5.1. For the two-variable binary logistic model $u_1 = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ where p is the probability of positive response with u_1 unrestricted in $\mathbb{R}$, and for the restriction $-b \leq u_2 = b_0 + b_1 x_1 + b_2 x_2 \leq b$ where $\dfrac{b_1}{b_2} \neq \dfrac{\beta_1}{\beta_2}$ and $b > 0$, the design

$$\xi_u^* = \left\{ \begin{array}{cccc} (-u^*,-b) & (u^*,-b) & (-u^*,b) & (u^*,b) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array} \right\}, \tag{5.3}$$

where $u^* = 1.22291$, is D-optimal on the (u_1, u_2)-space.

Proof

Let

$$M(\boldsymbol{u}) = \psi(u_1)\begin{bmatrix} 1 \\ u_1 \\ u_2 \end{bmatrix}\begin{bmatrix} 1 & u_1 & u_2 \end{bmatrix}$$

be the information matrix for $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^T$ evaluated at a one-point design with an arbitrary support point $\boldsymbol{u} = (u_1, u_2)^T \in \mathbb{R} \times [-b, b]$. The D-optimality of design (5.3) is proved using the Equivalence Theorem for the D-optimality criterion (3.3) by showing that the directional derivative function

$$\phi(\boldsymbol{u}, \xi_u^*) = 3 - \text{tr}\left[M^{-1}(\xi_u^*)M(\boldsymbol{u})\right] \geq 0$$

for all $\boldsymbol{u} = (u_1, u_2)^T \in \mathbb{R} \times [-b, b]$ with equality holding at the support points of the design ξ_u^*, or equivalently, the standardized variance function

$$d(\boldsymbol{u}, \xi_u^*) = \operatorname{tr}\left[M^{-1}(\xi_u^*)M(\boldsymbol{u})\right] \leq 3$$

for all $\boldsymbol{u} = (u_1, u_2)^T \in \mathbb{R} \times [-b, b]$ with equality holding at the support points of the design ξ_u^*.

The standardized variance function $d(\boldsymbol{u}, \xi_u^*)$ is given by

$$d(\boldsymbol{u}, \xi_u^*) = \frac{\psi(u_1)}{\psi(u^*)}\left\{1 + \frac{u_1^2}{u^{*2}} + \frac{u_2^2}{b^2}\right\} = \frac{e^{-u^*+u_1}(1+e^{u^*})^2}{(1+e^{u_1})^2}\left\{1 + \frac{u_1^2}{u^{*2}} + \frac{u_2^2}{b^2}\right\} \tag{5.4}$$

where $u^* = 1.22291$. Clearly, if $u_1 = \pm u^*$ and $u_2 = \pm b$, then $d(\boldsymbol{u}, \xi_u^*) = 3$ at the support points of design (5.3) as required. It remains to be shown that $d(\boldsymbol{u}, \xi_u^*) \leq 3$ at any other point $\boldsymbol{u} = (u_1, u_2)^T \in \mathbb{R} \times [-b, b]$.

Applying the fact that $-b \leq u_2 \leq b \Rightarrow u_2^2 \leq b^2$ to (5.4) provides the inequality

$$d(\boldsymbol{u}, \xi_u^*) \leq g(u_1) = \frac{\psi(u_1)}{\psi(u^*)}\left\{2 + \frac{u_1^2}{u^{*2}}\right\} = \frac{e^{-u^*+u_1}(1+e^{u^*})^2(2u^{*^2} + u_1^2)}{u^{*^2}(1+e^{u_1})^2}.$$

It has to be shown that $g(u_1) \leq 3$ for all $u_1 \in \mathbb{R}$. Note that $g(u_1)$ is an even function in u_1 since $g(-u_1) = g(u_1)$, and thus $g(u_1)$ is symmetric about $u_1 = 0$. Hence, $g(u_1)$ is to be investigated on $[0, \infty)$ since the conclusions will also apply to $g(u_1)$ on $(-\infty, 0]$.

The stationary points of $g(u_1)$ on $[0, \infty)$ are solutions for u_1 to the equation

$$g'(u_1) = \frac{dg(u_1)}{du_1} = \frac{e^{-u^*+u_1}(1+e^{u^*})^2[(2u^{*^2} + u_1^2 + 2u_1) - (2u^{*^2} + u_1^2 - 2u_1)e^{u_1}]}{u^{*^2}(1+e^{u_1})^3} = 0,$$

or equivalently, the equation

$$h(u_1) = (2u^{*^2} + u_1^2 + 2u_1) - (2u^{*^2} + u_1^2 - 2u_1)e^{u_1} = 0. \tag{5.5}$$

Note that $g'(0) = 0$ and $g'(u^*) = \dfrac{2 + 2e^{u^*} + 3u^* - 3u^*e^{u^*}}{u^*(1+e^{u^*})} = 0$ since u^* is the solution of equation (5.2). Therefore $u_1 = 0$ and $u_1 = u^*$ are stationary points of $g(u_1)$, or equivalently, $u_1 = 0$ and $u_1 = u^*$ are solutions of equation (5.5). The second derivative of $g(u_1)$ with respect to u_1 is

$$g''(u_1) = \frac{e^{u_1-u^*}(1+e^{u^*})^2[(2 + 2u^{*^2} + u_1^2)(1+e^{2u_1}) + 4u_1(1-e^{2u_1}) + 4(1 - 2u^{*2} - u_1^2)e^{u_1}]}{u^{*2}(1+e^{u_1})^4}.$$

Hence, $g''(0) = \dfrac{e^{-u^*}(1+e^{u^*})^2(2-u^{*2})}{4u^{*2}} = 0.480 > 0$. Therefore $g(u_1)$ takes on a local minimum value at 0 and $g(0) = \dfrac{e^{-u^*}(1+e^{u^*})^2}{2} \simeq 2.85 < 3$. Also,

$$g''(u^*) = \frac{2 + 4e^{u^*} + 2e^{2u^*} + 4u^* - 4u^*e^{2u^*} + 3u^{*2} - 12u^{*2}e^{u^*} + 3u^{*2}e^{2u^*}}{u^{*2}(1+e^{u^*})^2} = -0.608 < 0$$

which means $g(u_1)$ takes on its local or global maximum value at u^*, and $g(u^*) = 3$. It has to be shown that $g(u^*) = 3$ is the unique global maximum value of $g(u_1)$ on $[0, \infty)$. Figure 5.2 suggests that $g(u^*) = 3$ is indeed the global maximum value of $g(u_1)$ on $[0, \infty)$.

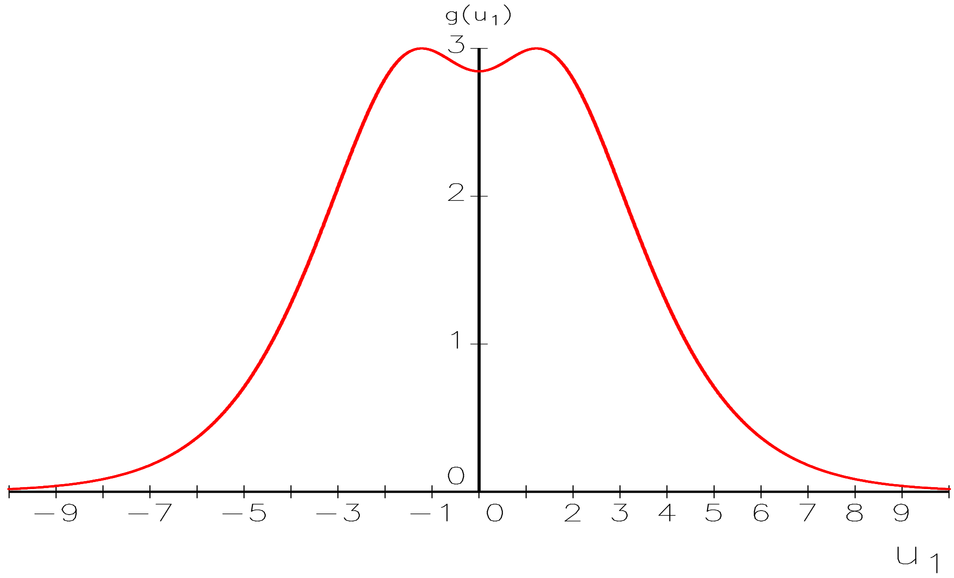


Figure 5.2: Plot of $g(u_1) = \dfrac{e^{u_1 - u^*}(1 + e^{u^*})^2(2u^{*2} + u_1^2)}{u^{*2}(1 + e^{u_1})^2}$ versus $u_1 \in \mathbb{R}$ with $u^* = 1.22291$.

To analytically prove that $g(u^*) = 3$ is the unique global maximum value of $g(u_1)$ on $[0, \infty)$, it has to be shown that $g'(u_1) \geq 0$ on $[0, u^*]$ and that $g'(u_1) \leq 0$ on $[u^*, \infty)$, or equivalently, that $h(u_1)$ given by (5.5) is nonnegative on $[0, u^*]$ and less than or equal to zero on $[u^*, \infty)$ since the sign of $g'(u_1)$ is determined by the sign of $h(u_1)$. Figure 5.3 (a) suggests that indeed, $h(u_1) \geq 0$ on $[0, u^*]$ and $h(u_1) \leq 0$ on $[u^*, \infty)$.

Note that $h(u_1) = h_1(u_1) - h_2(u_1)$, where $h_1(u_1) = u_1^2 + 2u_1 + 2u^{*2}$ is a parabola and $h_2(u_1) = (u_1^2 - 2u_1 + 2u^{*2})e^{u_1}$. Figure 5.3 (b) displays the graphs of $h_1(u_1)$ and $h_2(u_1)$. On $[0, \infty)$ both $h_1(u_1)$ and $h_2(u_1)$ are strictly increasing functions of u_1 since $h_1'(u_1) = 2(u_1 + 1) > 0$ and $h_2'(u_1) = (u_1^2 + 2u^{*2} - 2)e^{u_1} = (u_1^2 + 0.991)e^{u_1} > 0$. Clearly, $h_1(u_1)$ is a convex parabola on $[0, \infty)$ and also $h_2(u_1)$ is convex on $[0, \infty)$ since $h_2''(u_1) = (u_1^2 + 2u_1 + 0.991)e^{u_1} > 0$ for all $u_1 \in [0, \infty)$. This means that, on $[0, \infty)$, the graphs of $h_1(u_1)$ and $h_2(u_1)$ meet twice at most. Since 0 and u^* are solutions for u_1 to the equation $h(u_1) = 0$, as was shown above, it follows that $h_1(u_1)$ and $h_2(u_1)$ only meet at $u_1 = 0$ and $u_1 = u^*$. This and the fact that $\displaystyle\lim_{u_1 \to \infty} \frac{h_1(u_1)}{h_2(u_1)} = 0$ imply that $h(u_1) = h_1(u_1) - h_2(u_1) \leq 0$ on $[u^*, \infty)$, and that $h(u_1)$ must be greater than or equal to zero on $[0, u^*]$. ■

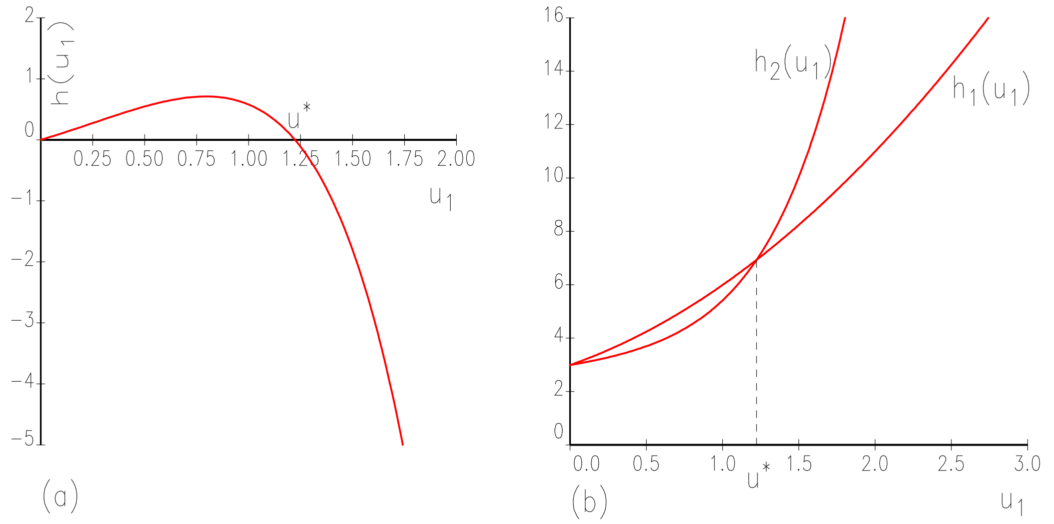


Figure 5.3: Plots of (a) $h(u_1) = 2u^{*^2} + u_1^2 + 2u_1 - (2u^{*^2} + u_1^2 - 2u_1)e^{u_1}$ versus $u_1 \geq 0$, and (b) $h_1(u_1) = 2u^{*^2} + u_1^2 + 2u_1$ and $h_2(u_1) = (2u^{*2} + u_1^2 - 2u_1)e^{u_1}$ versus $u_1 \geq 0$ with $u^* = 1.22291$.

5.3. D-optimal designs on actual dose design space

In this section D-optimal designs for the two-variable binary logistic model $u = \text{logit}(p) = \beta_0 + \beta_1 d_1 + \beta_2 d_2$ are constructed on the actual dose design space $\mathcal{D} = [0, \infty) \times [0, \infty)$. Specifically, it is shown that if $-\infty < \beta_0 < -1.5434$, then the D-optimal design has 4 support points located at the boundaries of the design space $\mathcal{D}$. This design is termed a trapezium design because its support points are vertices of a trapezium in $\mathcal{D}$. It is also shown that if $-1.5434 \leq \beta_0 \leq 0$, then the trapezium design degenerates to a 3-point D-optimal design, and if $\beta_0 \to -\infty$, then the trapezium design degenerates to the equally weighted 4-point D-optimal design as in Sitter and Torsney (1995a), Atkinson and Haines (1996), and Jia and Myers (2001).

5.3.1 The design problem in canonical form

For actual doses d_1 and d_2 of two drugs under study, model (4.18) becomes

$$u = \text{logit}(p) = \beta_0 + \beta_1 d_1 + \beta_2 d_2, \tag{5.6}$$

or equivalently,

$$u = \text{logit}(p) = \beta_0 + z_1 + z_2 \tag{5.7}$$

where $z_1 = \beta_1 d_1 \geq 0$ and $z_2 = \beta_2 d_2 \geq 0$ assuming that $\beta_0 < 0$, $\beta_1 > 0$ and $\beta_2 > 0$ as in Chapter 2, Section 2.4.4. The information matrix for the parameter vector $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^T$ evaluated at the 1-point design $\boldsymbol{d} = (d_1, d_2)^T$ for model (5.6) is $M(\boldsymbol{d}; \boldsymbol{\beta})$ given by the information matrix (4.3) evaluated at $\boldsymbol{x}_i = \boldsymbol{d}_i$ for $i = 1, 2, \ldots, d \geq 3$. Similarly evaluating the information matrix (4.3) at the 1-point design $\boldsymbol{z} = (z_1, z_2)^T = (\beta_1 d_1, \beta_2 d_2)^T$ for model (5.7) gives the information matrix

$$M(\boldsymbol{z}; \beta_0) = BM(\boldsymbol{d}; \boldsymbol{\beta})B^T = \frac{e^u}{(1+e^u)^2} \widetilde{\boldsymbol{z}}\widetilde{\boldsymbol{z}}^T \tag{5.8}$$

where $\widetilde{\boldsymbol{z}} = (1, \boldsymbol{z}^T)^T$ and $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \beta_1 & 0 \\ 0 & 0 & \beta_2 \end{bmatrix}$ is a nonsingular matrix since $|B| = \beta_1 \beta_2 > 0$.

A design ξ for model (5.6) is design (4.5) with the $\boldsymbol{x}_i$ replaced by $\boldsymbol{d}_i = (d_{i1}, d_{i2})^T$ for $i = 1, 2, \ldots, d$. The transformation of model (5.6) to (5.7) transforms the design ξ to

$$\xi_z = \left\{ \begin{array}{cccc} \boldsymbol{z}_1 & \boldsymbol{z}_2 & \ldots & \boldsymbol{z}_d \\ \lambda_1 & \lambda_2 & \ldots & \lambda_d \end{array} \right\} \tag{5.9}$$

where $\boldsymbol{z}_i = (z_{i1}, z_{i2})^T = (\beta_1 d_{i1}, \beta_2 d_{i2})^T$ with associated weights λ_i such that $\lambda_i > 0$ and $\sum_{i=1}^{d} \lambda_i = 1$ for $i = 1, 2, \ldots, d$. Note that under the assumption that $\beta_0 < 0$, $\beta_1 > 0$ and $\beta_2 > 0$, designs ξ and ξ_z are defined on the same design space $\mathcal{D} = [0, \infty) \times [0, \infty)$.

Let $M(\xi_z; \beta_0) = \sum_{i=1}^{d} \lambda_i M(\boldsymbol{z}_i; \beta_0)$ and $M(\xi; \boldsymbol{\beta}) = \sum_{i=1}^{d} \lambda_i M(\boldsymbol{d}_i; \boldsymbol{\beta})$. Then design ξ_z is D-optimal with respect to maximizing $|M(\xi_z; \beta_0)|$ if and only if design ξ is D-optimal with respect to maximizing $|M(\xi; \boldsymbol{\beta})|$. In fact, note that $|M(\xi_z; \beta_0)| = |BM(\xi; \boldsymbol{\beta})B^T| = |B|^2 |M(\xi; \boldsymbol{\beta})|$ where B is the matrix defined in (5.8). This implies that maximizing $|M(\xi; \boldsymbol{\beta})|$ is equivalent to maximizing $|M(\xi_z, \beta_0)|$. Furthermore, the directional derivative function for D-optimality associated with design ξ_z for model (5.7) is given by

$$\begin{aligned}
\phi(\boldsymbol{z}, \xi_z; \beta_0) &= 3 - \operatorname{tr}\left\{M^{-1}(\xi_z; \beta_0) M(\boldsymbol{z}; \beta_0)\right\} \\
&= 3 - \operatorname{tr}\left\{\left[B^{-1} M^{-1}(\xi; \boldsymbol{\beta}) B^{-1}\right] BM(\boldsymbol{d}; \boldsymbol{\beta})B\right\}, \text{ since } M(\boldsymbol{z}; \beta_0) = BM(\boldsymbol{d}; \boldsymbol{\beta})B \\
&= 3 - \operatorname{tr}\left\{M^{-1}(\xi; \boldsymbol{\beta}) M(\boldsymbol{d}; \boldsymbol{\beta})\right\} \\
&= \phi(\boldsymbol{d}, \xi; \boldsymbol{\beta})
\end{aligned}$$

which is the directional derivative function associated with design ξ. The equality of the directional derivative functions implies that design ξ_z is D-optimal with respect to maximizing $|M(\xi_z, \beta_0)|$ if and only if design ξ is D-optimal with respect to maximizing $|M(\xi; \boldsymbol{\beta})|$.

5.3.2 4-point D-optimal trapezium design

From the literature, as discussed in Section 4.5, it can be assumed that the support points of the candidate D-optimal design for the two-variable binary logistic model without interaction lie on two logit lines which are symmetric about $u = 0$. Furthermore, the support points located on the same logit line are equally weighted (see for example Sitter and Torsney (1995a), and Jia and Myers (2001)). In addition, it has been reported that the determinant of the information matrix for the parameter vector $\boldsymbol{\beta}$ is an increasing function of the size of the design space (e.g. see Jia and Myers (2001), and Myers et al. (2002, p. 240)). Hence, it can be conjectured that the support points of a candidate 4-point D-optimal design, defined on $\mathcal{D} = [0, \infty) \times [0, \infty)$, are located at the intersection of the logit lines $\pm u$ and the boundaries of the design space $\mathcal{D}$ as illustrated in Figure 5.4. Thus, under the above assumptions, the

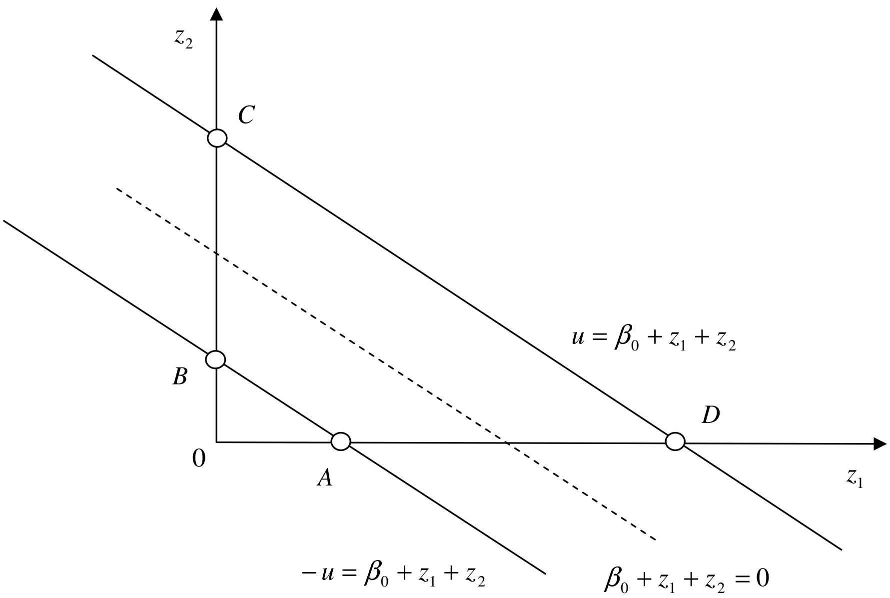


Figure 5.4: The pattern of a 4-point candidate D-optimal design for a two-variable binary logistic model without interaction $u = \beta_0 + z_1 + z_2$ on the design space $\mathcal{D} = [0, \infty) \times [0, \infty)$: A, B, C and D are support points.

candidate 4-point D-optimal design in (z_1, z_2)-space is given by

$$\xi_z = \left\{ \begin{array}{cccc} A & B & C & D \\ \frac{1-\lambda}{2} & \frac{1-\lambda}{2} & \frac{\lambda}{2} & \frac{\lambda}{2} \end{array} \right\} = \left\{ \begin{array}{cccc} (-u-\beta_0, 0) & (0, -u-\beta_0) & (0, u-\beta_0) & (u-\beta_0, 0) \\ \frac{1-\lambda}{2} & \frac{1-\lambda}{2} & \frac{\lambda}{2} & \frac{\lambda}{2} \end{array} \right\}$$

with $0 < \lambda < 1$ and $0 \leq u < -\beta_0$. The following theorem gives conditions on β_0, u and λ under which the design ξ_z given above is D-optimal.

Theorem 5.2. The respective optimal values of u and λ in the candidate D-optimal design of the form

$$\xi_z = \left\{ \begin{array}{cccc} (-u-\beta_0, 0) & (0, -u-\beta_0) & (0, u-\beta_0) & (u-\beta_0, 0) \\ \frac{1-\lambda}{2} & \frac{1-\lambda}{2} & \frac{\lambda}{2} & \frac{\lambda}{2} \end{array} \right\} \tag{5.10}$$

for the two-variable binary logistic model $u = \text{logit}(p) = \beta_0 + z_1 + z_2$ on the design space $[0, \infty) \times [0, \infty)$ with parameter $\beta_0 < -1.5434$ are u^* and λ^*, solutions for u and λ to the simultaneous equations

$$3(1 + u + e^u - ue^u)[(u + \beta_0)^2 - 4\lambda\beta_0 u] - (\beta_0^2 - u^2)(1 + e^u) = 0$$

and

$$(1 - 2\lambda)(u + \beta_0)^2 - 4\lambda(2 - 3\lambda)\beta_0 u = 0.$$

Furthermore, $u^* \in (1.22291; 1.5434)$ and $\lambda^* = \dfrac{\beta_0^2 + 6\beta_0 u^* + u^{*2} - \sqrt{\beta_0^4 + 14\beta_0^2 u^{*2} + u^{*4}}}{12\beta_0 u^*} \in \left(\dfrac{1}{2}, \dfrac{2}{3}\right)$.

Proof

Let $M(\xi_z; \beta_0)$ be the information of the vector of parameters $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^T$ evaluated at design ξ_z given in (5.10). Then,

$$M(\xi_z; \beta_0) = \frac{1}{2}\psi(u) \begin{bmatrix} 2 & -u-\beta_0+2\lambda u & -u-\beta_0+2\lambda u \\ -u-\beta_0+2\lambda u & (u+\beta_0)^2 - 4\lambda\beta_0 u & 0 \\ -u-\beta_0+2\lambda u & 0 & (u+\beta_0)^2 - 4\lambda\beta_0 u \end{bmatrix} \tag{5.11}$$

and the associated determinant is

$$D = \frac{(1-\lambda)\lambda u^2 e^{3u}[(u+\beta_0)^2 - 4\lambda\beta_0 u]}{(1+e^u)^6}. \tag{5.12}$$

The value and the interval for λ in the candidate D-optimal design ξ_z of the form (5.10) is calculated as follows. Differentiating D, in (5.12), with respect to λ gives

$$\frac{\partial D}{\partial \lambda} = \frac{u^2 e^{3u}\left\{(1-2\lambda)(u+\beta_0)^2 - 4\lambda(2-3\lambda)\beta_0 u\right\}}{(1+e^u)^6}.$$

Setting $\dfrac{\partial D}{\partial \lambda} = 0$ and solving for λ gives

$$(1 - 2\lambda)(u + \beta_0)^2 - 4\lambda(2 - 3\lambda)\beta_0 u = 0 \tag{5.13}$$

which is a quadratic equation in λ, and hence has at most two solutions for λ. Solving the equation (5.13) for λ gives

$$\lambda = \lambda^* = \frac{\beta_0^2 + 6\beta_0 u + u^2 - \sqrt{\beta_0^4 + 14\beta_0^2 u^2 + u^4}}{12\beta_0 u} \tag{5.14}$$

or

$$\lambda = \lambda^{**} = \frac{\beta_0^2 + 6\beta_0 u + u^2 + \sqrt{\beta_0^4 + 14\beta_0^2 u^2 + u^4}}{12\beta_0 u}. \tag{5.15}$$

where $\beta_0 < 0$ and $u > 0$. An admissible solution for λ in (5.13) must satisfy the condition $0 < \lambda < 1$.

Consider $\lambda = \lambda^{**}$. Now

$$\begin{aligned}\beta_0^4 + 14\beta_0^2 u^2 + u^4 &= (\beta_0^2 + 6\beta_0 u + u^2)^2 - 12\beta_0 u(u + \beta_0)^2 \\ &= (\beta_0^2 + 6\beta_0 u + u^2)^2 + 12|\beta_0| u(u + \beta_0)^2 \text{ since } \beta_0 < 0\end{aligned}$$

which implies that

$$\sqrt{\beta_0^4 + 14\beta_0^2 u^2 + u^4} \geq \sqrt{(\beta_0^2 + 6\beta_0 u + u^2)^2} = |\beta_0^2 + 6\beta_0 u + u^2| \geq \beta_0^2 + 6\beta_0 u + u^2, \tag{5.16}$$

and thus that

$$\beta_0^2 + 6\beta_0 u + u^2 + \sqrt{\beta_0^4 + 14\beta_0^2 u^2 + u^4} \geq \beta_0^2 + 6\beta_0 u + u^2 + |\beta_0^2 + 6\beta_0 u + u^2| \geq 0.$$

Hence, $\lambda^{**} \leq 0$ since in the expression (5.15) for λ^{**}, the numerator is nonnegative and the denominator $12\beta_0 u$ is negative. Thus, λ^{**} is not the required solution.

Consider $\lambda = \lambda^*$. It follows from (5.16) that $\beta_0^2 + 6\beta_0 u + u^2 - \sqrt{\beta_0^4 + 14\beta_0^2 u^2 + u^4} \leq 0$, and hence $\lambda^* \geq 0$ since in the expression (5.14) for λ^*, the numerator is not positive and the denominator $12\beta_0 u$ is negative. Furthermore,

$$\begin{aligned}(\beta_0^2 - 6\beta_0 u + u^2)^2 &= \beta_0^4 + 14\beta_0^2 u^2 + u^4 - 12\beta_0 u(u - \beta_0)^2 \\ &\geq \beta_0^4 + 14\beta_0^2 u^2 + u^4 \text{ since } \beta_0 < 0\end{aligned}$$

which implies

$$\sqrt{\beta_0^4 + 14\beta_0^2 u^2 + u^4} \leq |\beta_0^2 - 6\beta_0 u + u^2| = \beta_0^2 - 6\beta_0 u + u^2 \text{ since } \beta_0 < 0$$

and

$$12\beta_0 u = \beta_0^2 + 6\beta_0 u + u^2 - (\beta_0^2 - 6\beta_0 u + u^2) \leq \beta_0^2 + 6\beta_0 u + u^2 - \sqrt{\beta_0^4 + 14\beta_0^2 u^2 + u^4} \leq 0.$$

Hence $\lambda^* \leq 1$ since in the expression (5.14) for λ^*, the numerator and the denominator are negative with the absolute value of the numerator less than or equal to that of the denominator. Thus, for λ^* evaluated at $u = u^*$, the candidate D-optimal value of u, is the required candidate D-optimal value of λ.

For now, note that equation (5.13) can be written as

$$\frac{(u+\beta_0)^2}{4\lambda\beta_0 u} = \frac{2-3\lambda}{1-2\lambda}. \tag{5.17}$$

Since $(u+\beta_0)^2 > 0$, $4\beta_0 u \leq 0$ and $\lambda > 0$, it follows from (5.17) that $\frac{2-3\lambda}{1-2\lambda} < 0$ which implies that $\frac{1}{2} < \lambda < \frac{2}{3}$. Hence, the candidate D-optimal value of λ is λ^* given in (5.14) and $\lambda^* \in \left(\frac{1}{2}, \frac{2}{3}\right)$.

The value and the interval for u in the candidate D-optimal design ξ_z of the form (5.10) is calculated as follows. Differentiating D, in (5.12), with respect to u gives

$$\frac{\partial D}{\partial u} = \frac{(1-\lambda)\lambda u e^{3u}\left\{3(1+u+e^u-ue^u)[(u+\beta_0)^2 - 4\lambda\beta_0 u] + (\beta_0^2 - u^2)(1+e^u)\right\}}{(1+e^u)^7}.$$

Setting $\frac{\partial D}{\partial u} = 0$ and solving for u gives $u = 0$ or u satisfying the equation

$$3(1+u+e^u-ue^u)[(u+\beta_0)^2 - 4\lambda\beta_0 u] - (\beta_0^2 - u^2)(1+e^u) = 0 \tag{5.18}$$

Equation (5.18) cannot be solved analytically for u since its left hand side is a transcendental expression in u. However, as it will be shown later, numerical solutions can be found for given values of β_0. For now, note that $\left.\frac{\partial^2 D}{\partial u^2}\right|_{u=0} = \frac{(1-\lambda)\lambda\beta_0^2}{32} > 0$, and hence D takes on a minimum value at $u = 0$. Note also that dividing both sides of (5.13) and (5.18) by β_0^2 and calculating the limits of the results as β_0 goes to $-\infty$ gives $\lambda = \frac{1}{2}$ and

$$2 + 3u + 2e^u - 3ue^u = 0. \tag{5.19}$$

The equation (5.19) is the same as (4.27) for which the unique solution for u on $[0,\infty)$ is $u = 1.22291$. Hence, when β_0 tends to $-\infty$, the trapezium design ξ_z (5.10) tends to the equally weighted 4-point parallelogram design discussed in Section 5.2.

Since $-\beta_0 > u$ and $-4\lambda\beta_0 u > 0$, it follows from equation (5.18) that

$$1 + u + e^u - ue^u = \frac{(\beta_0^2 - u^2)(1+e^u)}{3[(u+\beta_0)^2 - 4\lambda\beta_0 u]} > 0. \tag{5.20}$$

Hence, suitable optimal values of u must satisfy the inequality

$$1 + u + e^u - ue^u > 0. \tag{5.21}$$

Consider the function $f(u) = 1 + u + e^u - ue^u$ on $[0,\infty)$. The equation $f(u) = 0$ is the same as (4.14) for which the unique solution on $[0,\infty)$ was found to be $u = 1.5434$. Furthermore, it was shown in Theorem 4.1 that $f(u) > 0$ on $[0, 1.5434]$ and $f(u) < 0$ on $(1.5434, \infty)$. Hence, for all $\beta_0 \in (-\infty; -1.5434)$, the candidate D-optimal value of u is $u^* \in (1.22291, 1.5434)$. ■

Numerical approximation of λ^* and u^* for given $\beta_0 \in (-\infty; -1.5434)$

For a given value of β_0, the values of λ^* and u^* are the respective numerical solutions for λ and u to the simultaneous equations (5.13) and (5.18). Table 5.1 gives the solutions λ^* and u^* for selected values of β_0. Figure 5.5 displays the scatter plots of u^* and λ^* versus β_0 in

Table 5.1: The relationship between β_0, u^* and λ^*.

β_0	u^*	λ^*	β_0	u^*	λ^*	β_0	u^*	λ^*
-500	1.2229	0.5012	-30	1.2252	0.5203	-6.5	1.2666	0.5856
-400	1.2229	0.5013	-20	1.2280	0.5302	-6	1.2734	0.5914
-300	1.2229	0.5020	-15	1.2318	0.5400	-5.5	1.2817	0.5978
-200	1.2230	0.5031	-12.5	1.2356	0.5476	-5	1.2923	0.6051
-100	1.2231	0.5061	-10	1.2425	0.5587	-4.5	1.3056	0.6132
-90	1.2232	0.5068	-9.5	1.2445	0.5615	-4	1.3230	0.6223
-80	1.2232	0.5076	-9	1.2468	0.5646	-3.5	1.3458	0.6323
-70	1.2233	0.5087	-8.5	1.2495	0.5680	-3	1.3763	0.6430
-60	1.2235	0.5102	-8	1.2527	0.5717	-2.5	1.4176	0.6537
-50	1.2237	0.5122	-7.5	1.2565	0.5759	-2	1.4741	0.6628
-40	1.2242	0.5152	-7	1.2611	0.5805	-1.6	1.5338	0.6666

the interval $[-10, -1.6]$. Both Table 5.1 and Figure 5.5 indicate that λ^* and u^* are increasing functions of β_0.

Now, let ξ_z^* be the design ξ_z given in (5.10) with $\lambda = \lambda^*$ and $u = u^*$ respective solutions for (λ, u) to the simultaneous equations (5.13) and (5.18). The D-optimality of design ξ_z^* is proved in Theorem 5.3 by showing that the standardized variance function $d(\boldsymbol{z}, \xi_z^*; \beta_0) = \mathrm{tr}\{M^{-1}(\xi_z^*; \beta_0)M(\boldsymbol{z}; \beta_0)\} \leq 3$ for all $\boldsymbol{z} = (z_1, z_2)^T \in [0, \infty) \times [0, \infty)$.

Theorem 5.3. The design

$$\xi_z^* = \left\{ \begin{array}{cccc} (-u^* - \beta_0, 0) & (0, -u^* - \beta_0) & (0, u^* - \beta_0) & (u^* - \beta_0, 0) \\ \frac{1-\lambda^*}{2} & \frac{1-\lambda^*}{2} & \frac{\lambda^*}{2} & \frac{\lambda^*}{2} \end{array} \right\} \tag{5.22}$$

for the two-variable binary logistic model $u = \mathrm{logit}(p) = \beta_0 + z_1 + z_2$ on the design space $[0, \infty) \times [0, \infty)$ with parameter $\beta_0 < -1.5434$, where u^* and λ^* satisfy the equalities

$$(1 - 2\lambda^*)(u^* + \beta_0)^2 - 4\lambda^*(2 - 3\lambda^*)\beta_0 u^* = 0$$

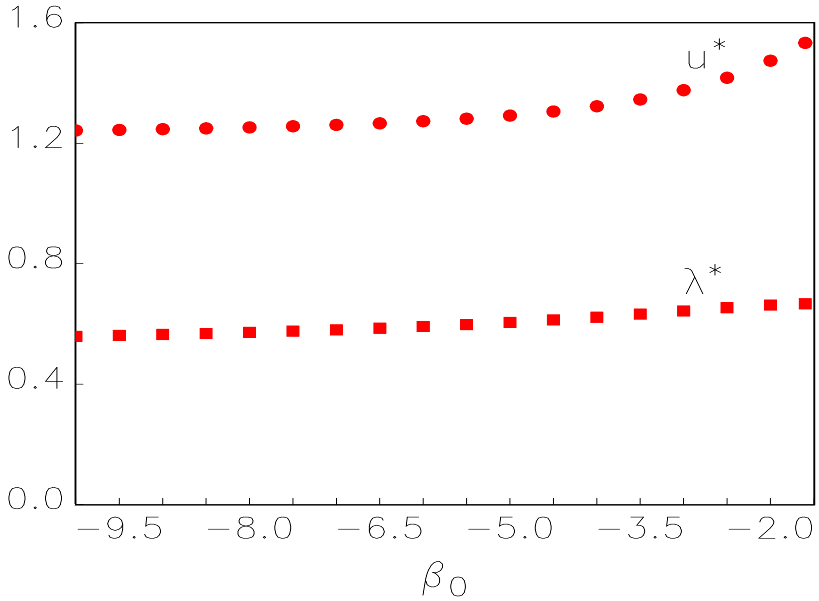


Figure 5.5: The scatter plots of λ^* (solid squares) and u^* (solid circles) versus β_0. Values of λ^* and u^* are respective solutions for λ and u in the simultaneous equations (5.13) or (5.14) and (5.18) respectively.

and

$$3(1+u^*+e^{u^*}-u^*e^{u^*})[(u^*+\beta_0)^2-4\lambda^*\beta_0u^*]-(\beta_0^2-u^{*2})(1+e^{u^*})=0,$$

is D-optimal.

Proof

Let $M(\xi_z^*;\beta_0)$ be the information matrix (5.11) evaluated at design ξ_z^* given in (5.22), and let $M(\boldsymbol{z},\beta_0)$ be the information matrix (5.8) evaluated at a one-point design with an arbitrary support point $\boldsymbol{z}=(z_1,z_2)^T=(z_1,u-\beta_0-z_1)^T\in\mathcal{D}=[0,\infty)\times[0,\infty)$ where $u=\mathrm{logit}(p)=\beta_0+z_1+z_2\geq\beta_0$ since $z_1\geq 0$ and $z_2\geq 0$. Theorem 5.3 is proved using the Equivalence Theorem for the D-optimality criterion (3.3) by showing that the standardized variance function

$$d(\boldsymbol{z},\xi_z^*;\beta_0)=\mathrm{tr}\left[M^{-1}(\xi_z^*;\beta_0)M(\boldsymbol{z};\beta_0)\right]\leq 3$$

for all $\boldsymbol{z}=(z_1,z_2)^T\in\mathcal{D}=[0,\infty)\times[0,\infty)$ with equality holding at the support points of the design ξ_z^*.

The standardized variance function $d(\boldsymbol{z}, \xi_z^*; \beta_0)$ simplifies to

$$d(\boldsymbol{z}, \xi_z^*; \beta_0) = \frac{e^{u-u^*}(1+e^{u^*})^2}{(1+e^u)^2}\left\{\frac{(u+u^*)^2 - 4\lambda^* uu^*}{4\lambda^*(1-\lambda^*)u^{*2}} + \frac{(u-\beta_0)^2 + 4[z_1^2 - (u-\beta_0)z_1]}{(u^*+\beta_0)^2 - 4\lambda^*\beta_0 u^*}\right\} \tag{5.23}$$

$$\leq g(u) = \frac{e^{u-u^*}(1+e^{u^*})^2}{(1+e^u)^2}\left\{\frac{(u+u^*)^2 - 4\lambda^* uu^*}{4\lambda^*(1-\lambda^*)u^{*2}} + \frac{(u-\beta_0)^2}{(u^*+\beta_0)^2 - 4\lambda^*\beta_0 u^*}\right\}. \tag{5.24}$$

The inequality (5.24) follows from the fact that for fixed $u \in [\beta_0, \infty)$ and $z_1 \in [0, u-\beta_0]$, the parabola $z_1^2 - (u-\beta_0)z_1$ takes on its maximum value of 0 at $z_1 = 0$ and $z_1 = u - \beta_0$.

Let $\boldsymbol{z}^* = (z_1^*, z_2^*)^T$ represent the support points of design ξ_z^*. It has to be shown that $d(\boldsymbol{z}^*, \xi_z^*; \beta_0) = 3$ and $g(u) \leq 3$ for all $u \in [\beta_0, \infty)$. The expressions $d(\boldsymbol{z}, \xi_z^*; \beta_0)$ given in (5.23) and $g(u)$ given in (5.24) are simplified using an identity derived as follows.

The expression (5.14) for λ^* gives

$$\begin{aligned}
\frac{\lambda^*}{3\lambda^*-1} &= \frac{\beta_0^2 + 6\beta_0 u^* + u^{*2} - \sqrt{\beta_0^4 + 14\beta_0^2 u^{*2} + u^{*4}}}{3(\beta_0^2 + 2\beta_0 u^* + u^{*2} - \sqrt{\beta_0^4 + 14\beta_0^2 u^{*2} + u^{*4}})} \\
&= \frac{(\beta_0^2 + 6\beta_0 u^* + u^{*2} - \sqrt{\beta_0^4 + 14\beta_0^2 u^{*2} + u^{*4}})(\beta_0^2 + 2\beta_0 u^* + u^{*2} + \sqrt{\beta_0^4 + 14\beta_0^2 u^{*2} + u^{*4}})}{3(\beta_0^2 + 2\beta_0 u^* + u^{*2} - \sqrt{\beta_0^4 + 14\beta_0^2 u^{*2} + u^{*4}})(\beta_0^2 + 2\beta_0 u^* + u^{*2} + \sqrt{\beta_0^4 + 14\beta_0^2 u^{*2} + u^{*4}})} \\
&= \frac{4\beta_0 u^*(2u^{*2} + 2\beta_0^2 + \sqrt{\beta_0^4 + 14\beta_0^2 u^{*2} + u^{*4}})}{12\beta_0 u^*(u^*-\beta_0)^2} \\
&= \frac{2(u^{*2} + \beta_0^2) + \sqrt{\beta_0^4 + 14\beta_0^2 u^{*2} + u^{*4}}}{3(u^*-\beta_0)^2}.
\end{aligned}$$

Hence,

$$\frac{\lambda^*(u^*-\beta_0)^2}{3\lambda^*-1} = \frac{2(u^{*2} + \beta_0^2) + \sqrt{\beta_0^4 + 14\beta_0^2 u^{*2} + u^{*4}}}{3}. \tag{5.25}$$

Also,

$$(u^*+\beta_0)^2 - 4\lambda^*\beta_0 u^* = \frac{2(u^{*2} + \beta_0^2) + \sqrt{\beta_0^4 + 14\beta_0^2 u^{*2} + u^{*4}}}{3}. \tag{5.26}$$

Thus, comparing (5.25) and (5.26) gives the identity

$$(u^*+\beta_0)^2 - 4\lambda^*\beta_0 u^* = \frac{\lambda^*(u^*-\beta_0)^2}{3\lambda^*-1}. \tag{5.27}$$

Substituting the right hand side of (5.27) with the left hand side of (5.27) in (5.23) and in (5.24) gives

$$d(\boldsymbol{z}, \xi_z^*; \beta_0) = \frac{e^{u-u^*}(1+e^{u^*})^2}{(1+e^u)^2}\left\{\frac{(u+u^*)^2 - 4\lambda^* uu^*}{4\lambda^*(1-\lambda^*)u^{*2}} + \frac{(3\lambda^*-1)\left\{(u-\beta_0)^2 + 4[z_1^2 - (u-\beta_0)z_1]\right\}}{\lambda^*(u^*-\beta_0)^2}\right\} \tag{5.28}$$

$$\leq g(u) = \frac{e^{u-u^*}(1+e^{u^*})^2}{(1+e^u)^2}\left\{\frac{(u+u^*)^2 - 4\lambda^* uu^*}{4\lambda^*(1-\lambda^*)u^{*2}} + \frac{(3\lambda^*-1)(u-\beta_0)^2}{\lambda^*(u^*-\beta_0)^2}\right\} \text{ for } u \geq \beta_0. \tag{5.29}$$

Evaluating (5.28) at $\boldsymbol{z} = \boldsymbol{z}^* = (z_1^*, z_2^*)^T$, the support points of design ξ_z^*, is equivalent to substituting (u, z_1) with $(u^*, 0)$ and with $(u^*, u^* - \beta_0)$. The two substitutions give $d(\boldsymbol{z}^*, \xi_z^*; \beta_0) = 3$ as required. What remains to be shown is that $g(u)$ given by (5.29) is less than or equal to 3 for all $u \in [\beta_0, \infty)$ and $\beta_0 \in (-\infty, -1.5434)$. The following is an algebraic proof that $g(u) \leq 3$ for all $u \in [\beta_0, \infty)$ and $\beta_0 \in (-\infty, -1.5434)$.

Consider $g(u)$ given by (5.29). The function $g(u)$ is even on $[\beta_0, -\beta_0]$ since u^* the solution for u to equations (5.13) implies that

$$g(u) - g(-u) = \frac{ue^{u-u^*}(1+e^{u^*})^2\left\{(1-2\lambda^*)(u^*+\beta_0)^2 - 4\lambda^*(2-3\lambda^*)\beta_0 u^*\right\}}{u^*\lambda^*(1-\lambda^*)(u^*-\beta_0)^2(1+e^u)^2} = 0,$$

and thus $g(u) = g(-u)$. The symmetry of $g(u)$ on $[\beta_0, -\beta_0]$ implies that if $g(u) \leq 3$ on $[0, -\beta_0]$, then $g(u) \leq 3$ on $[\beta_0, 0]$. Hence, it has only to be shown that $g(u) \leq 3$ on $[0, \infty)$.

Consider $g(u)$ on $[0, \infty)$. Clearly, $g(u^*) = 3$. It has to be shown that $g'(u) \geq 0$ on $[0, u^*]$ and that $g'(u) < 0$ on (u^*, ∞). After some simplifications, using the fact that

$$(1-2\lambda^*)(u^*+\beta_0)^2 - 4\lambda^*(2-3\lambda^*)\beta_0 u^* = 0,$$

the first derivative of $g(u)$ with respect to u is

$$g'(u) = \frac{e^{u-u^*}(1+e^{u^*})^2 h(u)}{4\lambda^*(1-\lambda^*)u^{*2}(u^*-\beta_0)^2(1+e^u)^3} \tag{5.30}$$

where

$$\begin{aligned} h(u) \;=\; & (u^*-\beta_0)^2(u^{*2}+u^2+2u) + 4u^{*2}(3\lambda^*-1)(1-\lambda^*)(\beta_0^2+u^2+2u) \\ & -e^u\left[(u^*-\beta_0)^2(u^{*2}+u^2-2u) + 4u^{*2}(3\lambda^*-1)(1-\lambda^*)(\beta_0^2+u^2-2u)\right]. \end{aligned} \tag{5.31}$$

Hence, showing that $g'(u) \geq 0$ on $[0, u^*]$ and $g'(u) < 0$ on (u^*, ∞) are respectively equivalent to showing that $h(u) \geq 0$ on $[0, u^*]$ and $h(u) < 0$ on (u^*, ∞).

Consider $h(u)$ on $[0, \infty)$. Clearly, $h(0) = 0$. Furthermore, the equality

$$3(1+u^*+e^{u^*}-u^*e^{u^*})[(u^*+\beta_0)^2 - 4\lambda\beta_0 u^*] - (\beta_0^2 - u^{*2})(1+e^{u^*}) = 0$$

can be written as

$$e^{u^*} = \frac{u^{*2}(3u^*+4) + \beta_0^2(3u^*+2) - 6\beta_0 u^*(2\lambda-1)(u^*+1)}{u^{*2}(3u^*-4) + \beta_0^2(3u^*-2) - 6\beta_0 u^*(2\lambda-1)(u^*-1)}. \tag{5.32}$$

Then, substituting e^{u^*} in (5.31) gives

$$h(u^*) = \frac{4(5-6\lambda^*)u^*(u^{*2}-\beta_0^2)\left[(1-2\lambda^*)(u^*+\beta_0)^2 - 4\lambda^*(2-3\lambda^*)u^*\beta_0\right]}{\beta_0^2(2-3u^*) + 6\beta_0(2\lambda^*-1)(u^*-1)u^* + (4-3u^*)u^{*2}} = 0$$

since $(u^* + \beta_0)^2(1 - 2\lambda^*) - 4\lambda^*(2 - 3\lambda^*)\beta_0 u^* = 0$. Hence, $h(u) = 0$ has two solutions $u = 0$ and $u = u^*$. The fact that $h(0) = h(u^*) = 0$ implies that $h(u)$ has at least one stationary point in the interval $(0, u^*)$ since $h(u)$ is continuous on $[0, \infty)$. Furthermore, $h(u)$ goes to $-\infty$ as u goes to ∞ since

$$\lim_{u\to\infty} h(u) = -\lim_{u\to\infty} [(u^* - \beta_0)^2 + 4u^{*2}(3\lambda^* - 1)(1 - \lambda^*)]u^2 e^u = -\infty.$$

As a consequence, if there is only one stationary point for $u \in (0, \infty)$ then the stationary point would be a maximum and the equation $h(u) = 0$ would have only two solutions $u = 0$ and $u = u^*$. For illustration purpose, Figure 5.6 (a) indicates the profile of function $h(u)$ when for example $\beta_0 = -3$.

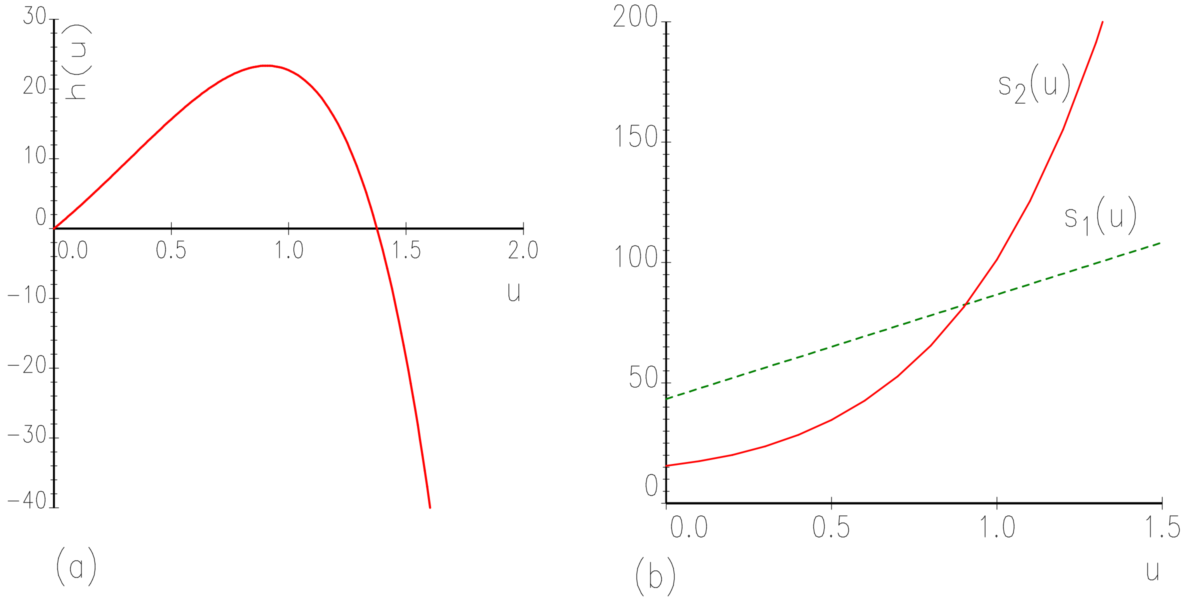


Figure 5.6: (a) Function $h(u)$ and (b) functions $s_1(u)$ (dashed curve) and $s_2(u)$ (solid curve) for $\beta_0 = -3$.

Stationary points of $h(u)$ are obtained by solving $s(u) = h'(u) = 0$. In the present case $s(u) = h'(u) = s_1(u) - s_2(u)$ where $s_1(u) = 2(u+1)[(u^* - \beta_0)^2 + 4u^{*2}(3\lambda^* - 1)(1 - \lambda^*)]$ is a straight line with both slope and intercept equal to $2[(u^* - \beta_0)^2 + 4u^{*2}(3\lambda^* - 1)(1 - \lambda^*)]$ and $s_2(u) = [(u^* - \beta_0)^2(u^{*2} + u^2 - 2) + 4u^{*2}(3\lambda^* - 1)(1 - \lambda^*)(\beta_0^2 + u^2 - 2)]e^u$ is a convex function on $[0, \infty)$ since $s_2''(u) = [(u^* - \beta_0)^2(u^{*2} + u^2 + 4u) + 4u^{*2}(3\lambda^* - 1)(1 - \lambda^*)(\beta_0^2 + u^2 + 4u)]e^u > 0$. In addition, note that $s_1(0) > s_2(0)$. In fact,

$$s_1(0) - s_2(0) = (u^* - \beta_0)^2(4 - u^{*2}) + 4u^{*2}(3\lambda^* - 1)(1 - \lambda^*)(4 - \beta_0^2)$$

which can be written as:

$$s_1(0) - s_2(0) = (u^{*2} - 2u^*\beta_0)(4 - u^{*2}) + 16u^{*2}(3\lambda^* - 1)(1 - \lambda^*) + \beta_0^2[4 + 3u^{*2} + u^{*2}(3\lambda^{*2} - 4\lambda^*)].$$

Clearly, $u^{*2} - 2u^*\beta_0 > 0$ because $-\beta_0 > 0$ and $u^* > 0$. Furthermore, the conditions $\frac{1}{2} < \lambda^* < \frac{2}{3}$ and $1.22291 < u^* < 1.5434$ shown in Theorem 5.2 imply that $4 - u^{*2} > 0$ and $(3\lambda^* - 1)(1 - \lambda^*) > 0$. Moreover, the quadratic convex polynomial $f(\lambda^*) = 3\lambda^{*2} - 4\lambda^*$ is negative on $\left[\frac{1}{3}, \frac{2}{3}\right]$ and has its minimum value at the boundary $\lambda^* = \frac{2}{3}$ which corresponds to $u^* = -\beta_0 = 1.5434$. Evaluating $s_1(0) - s_2(0)$ at $\lambda^* = \frac{2}{3}$ and $u^* = -\beta_0 = 1.5434$ gives $s_1(0) - s_2(0) = 20.5547 > 0$. Hence, the straight line $s_1(u)$ is above the convex curve $s_2(u)$ at $u = 0$. Moreover, $\lim_{u\to\infty}[s_2(u) - s_1(u)] = \lim_{u\to\infty}(u^* - \beta_0)^2u^2e^u = \infty$. Hence, the convex curve $s_2(u)$ passes above line $s_1(u)$ as $u \to \infty$. The alternating values of the straight line $s_1(u)$ and the convex curve $s_2(u)$ at $u = 0$ and $u = \infty$ imply that line $s_1(u)$ and curve $s_2(u)$ intersect only once on $(0, \infty)$, and in particular only once on $(0, -\beta_0)$. Figure 5.6 (b) illustrates the profiles of $s_1(u)$ and $s_2(u)$ when $\beta_0 = -3$. The fact that $s_1(u)$ and $s_2(u)$ intersects only once on $[0, \infty)$ implies that the equation $h'(u) = 0$ has only one solution for $u \in [0, \infty)$. In other words, $h(u)$ has a single stationary point on $[0, \infty)$ and the stationary point is a maximum located in the interval $[0, u^*]$. Therefore, $h(u) \geq 0$ on $[0, u^*]$ and $h(u) < 0$ on (u^*, ∞), or equivalently, the sign of $g'(u)$ changes from positive to negative only once at $u = u^*$. Thus, overall the function $g(u)$ in (5.29) has a local minimum $g(0)$ at $u = 0$ and a unique maximum $g(u) = 3$ at $u = u^*$ on $[0, \infty)$, and by symmetry about $u = 0$, $g(u)$ has a unique maximum $g(u) = 3$ at $u = -u^*$ on $(\beta_0, 0]$. ■

Numerical check of D-optimality for design ξ_z^* given in (5.22)

D-optimality of a candidate optimal design is often checked numerically by verifying on a plot that the standardized variance function is less than or equal to the number of model parameters with equality holding at the support points of the candidate optimal design. Figure 5.7 displays the graphs of $g(u)$ given in (5.29) versus $u \geq \beta_0 = \{-8, -4, -2\}$. Note that $g(u)$ was evaluated at these values of β_0 and the corresponding u^* values that are displayed in Table 5.1. Figure 5.7 indicates that $g(u) \leq 3$ for all $u \geq \beta_0 = \{-8, -4, -2\}$, with $g(u) = 3$ holding only at $u = \pm u^*$, and hence suggests that the 4-point designs ξ_z^* in (5.22) associated with these values of β_0 are D-optimal.

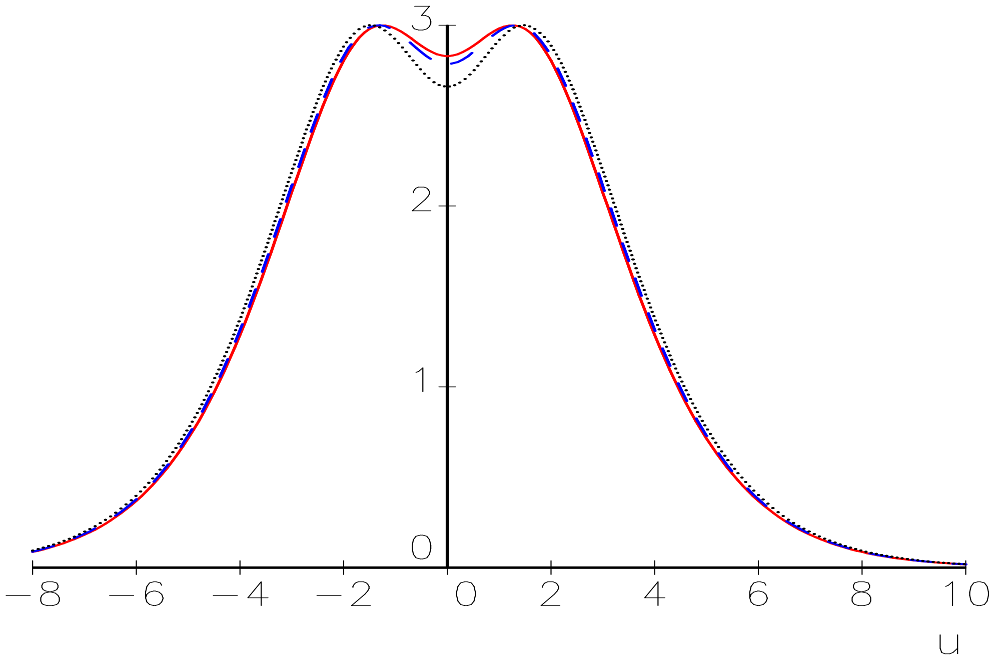


Figure 5.7: Function $g(u)$ given by (5.29) with $\beta_0 = -8$ (solid line), $\beta_0 = -4$ (dashed line), $\beta_0 = -2$ (dotted line).

Approximate relationship between β_0 and u^*

The values of u^* and λ^* in Table 5.1 were found for just the selected values of β_0. Since all values of β_0 with $\beta_0 \in (-\infty, -1.5434)$ are not enumerated, it may be of interest to approximate the relationship between β_0 and u^* by a regression equation. Then, the value of λ^* can be derived using the straightforward equation (5.14). The scatter plot of u^* versus β_0 illustrated in Figure 5.5 indicates that an exponential function with a lower asymptote can represent u^* in terms of β_0. Fitting an exponential function of the form $u = \theta_0 + \theta_1 e^{\theta_2 \beta_0}$ to the data using moderate values of β_0 (here, from $\beta_0 = -500$ with steps of 10, then a change of steps to 1 from $\beta_0 = -60$ to -20, then steps smaller or equal to 0.5 beyond -20) gives the fitted model

$$\widehat{u} = 1.2260 + 0.5716 e^{0.4256 \beta_0} \tag{5.33}$$

with $R^2 \simeq 0.993$. It can be important to evaluate the efficiencies of the design generated using model (5.33) when compared to designs generated using solutions of the equation (5.18) as in Table 5.1. Recall from Section 3.7.2 that D-efficiency of ξ relative to ξ_z^* is defined by

$$D_{\text{eff}} = \left\{ \frac{|M(\xi; \boldsymbol{\beta})|}{|M(\xi_z^*, \boldsymbol{\beta})|} \right\}^{1/p},$$

where ξ is an arbitrary design, ξ_z^* is the optimal design and p is the number of the parameters of the model. In the present case $p = 3$, ξ_z^* is the design (5.10) with λ and u solutions of the equations (5.13) and (5.18) respectively. Now $\xi = \xi^{**}$ is the counterpart of design (5.10) obtained using $\widehat{u}$ given by (5.33), and λ obtained from (5.14) evaluated at $u = \widehat{u}$. Table 5.2 provides the efficiencies for selected values of β_0. Clearly, all the efficiencies are close to 1.

Table 5.2: Values of u^*, $\widehat{u}$ and D-efficiencies for selected values of β_0.

β_0	u^*	$\widehat{u}$	D_{eff}
-10	1.2425	1.2341	0.999973
-9	1.2468	1.2384	0.999973
-8	1.2527	0.2450	0.999977
-7	1.2611	1.2551	0.999986
-6	1.2734	1.2705	0.999997
-5	1.2923	1.2941	0.999999
-4	1.3230	1.3302	0.999982
-3	1.3763	1.3854	0.999972
-2	1.4741	1.4700	0.999995

Therefore, in practice there is little loss in efficiency if the optimal value of u is estimated using model (5.33) instead of exact optimal value of u calculated using equation (5.18).

5.3.3 3-point design

In Section 5.3.2 it was shown that the 4-point D-optimal trapezium design exists if $-\infty < \beta_0 < -1.5434$. In this section it is shown that if $-1.5434 \leq \beta_0 \leq 0$, then the 4-point D-optimal trapezium design degenerates to a 3-point D-optimal design.

Consider Figures 5.4 (a). Note that $(z_1, z_2) \in [0, \infty) \times [0, \infty)$, and hence the condition for the 4-point design displayed in the figure is that $-u - \beta_0 = z_1 + z_2 > 0$. However, if $-u - \beta_0 = z_1 + z_2 \leq 0$, then $z_1 = z_2 = 0$ and $u \geq -\beta_0$, and hence the 4-point design becomes the 3-point design with optimal logits $\beta_0 \leq 0$ and $u > -\beta_0 > 0$. Figure 5.8 (a) displays the layout of the 3-point design. The 3-point design displayed in the figure has the following form:

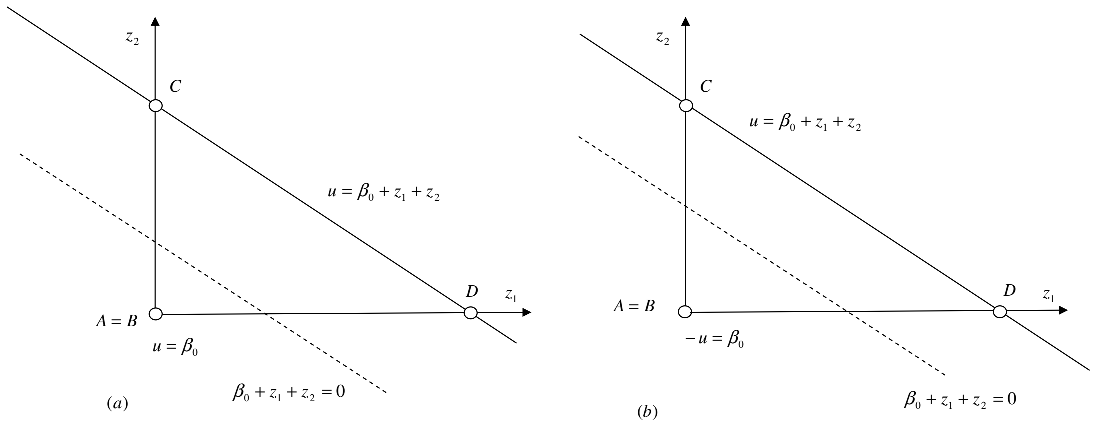


Figure 5.8: Pattern of a 3-point D-optimal design for a two-variable binary logistic model without interaction $u = \beta_0 + z_1 + z_2$ on the design space $\mathcal{D} = [0, \infty) \times [0, \infty)$: (a) $u = \beta_0$ and (b) $u = -\beta_0$.

$$\xi_z = \begin{Bmatrix} (z_1, z_2) \\ \lambda \\ u \end{Bmatrix} = \begin{Bmatrix} (0,0) & (0, u-\beta_0) & (u-\beta_0, 0) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \beta_0 & u & u \end{Bmatrix} \tag{5.34}$$

where $u > 0$. Let $M(\xi_z; \beta_0)$ be the information matrix for the parameter vector $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^T$, on the (z_1, z_2)-space, evaluated at design ξ_z given by (5.34). Then,

$$M(\xi_z; \beta_0) = \frac{1}{3} \begin{bmatrix} \frac{e^{\beta_0}}{(1+e^{\beta_0})^2} + 2\frac{e^u}{(1+e^u)^2} & \frac{e^u(u-\beta_0)}{(1+e^u)^2} & \frac{e^u(u-\beta_0)}{(1+e^u)^2} \\ \frac{e^u(u-\beta_0)}{(1+e^u)^2} & \frac{e^u(u-\beta_0)^2}{(1+e^u)^2} & 0 \\ \frac{e^u(u-\beta_0)}{(1+e^u)^2} & 0 & \frac{e^u(u-\beta_0)^2}{(1+e^u)^2} \end{bmatrix} \tag{5.35}$$

and its determinant is given by

$$D = \frac{(u-\beta_0)^4 e^{2u+\beta_0}}{27(1+e^{\beta_0})^2(1+e^u)^4}.$$

Maximizing D with respect to u is equivalent to maximizing

$$D_s = \frac{(u-\beta_0)^4 e^{2u}}{(1+e^u)^4} \tag{5.36}$$

with respect to u. Differentiating the expression (5.36) with respect to u gives

$$\frac{\partial D_s}{\partial u} = -\frac{2e^{2u}(u-\beta_0)^3(2-\beta_0+u+2e^u+\beta_0 e^u - ue^u)}{(1+e^u)^5}. \tag{5.37}$$

Hence, the candidate D-optimal value of u is a solution for u to the equation

$$2-\beta_0+u+2e^u+\beta_0 e^u - ue^u = 0 \tag{5.38}$$

which is greater than or equal to $-\beta_0$, and has a unique solution on $[-\beta_0, \infty)$ since it has unique solution on $[0, \infty)$. In fact, consider the continuous function of u, $f(u) = 2 - \beta_0 + u + 2e^u + \beta_0 e^u - ue^u$ on $[0, \infty)$. Clearly, $f(0) = 4 > 0$ and $\lim_{u \to \infty} f(u) = -\lim_{u \to \infty} ue^u = -\infty$, i.e. $f(u) = 0$ has at least one solution on $[0, \infty)$. In addition, $f''(u) = (\beta_0 - u)e^u < 0$ since $u \geq 0$ and $\beta_0 < 0$. Hence, $f(u)$ is concave on $[0, \infty)$ and thus, $f(u) = 0$ or equation (5.38) has a unique solution for u on $[0, \infty)$.

Consider the special case when $\beta_0 = -u$ with $u > 0$. Figure 5.8 (b) is a graphical representation of the design for this case. The support points of the candidate D-optimal design are A, B and C symbolized by circles. Setting $\beta_0 = -u$ in equation (5.38) gives $1 + u + e^u - ue^u = 0$ which is the same as equation (4.14) which was shown to have the unique solution $u^* = 1.5434$ for $u \geq 0$. Furthermore, the equation is the same as equation (5.18) evaluated at $\beta_0 = -u$ and at $\lambda = \frac{2}{3}$, the solution for λ to the equation (5.13) when $\beta_0 = -u$. Thus, from this observation and Theorem 5.2 it is deduced that $\beta_0 = -u = -1.5434$ forms the "boundary" between the 3-point and the 4-point trapezium design. Table 5.3, second row, displays the numerical solutions, u^*, of equation (5.38) evaluated at selected values of $\beta_0 \in [-1.5434, 0]$. The table suggests that the solutions u^* of equation (5.38) form an increasing function of $\beta_0 \in [-1.5434, 0]$. What remains is to prove the D-optimality of the design (5.34) with $u = u^*$ a solution of equation (5.38) for all $\beta_0 \in [-1.5434, 0]$. The D-optimality of design ξ_z^*

Table 5.3: Relationship between β_0 and u^*, and a local minimum $u = u_m$ of $g(u)$ where $u = \beta_0 + \alpha(u^* - \beta_0)$ and $0 \leq \alpha \leq 1$: 3-point D-optimal design.

β_0	-1.543	-1.5	-1.25	-1	-0.75	-0.5	-0.25	-0.15	-0.1	0
$u = u^*$	1.543	1.562	1.674	1.796	1.930	2.075	2.231	2.297	2.331	2.399
$u = u_m$	0	0.001	0.043	0.132	0.255	0.404	0.571	0.642	0.679	0.753

given in (5.34) is proved in Theorem 5.4 by showing that the standardized variance function $d(\boldsymbol{z}, \xi_z^*; \beta_0) = \text{tr}\{M^{-1}(\xi_z^*; \beta_0)M(\boldsymbol{z}; \beta_0)\} \leq 3$ for all $\boldsymbol{z} = (z_1, z_2)^T \in [0, \infty) \times [0, \infty)$.

Theorem 5.4. The design

$$\xi_z^* = \left\{ \begin{array}{ccc} (0,0) & (0, u^* - \beta_0) & (u^* - \beta_0, 0) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \beta_0 & u & u \end{array} \right\} \tag{5.39}$$

for the two-variable binary logistic model $u = \text{logit}(p) = \beta_0 + z_1 + z_2$ on the design space $[0, \infty) \times [0, \infty)$ with parameter $\beta_0 \in [-1.5434, 0]$, where $u^* > 0$ satisfies the equation $2 - \beta_0 + u^* + 2e^{u^*} + \beta_0 e^{u^*} - u^* e^{u^*} = 0$, is D-optimal.

Proof

Let $M(\xi_z^*; \beta_0)$ be the information matrix (5.35) evaluated at design ξ_z^* given in (5.39), and let $M(\boldsymbol{z}, \beta_0)$ be the information matrix (5.8) evaluated at a one-point design with an arbitrary support point $\boldsymbol{z} = (z_1, z_2)^T = (z_1, u - \beta_0 - z_1)^T \in \mathcal{D} = [0, \infty) \times [0, \infty)$ where $u = \text{logit}(p) = \beta_0 + z_1 + z_2 \geq \beta_0$ since $z_1 \geq 0$ and $z_2 \geq 0$. Theorem 5.4 is proved using the Equivalence Theorem for the D-optimality criterion (3.3) by showing that the condition

$$d(\boldsymbol{z}, \xi_z^*; \beta_0) = \text{tr}\left[M^{-1}(\xi_z^*; \beta_0) M(\boldsymbol{z}; \beta_0)\right] \leq 3$$

holds for all $\boldsymbol{z} = (z_1, z_2)^T \in \mathcal{D} = [0, \infty) \times [0, \infty)$ with equality holding at the support points of the design ξ_z^*. Simple algebraic calculations show that for all $\boldsymbol{z} = (z_1, z_2)^T \in [0, \infty) \times [0, \infty)$,

$$\begin{aligned} d(\boldsymbol{z}, \xi_z^*; \beta_0) =& \frac{3e^u \left\{e^{-\beta_0}(1+e^{\beta_0})^2(u-u^*)^2 + e^{-u^*}(1+e^{u^*})^2[(u-\beta_0)^2 + 2[z_1^2 - (u-\beta_0)z_1]]\right\}}{(u^*-\beta_0)^2(1+e^u)^2} && (5.40) \\ \leq & g(u) = \frac{3e^u \left\{e^{-\beta_0}(1+e^{\beta_0})^2(u-u^*)^2 + e^{-u^*}(1+e^{u^*})^2(u-\beta_0)^2\right\}}{(u^*-\beta_0)^2(1+e^u)^2} \text{ for } u \geq \beta_0. && (5.41) \end{aligned}$$

The inequality (5.41) follows from the fact that for fixed $u \geq \beta_0$ and $z_1 \in [0, u - \beta_0]$, the parabola $z_1^2 - (u - \beta_0)z_1$ takes on its maximum value of 0 at $z_1 = 0$ and $z_1 = u - \beta_0$. Evaluating (5.40) at $\boldsymbol{z} = \boldsymbol{z}^* = (z_1, z_2)^T$, the support points of design ξ_z^*, is equivalent to substituting (u, z_1) with $(\beta_0, 0)$, $(u^*, 0)$ and $(u^*, u^* - \beta_0)$. All the three substitutions give $d(\boldsymbol{z}, \xi_z^*; \beta_0) = 3$ as required at the support of a D-optimal design. What remains to be shown is that $g(u)$ given by (5.41) is less than or equal to 3 for all $u \geq \beta_0 \in [-1.5434, 0]$.

Consider the function $g(u)$ given by (5.41) and $u^* = -\beta_0 = 1.5434$. Then, (5.41) reduces to

$$g(u) = \frac{3e^{u-u^*}(1+e^{u^*})^2(u^2+u^{*2})}{2u^{*2}(1+e^u)^2} \leq 3$$

since $\dfrac{e^{u-u^*}(1+e^{u^*})^2(u^2+u^{*2})}{u^{*2}(1+e^u)^2} \leq 2$ for $u^* = 1.5434$ and all $u \in [\beta_0, \infty)$ as was shown in Theorem 4.1.

Consider the case of $u^* > -\beta_0$ with $\beta_0 \in [-1.5434, 0]$. It has to be shown that $g(u) \leq 3$ with equality holding only at $u = \beta_0$ and $u = u^*$. Since $u \geq \beta_0$, u can be written as $u = \beta_0 + \alpha(u^* - \beta_0)$ where $\alpha \in [0, \infty)$. In particular, $u = \beta_0$ and $u = u^*$ respectively correspond to $\alpha = 0$ and $\alpha = 1$. Hence, function (5.41) can be written in terms of α as

$$g(\alpha) = \frac{3e^{\beta_0+\alpha(u^*-\beta_0)}\left\{\alpha^2 e^{-u^*}(1+e^{u^*})^2 + (\alpha-1)^2 e^{-\beta_0}(1+e^{\beta_0})^2\right\}}{\left[1+e^{\beta_0+\alpha(u^*-\beta_0)}\right]^2}. \tag{5.42}$$

The aim of the following derivations is to show that $g(\alpha)$ attains the maximum 3 only at $\alpha = 0$ and $\alpha = 1$, a minimum at $\alpha \in (0, 1)$, and that $g(\alpha)$ goes to 0 as α gets larger and larger.

Clearly, $g(0) = g(1) = 3$ as required at the support points of a D-optimal design. The stationary points of $g(\alpha)$ are zeros of the function

$$g'(\alpha) = \frac{3e^{\beta_0+\alpha(u^*-\beta_0)}\left[h_1(\alpha) - h_2(\alpha)\right]}{\left[1 + e^{\beta_0+\alpha(u^*-\beta_0)}\right]^3}$$

where

$$h_1(\alpha) = \left[\alpha^2(u^* - \beta_0) + 2\alpha\right] e^{-u^*}(1 + e^{u^*})^2 + \left[(\alpha - 1)^2(u^* - \beta_0) + 2(\alpha - 1)\right] e^{-\beta_0}(1 + e^{\beta_0})^2 \tag{5.43}$$

is a convex parabola in α and

$$h_2(\alpha) = e^{u}\left\{\left[\alpha^2(u^* - \beta_0) - 2\alpha\right] e^{-u^*}(1 + e^{u^*})^2 + \left[(\alpha - 1)^2(u^* - \beta_0) - 2(\alpha - 1)\right] e^{-\beta_0}(1 + e^{\beta_0})^2\right\} \tag{5.44}$$

where $u = \beta_0 + \alpha(u^* - \beta_0)$. The number of stationary points of $g(\alpha)$ given in (5.42) is the number of solutions of $g'(\alpha) = 0$, or, equivalently of $h_1(\alpha) = h_2(\alpha)$. It follows from $g'(\alpha)$ and the fact that u^* is solution of equation (5.38) that

$$g'(1) = \frac{3\left\{2 - \beta_0 + u^* + (2 + \beta_0 - u^*)e^{u^*}\right\}}{1 + e^{u^*}} = 0.$$

Hence, $g(\alpha)$ has a stationary point at $\alpha = 1$, or equivalently, $h_1(\alpha)$ and $h_2(\alpha)$ given by (5.43) and (5.44) meet at $\alpha = 1$. Curves $h_1(\alpha)$ and $h_2(\alpha)$ intersect only twice on $[0, \infty)$ as it is shown below.

First, note that $h_2(0) > h_1(0)$. In fact, the condition $u^* > -\beta_0$ implies that $(u^* + \beta_0) > 0$, and u^* solution of equation (5.38) implies that $(u^* - \beta_0 - 2) = (u^* - \beta_0 + 2)e^{-u^*} > 0$. Hence,

$$\begin{aligned} h_2(0) - h_1(0) &= (1 + e^{\beta_0})^2\left\{(u^* - \beta_0 + 2) - (u^* - \beta_0 - 2)e^{-\beta_0}\right\} \\ &= (1 + e^{\beta_0})^2(u^* - \beta_0 + 2)(1 - e^{-u^*-\beta_0}) > 0. \end{aligned}$$

Thus, curve $h_2(\alpha)$ is above the parabola $h_1(\alpha)$ at $\alpha = 0$. Furthermore, curve $h_2(\alpha)$ has only one point of inflexion on $[0, \infty)$, and hence intersects the parabola $h_1(\alpha)$ in only one point α other than $\alpha = 1$. In fact, the second derivative of $h_2(\alpha)$ with respect to α is

$$\begin{aligned} h_2''(\alpha) = (u^* - \beta_0)e^{\beta_0+\alpha(u^*-\beta_0)}\Big\{&(u^* - \beta_0)^2\left[e^{-u^*}(1 + e^{u^*})^2 + e^{-\beta_0}(1 + e^{\beta_0})^2\right]\alpha^2 \\ &+ 2\alpha(u^* - \beta_0)\left[e^{-u^*}(1 + e^{u^*})^2 + (1 - u^* + \beta_0)e^{-\beta_0}(1 + e^{\beta_0})^2\right] \\ &- 2e^{-u^*}(1 + e^{u^*})^2 - 2(u^* - \beta_0 + 1)e^{-\beta_0}(1 + e^{\beta_0})^2 + (u^* - \beta_0)^2e^{-\beta_0}(1 + e^{\beta_0})^2\Big\} \end{aligned}$$

The points of inflexion of $h_2(\alpha)$ are solutions of equation $h_2''(\alpha) = 0$, and change of concavity are determined by the sign of $h_2''(\alpha)$. Since $(u^* - \beta_0)e^{\beta_0+\alpha(u^*-\beta_0)} > 0$, zeros of $h_2''(\alpha)$ are those of the quadratic polynomial in α:

$$\begin{aligned} h_3(\alpha) = {} & (u^* - \beta_0)^2 \left[e^{-u^*}(1+e^{u^*})^2 + e^{-\beta_0}(1+e^{\beta_0})^2\right] \alpha^2 \\ & + 2\alpha(u^* - \beta_0)\left[e^{-u^*}(1+e^{u^*})^2 + (1 - u^* + \beta_0)e^{-\beta_0}(1+e^{\beta_0})^2\right] \\ & - 2e^{-u^*}(1+e^{u^*})^2 - 2(u^* - \beta_0 + 1)e^{-\beta_0}(1+e^{\beta_0})^2 + (u^* - \beta_0)^2 e^{-\beta_0}(1+e^{\beta_0})^2. \end{aligned} \tag{5.45}$$

Using properties of quadratic equations, the product of the roots of $h_3(\alpha) = 0$ is

$$s(\beta_0, u^*) = \frac{-2e^{-u^*}(1+e^{u^*})^2 - 2(u^* - \beta_0 + 1)e^{-\beta_0}(1+e^{\beta_0})^2 + (u^* - \beta_0)^2 e^{-\beta_0}(1+e^{\beta_0})^2}{(u^* - \beta_0)^2\left[e^{-u^*}(1+e^{u^*})^2 + e^{-\beta_0}(1+e^{\beta_0})^2\right]} \tag{5.46}$$

Clearly, the denominator of expression (5.46) is positive, and the two first terms of the numerator are negative. The third term in the numerator of (5.46) is positive, but as $\beta_0 < 0$, then $e^{-\beta_0}(1+e^{\beta_0})^2$ is a decreasing function of β_0 with $\beta_0 \in [-1.5434, 0]$. Hence, the numerator attains its largest value -4.449 at $\beta_0 = -1.5434$. Therefore, $s(\beta_0, u^*) < 0$ for $\beta_0 \in [-1.5434, 0]$ which implies that $h_2(\alpha)$ has two points of inflexion of opposite signs, say $\alpha_1 < 0$ and $\alpha_2 > 0$. However, α_1 is to be rejected because it is not in the domain $[0, \infty)$. Hence, $h_2(\alpha)$ changes its concavity only once on $[0, \infty)$ because $h_2(\alpha)$ has only one point of inflexion α_2 on $[0, \infty)$. Furthermore, $\alpha = 1$ belongs to the subset of $[0, \infty)$ where $h_2(\alpha)$ is convex. In fact $u^* > -\beta_0 > 0$ and $u^* - \beta_0 - 2 > 0$ imply that $e^{-u^*}(1+e^{u^*})^2 > e^{-\beta_0}(1+e^{\beta_0})^2$, and hence

$$h_3(1) = \left[(u^* - \beta_0)^2 + 2(u^* - \beta_0 - 1)\right] e^{-u^*}(1+e^{u^*})^2 - 2e^{-\beta_0}(1+e^{\beta_0})^2 > 0.$$

Hence, $h_2(\alpha)$ is concave on $[0, \alpha_2]$ and convex on $[\alpha_2, \infty)$. As a consequence, the curve $h_2(u)$ and the parabola $h_1(\alpha)$ cross each other at $\alpha = 1$ and at another point $\alpha_m \in (0, 1)$. The above derivations can be summarized as follows.

- If $\alpha \in [0, \alpha_m]$, then $h_1(\alpha) - h_2(\alpha) < 0$, i.e. $g(\alpha)$ decreases on $[0, \alpha_m]$.
- If $\alpha \in [\alpha_m, 1]$, then $h_1(\alpha) - h_2(\alpha) > 0$, i.e. $g(\alpha)$ increases on $[\alpha_m, 1]$.
- If $\alpha \in [1, \infty]$, then $h_1(\alpha) - h_2(\alpha) < 0$, i.e. $g(\alpha)$ decreases on $[1, \infty]$.
- Figure 5.9 shows the pattern of $h_1(\alpha)$ and $h_2(\alpha)$ for $\beta_0 = -1$. The signs of the graph are those of $h_1(\alpha) - h_2(\alpha)$.

Hence, $g(\alpha)$ has maximum $g(0) = g(1) = 3$ and a local minimum $g(\alpha_m)$, and decays to zero as α gets larger and larger. In other words, $g(\alpha) \leq 3$ for all $\alpha \in [0, \infty)$, or equivalently, $g(u) \leq 3$ for all $u \in [\beta_0, \infty)$ where $g(u)$ is given by (5.41). Hence, the design of the form (5.34) is D-optimal. ■

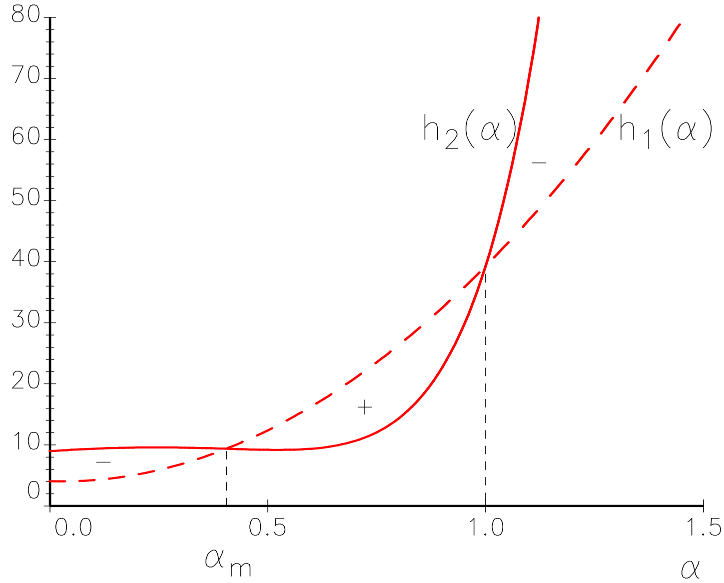


Figure 5.9: Plots of functions $h_1(\alpha)$ (dashed line) and $h_2(\alpha)$ (solid line) respectively given by (5.43) and (5.44) for $\beta_0 = -1$ ($u^* = 1.796$). The sign + and − are those of $h_1(\alpha) - h_2(\alpha)$.

Numerical check of D-optimality of design ξ_z^* given in (5.34)

D-optimality of a design of the form (5.34) could be checked numerically by verifying that the plot of the standardized variance function is less than or equal to the number of model parameters with equality holding at the support points of the candidate optimal design. Figure 5.10 displays the graphs of $g(u)$ versus $u \geq \beta_0 = \{-1.5434, -1, -0.25, 0\}$. Note that $g(u)$ was evaluated at these values of β_0 and the corresponding u^* values that are displayed in Table 5.3. Figure 5.10 indicates that $g(u) \leq 3$ for all $u \geq \beta_0 = \{-1.5434, -1, -0.25, 0\}$, and hence the figure indicates that the 3-point designs ξ_z^* associated with these values of β_0 are D-optimal. Note that unlike the case of the 4-point trapezium design discussed in Section 5.3.2, the local minimum of $g(u)$ is not necessarily attained at $u = 0$ as evidenced by Figure 5.10. Table 5.3, last row, contains the local minimum values $u_m = \beta_0 + \alpha_m(u^* - \beta_0)$ of $g(u)$ obtained by numerically solving equation $g'(u) = 0$ on $[\beta_0, \infty)$ for some selected values of β_0. Only the case $\beta_0 = -1.5434$ leads to a local minimum at $u = 0$. Values of u_m increase from 0 to 0.753 as β_0 increases from -1.5434 to 0

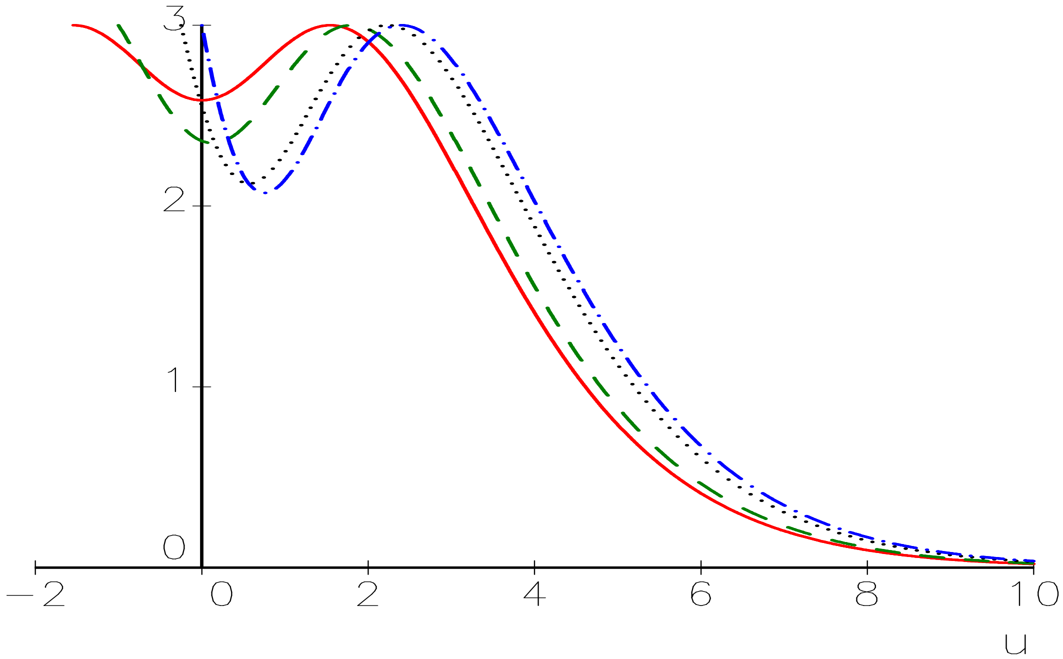


Figure 5.10: Function $g(u)$ given by (5.41) with $\beta_0 = -1.5434$ (solid line), $\beta_0 = -1$ (dashed line), $\beta_0 = -0.25$ (dotted line) and $\beta_0 = 0$ (dashed and dotted line).

Approximate relationship between β_0 and u^*

An approximate equation for the relationship between β_0 and u^* can be found as follows. The scatter plot of the data of Table 5.3 is represented by "circles" in Figure 5.11. The scatter plot indicates that a straight line can fit the data. Fitting a simple linear model to the data gives

$$\widehat{u} = 2.3738 + 0.5538\beta_0 \tag{5.47}$$

with $R^2 = 0.996$. The fitted values of model (5.47) are represented by "plus" in Figure 5.11, and clearly fitted values are not discernable from the observed values.

Finally, Table 5.4 provides the values of u^* solution of (5.38), $\widehat{u}$ given by (5.47) and the efficiencies for selected values of β_0. Table 5.4 reveals that all the efficiencies are close to 1. Therefore, in practical situations there is no loss in efficiency if u estimated using model (5.47) is used instead exact values of u calculated using equation (5.38).

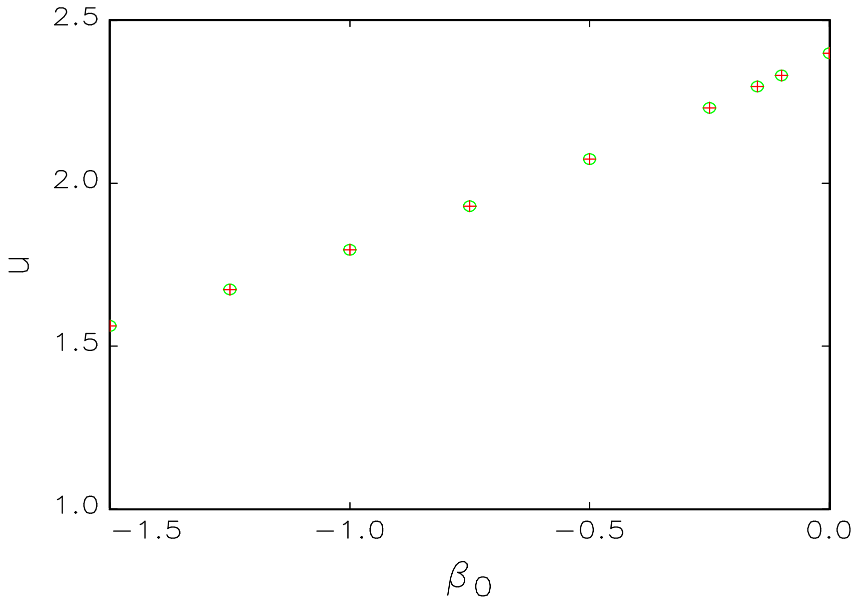


Figure 5.11: Scatter plot of the values of u for given values of β_0: O represent the observed values of u, and + represent the fitted values of u.

Table 5.4: Values of u^*, $\widehat{u}$ and D-efficiencies for selected values of β_0 in model $u = \text{logit}(p) = \beta_0 + z_1 + z_2$ where $z_1 \geq 0$ and $z_2 \geq 0$.

β_0	u^*	u^{**}	D_{eff}
-1.5434	1.5434	1.5191	0.999901
-1.5	1.5618	1.5431	0.999942
-1	1.7960	1.8200	0.999904
-0.75	1.9295	1.9585	0.999862
-0.5	2.0746	2.0969	0.999918
-0.25	2.2313	2.2354	0.999997
0	2.3994	2.3738	0.999890

5.4. Conclusions

In this chapter, the construction of D-optimal designs for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the $(-\infty, \infty) \times [-b, b]$ and the $[0, \infty) \times [0, \infty)$ design spaces as follows.

The design space $(-\infty, \infty) \times [-b, b]$ is the (u_1, u_2)-design space where u_1 is the logit line

$u = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ assumed to take any value on $(-\infty, \infty)$ and $u_2 = b_0 + b_1 x_1 + b_2 x_2$ is a straight line in (x_1, x_2)-space $\mathbb{R}^2$ for a given value of u_2 such that lines $u = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ and $u_2 = b_0 + b_1 x_1 + b_2 x_2$ are not parallel. Two parallel line restrictions were imposed by setting $-b \leq u_2 \leq b$ where $b > 0$, then D-optimal designs were shown to have 4 equally weighted support points located on the intersection of the parallel lines $u_2 = \pm b$ and the two parallel logit lines $u = \pm 1.22291$. This design case corresponds to the 4-point parallelogram design of Sitter and Torsney (1995a) with design locus G_w which is equivalent to the 4-point parallelogram design introduced by Jia and Myers (2001), and a special case of the 4-point design of Atkinson and Haines (1996). This design case is also reported in Myers et al. (2002, p. 240-241).

The design space $[0, \infty) \times [0, \infty)$ is the (d_1, d_2) actual dose. On the design space $[0, \infty) \times [0, \infty)$, two new D-optimal designs, a 4-point trapezium design with support points at the intersection of the boundary of the design space and two parallel logit lines, and an equally weighted 3-point were found under the respective conditions that $-\infty < \beta_0 < -1.5434$ and that $-1.5434 \leq \beta_0 \leq 0$ about the intercept parameter β_0. These designs were proved algebraically and illustrated numerically to be D-optimal. The support points of the trapezium D-optimal designs are points of intersection between the boundary of the design space $[0, \infty) \times [0, \infty)$ and two parallel logits lines $u = \pm u^*$ with u^* positive and not necessary equal to 1.22291. It was found that only support points located on the same logit line have equal weights, and that if β_0 tends to $-\infty$, then u^* tends to 1.22291, which means that the 4-point trapezium design introduced in this thesis tends to the 4-point parallelogram designs of Sitter and Torsney (1995a), also in found in Jia and Myers (2001), Myers et al. (2002), and in Atkinson and Haines (1996). For the 4-point D-optimal trapezium designs, closed expressions of weights associated with the support points were derived algebraically. However, it was not possible to express optimal values of u in closed forms. Instead a regression approach was used to provide approximate D-optimal values u as functions of β_0. The efficiencies of the designs constructed using these approximations relative to the optimal designs were found to be very close to unity for selected values of the parameter β_0. A subset of results found in this chapter are in Haines, Kabera, Ndlovu and O'Brien (2007).

6

D-optimal Designs for the Two-Variable Binary Logistic Model without Interaction: Empirical Results

6.1. Introduction

Chapter 5 discussed the construction of D-optimal designs for the two-variable binary logistic model without interaction

$$u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \tag{6.1}$$

on the design space $[0, \infty) \times [0, \infty)$. The design space $[0, \infty) \times [0, \infty)$ is unbounded in the sense that the upper limit on each axis is not specified. However, in practice such a limit is predetermined leading to a rectangular design space of the form

$$\mathfrak{X} = \{(x_1, x_2) : x_1 \in [a, b] \text{ and } x_2 \in [c, d]\} \tag{6.2}$$

where x_1 and x_2 are either doses or log-doses, and a, b, c and d are real numbers. For example, doses of drugs given to patients generally have predetermined lower and upper limits. This chapter discusses the construction of D-optimal design for the two-variable binary logistic model without interaction (6.1) on the rectangular design space $\mathfrak{X}$ given in (6.2) which, for simplicity of derivations, will be transformed to a square design space. As in Sections 2.4.4 and 5.3.1, it will be assumed the parameter values in model (6.1) satisfy the conditions $\beta_0 < 0$, $\beta_1 > 0$ and $\beta_2 > 0$.

The chapter contains the following sections. Section 6.2 establishes the similarities of the design spaces used by Sitter and Torsney (1995a), Atkinson and Haines (1996) and Jia and

Myers (2001), then provides transformations of one rectangular design space to another rectangular design space. Section 6.3 discusses the construction of a 4-point trapezium and 3-point D-optimal designs on a rectangular design space. Section 6.4 provides cases of non-trapezium D-optimal designs on a rectangular design space. Section 6.5 discusses the construction of 4-point parallelogram D-optimal designs. Section 6.6 discusses the construction of D-optimal designs with support points varying from 4 to 6. Section 6.7 provides a summary chart of the designs patterns of the D-optimal designs for model (6.1) on the design space $[0, b] \times [0, b]$ satisfying condition $2\beta_0 + b(\beta_1 + \beta_2) = 0$. Section 6.8 discusses the construction of 4-point non-parallelogram designs. Section 6.9 gives a practical example of the application of some of the design theory in this chapter. Finally, Section 6.10 contains the conclusions for this chapter.

All candidate D-optimal designs for model (6.1) discussed throughout this chapter are classified as related to the trapezium design, or to the parallelogram, except cases when the number of support points is greater than 4. D-optimal designs for simple cases are constructed analytically using the Equivalence Theorem for D-optimality. In complex cases the patterns of the candidate D-optimal designs are conjectured under specific conditions, and then the D-optimal designs are constructed numerically using the Gauss program given in Appendix B. D-optimality of the designs is checked by plotting the standardized variance function which is p *minus the directional derivative function* on the design space, where p is the number of model parameters (see Atkinson et al. (2007, p. 124)).

6.2. Preliminaries

Consider the two-variable binary logistic model without interaction (6.1) where the explanatory variables x_1 and x_2 are elements of the rectangular design space $\mathfrak{X}$ given by (6.2), and are considered as dose or log-dose concentrations of two drugs. Candidate D-optimal designs for model (6.1) were constructed by Sitter and Torsney (1995a), Atkinson and Haines (1996), and Jia and Myers (2001) but with different designs spaces. The candidate D-optimal designs of Jia and Myers (2001) for model (6.1) can also be found in Myers et al. (2002).

The D-optimal designs of this chapter will be constructed on the (x_1, x_2) design space with $a \leq x_1 \leq b$ and $c \leq x_2 \leq d$. The (x_1, x_2) design space $[a, b] \times [c, d]$ underlie the (u_1, u_2) design space used by Sitter and Torsney (1995a), and (x_1, x_2) designs spaces used by Atkinson and Haines (1996), and Jia and Myers (2001). In fact, consider the two-variable binary logistic model $u = \text{logit}(p) = u_1 = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ and a straight line of the $u_2 = b_0 + b_1 x_1 + b_2 x_2$

for fixed u_2 where $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$ and $\frac{b_1}{b_2} \neq \frac{\beta_1}{\beta_2}$. If $a \leq x_1 \leq b$ and $c \leq x_2 \leq d$, then $\widetilde{a} \leq u_1 \leq \widetilde{b}$, where $\widetilde{a} = \beta_0 + \beta_1 a + \beta_2 c$ and $\widetilde{b} = \beta_0 + \beta_1 b + \beta_2 d$. Furthermore, the restriction $c \leq x_2 \leq d$, can be written as $c \leq u_2 = x_2 \leq d$ where $b_0 = b_1 = 0$ and $b_2 = 1$ in $u_2 = b_0 + b_1 x_1 + b_2 x_2$. Hence, there is a one-to-one correspondence between a D-optimal designs in terms of $(x_1, x_2) \in [a, b] \times [c, d]$ and $(u_1, u_2) \in [\widetilde{a}, \widetilde{b}] \times [c, d]$.

Sitter and Torsney (1995a) considered two design cases in constructing D-optimal designs for model (6.1). The first design case was reviewed in Section 4.5.1 in which u_1 is unrestricted, but u_2 restricted and this case was shown in Section 5.2 to be the same as the design case used by Jia and Myers (2001). The candidate D-optimal design of Sitter and Torsney (1995a) with u_1 unrestricted and Jia and Myers (2001) was an equally weighted 4-point design with support points on the logit lines $u^* = \pm 1.22291$. The second design case of Sitter and Torsney (1995a) reviewed in Section 4.5.1 was also a 4-point D-optimal design in the (u_1, u_2)-space, but with both u_1 and u_2 restricted. Four candidate D-optimal designs were identified, but the D-optimality proved in Section 5.2 was for only one of those cases. The construction of the D-optimal designs for the other cases is discussed in subsequent sections of this chapter. The main difficulty encountered in the construction of these designs is due to the dependence of the candidate D-optimal designs on the model parameters and on the design space.

Atkinson and Haines (1996) considered numerical construction of D-optimal designs for model (6.1) on the (x_1, x_2) design space $[-1, 1] \times [-1, 1]$. Clearly, the design space $[-1, 1] \times [-1, 1]$ is the design space (6.2) with $a = c = -1$ and $b = d = 1$.

For ease of calculations, it is usual to transform a rectangular design space such as $[a, b] \times [c, d]$ to another rectangular design space.

Generally, a linear transformation from one rectangular design space to another rectangular or square design space can be constructed as follows. Consider the model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$, and the transformation from the (x_1, x_2) design space $[a, b] \times [c, d]$ to the (x_{1new}, x_{2new}) design space $[a_1, b_1] \times [c_1, d_1]$. The relationships between points (x_1, x_2) and (x_{1new}, x_{2new}), and also between model parameters in design spaces $[a, b] \times [c, d]$ and $[a_1, b_1] \times [c_1, d_1]$ are

$$\begin{cases} x_{1new} = \frac{(b_1 - a_1)x_1 - (ab_1 - a_1 b)}{b - a} \\ x_{2new} = \frac{(d_1 - c_1)x_2 - (cd_1 - c_1 d)}{d - c} \end{cases} \tag{6.3}$$

and

$$\begin{cases} \beta_{1new} = \frac{b-a}{b_1-a_1}\beta_1 \\ \beta_{2new} = \frac{d-c}{d_1-c_1}\beta_2 \\ \beta_{0new} = \beta_0 + \frac{ab_1-a_1b}{b_1-a_1}\beta_1 + \frac{cd_1-c_1d}{d_1-c_1}\beta_2. \end{cases} \tag{6.4}$$

In order to link results of the subsequent sections of this chapter with those in Section 5.2 and Section 5.3, the (x_1, x_2) design space $[a, b] \times [c, d]$ that will be used is the square design space $[0, b] \times [0, b]$ obtained by setting $a = c = 0$ and $b = d$. If required, the transformations (6.3) and (6.4) will be used to transform the designs from the (x_1, x_2) design space $[0, b] \times [0, b]$ to the (x_{1new}, x_{2new}) design space $[a, b] \times [c, d]$.

6.3. 4- and 3-point trapezium designs

6.3.1 4-point trapezium designs

The trapezium design was discussed in Section 5.3 where the design space was $[0, \infty) \times [0, \infty)$. However, on the design space $[0, b] \times [0, b]$, the existence of a trapezium design depends on the location of the optimal logit lines $\pm u^*$ relative to the vertices of the square $[0, b] \times [0, b]$. Figure 6.1 (a) and Figure 6.1 (b) display the design patterns of candidate D-optimal designs for arbitrary $\pm u = \beta_0 + \beta_1 x_2 + \beta_2 x_2$.

Consider Figure 6.1 (a). Note that the logit line CD is below the diagonal line FH of the square whose vertices are E, F, G and H, and the logit line AB is above the vertex E. In addition, and more importantly note that the position of logit $u = 0$ in Figure 6.1 (a) implies that the logit values at points $F(0, b)$ and $H(b, 0)$ are both positive, i.e. respectively $\beta_0 + \beta_2 b > 0$ and $\beta_0 + \beta_1 b > 0$. This design pattern is similar to the design pattern of Figure 5.4, and hence the 4-point D-optimal trapezium design exists if the following conditions are satisfied: $-\infty < \beta_0 < -1.5434$, $\beta_0 + \beta_1 b \geq 1.5434$ and $\beta_0 + \beta_2 b \geq 1.5434$, or, equivalently $1.5434 < -\beta_0 < \infty$ and $1.5434 \leq \min\{\beta_0 + \beta_1 b, \beta_0 + \beta_2 b\}$ which implies that

$$1.5434 \leq \min\{-\beta_0, \beta_0 + \beta_1 b, \beta_0 + \beta_2 b\}. \tag{6.5}$$

In this case, the D-optimal design is of the form

$$\xi^* = \left\{ \begin{array}{cccc} \left(\frac{-u^*-\beta_0}{\beta_1}, 0\right) & \left(0, \frac{-u^*-\beta_0}{\beta_2}\right) & \left(0, \frac{u^*-\beta_0}{\beta_2}\right) & \left(\frac{u^*-\beta_0}{\beta_1}, 0\right) \\ \frac{1-\lambda^*}{2} & \frac{1-\lambda^*}{2} & \frac{\lambda^*}{2} & \frac{\lambda^*}{2} \end{array} \right\} \tag{6.6}$$

where for a given value of β_0, optimal values of λ^* and u^* are solutions to the simultaneous equations (5.13) and (5.18). D-optimality of a 4-point trapezium design was proved in

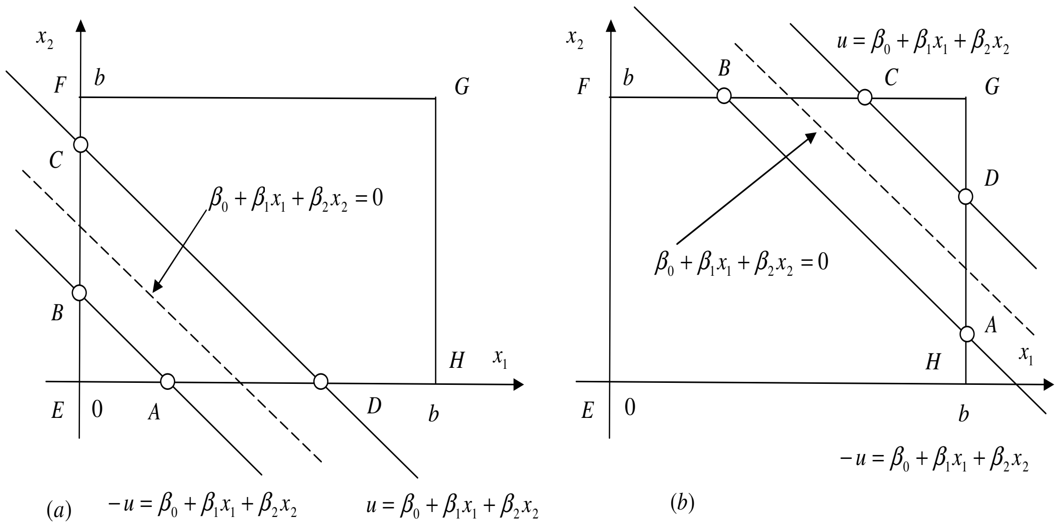


Figure 6.1: Design patterns for the candidate D-optimal trapezium designs for the two-variable binary logistic model without interaction $u = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the dose space $\mathcal{D} = [0, b] \times [0, b]$: (a) $1.5434 \leq \min\{-\beta_0, \beta_0 + \beta_1 b, \beta_0 + \beta_2 b\}$ and (b) $1.5434 \leq \min\{-\beta_0 - \beta_1 b, -\beta_0 - \beta_2 b, \beta_0 + (\beta_1 + \beta_2)b\}$. Circles represent support points.

Theorem 5.3.

Now, consider Figure 6.1 (b). The D-optimal trapezium design of this figure can be taken as a symmetric reflection of the design of Figure 6.1 (a) with respect to the diagonal line FH. Especially the position of the logit $u = 0$ implies that the logit values $\beta_0 + \beta_2 b$ and $\beta_0 + \beta_1 b$ respectively at points $F(0, b)$ and $H(b, 0)$ are now both negative. Reasoning as for the design of Figure 6.1 (a), a 4-point D-optimal trapezium design is obtained if

$$1.5434 \leq \min\{-\beta_0 - \beta_1 b, -\beta_0 - \beta_2 b, \beta_0 + (\beta_1 + \beta_2)b\}. \tag{6.7}$$

The D-optimal trapezium design is of the form

$$\xi^* = \left\{ \begin{array}{cccc} \left(b, \frac{-u-\beta_0-\beta_1 b}{\beta_2}\right) & \left(\frac{-u-\beta_0-\beta_2 b}{\beta_1}, b\right) & \left(\frac{u-\beta_0-\beta_2 b}{\beta_1}, b\right) & \left(b, \frac{u-\beta_0-\beta_1 b}{\beta_2}\right) \\ \frac{\lambda^*}{2} & \frac{\lambda^*}{2} & \frac{1-\lambda^*}{2} & \frac{1-\lambda^*}{2} \end{array} \right\}. \tag{6.8}$$

Observe that the results obtained from the design of Figure 6.1 (a) can be used to calculate the design of Figure 6.1 (b) since the design of Figure 6.1 (b) is a reflection of the design of Figure 6.1 (a) about line FH. In fact, if (x_1, x_2) are coordinates of a support point of the D-optimal trapezium design in Figure 6.1 (b) and (y_1, y_2) are corresponding coordinates of a support point of the D-optimal trapezium design in Figure 6.1 (a), these coordinates are linked by the following transformation

$$(x_1, x_2) = (b - y_2, b - y_1) \tag{6.9}$$

Suppose that u is the logit value in the design of Figure 6.1 (b). If $\widetilde{u}$ is the corresponding logit value in the design of Figure 6.1 (a) then the transformation (6.9) takes $\widetilde{u}$ to $-u$ and vice versa. Specifically assume that $u = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ and that $\widetilde{u} = \beta_0' + \beta_1' y_1 + \beta_2' y_2$. The transformation (6.9) gives

$$\begin{aligned}
\beta_0' + \beta_1' y_1 + \beta_2' y_2 = \widetilde{u} &= -u \\
&= -\beta_0 - \beta_1 x_1 - \beta_2 x_2 \\
&= -\beta_0 - \beta_1(b - y_2) - \beta_2(b - y_1) \\
&= -\beta_0 - b(\beta_1 + \beta_2) + \beta_2 y_1 + \beta_1 y_2
\end{aligned}$$

which implies that

$$\begin{cases} \beta_0' = -\beta_0 - (\beta_1 + \beta_2)b \\ \beta_1' = \beta_2 \\ \beta_2' = \beta_1. \end{cases} \tag{6.10}$$

The following example illustrates a situation where the parameter values in model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ lead to either the D-optimal trapezium design (6.6) or the D-optimal trapezium design (6.8).

Example 6.1. Consider the two-variable binary logistic models without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ with parameter values $\beta_1 = \beta_2 = 2$, and a) $\beta_0 = -2$ and (b) $\beta_0 = -6$ on the design space $[0, 2] \times [0, 2]$.

In part (a), condition (6.5) is satisfied since

$$1.5434 < \min\{-\beta_0 = 2, \beta_0 + \beta_1 b = 2, \beta_0 + \beta_2 b = 2\} = 2.$$

Hence, the appropriate D-optimal design is the 4-point D-optimal trapezium design (6.6), i.e.

$$\xi^* = \left\{ \begin{array}{cccc} (0.263, 0) & (0, 0.263) & (0, 1.737) & (1.737, 0) \\ 0.1686 & 0.1686 & 0.3314 & 0.3314 \end{array} \right\} \tag{6.11}$$

since $u^* = 1.4741$ for $\beta_0 = -2$ (see Table 5.1). If the design space $[-1, 1] \times [-1, 1]$ was used instead of the design space $[0, 2] \times [0, 2]$, the transformations (6.3) and (6.4) give

$$\begin{cases} x_{1new} = x_1 - 1 \\ x_{2new} = x_2 - 1 \end{cases}$$

and

$$\begin{cases} \beta_{1new} = 2 \\ \beta_{2new} = 2 \\ \beta_{0new} = 2. \end{cases}$$

Hence, the D-optimal trapezium design (6.11) becomes

$$\xi^* = \left\{ \begin{array}{cccc} (-0.737, 0) & (0, -0.737) & (0, 0.737) & (0.737, 0) \\ 0.1686 & 0.1686 & 0.3314 & 0.3314 \end{array} \right\}$$

which is the same as the D-optimal design found numerically by Atkinson and Haines (1996) for $\boldsymbol{\beta} = (2, 2, 2)^T$.

The parameters of the model in part (b) satisfy condition (6.7) since

$$1.5434 < \min\{-\beta_0 - \beta_1 b = 2, -\beta_0 - \beta_2 b = 2, \beta_0 + (\beta_1 + \beta_2)b = 2\} = 2.$$

In addition, the parameters of the model in part (b) are linked to those of the model in part (a) by the transformation (6.10) by writing the parameters of the model in part (a) as $\beta_0' = -2$ and $\beta_1' = \beta_2' = 2$. Hence, the D-optimal design for the parameters of the model in part (b) obtained from the 4-point D-optimal trapezium design (6.11) using the transformation (6.9) is

$$\xi^* = \left\{ \begin{array}{cccc} (2, 1.737) & (1.737, 2) & (2, 0.263) & (0.263, 2) \\ 0.1686 & 0.1686 & 0.3314 & 0.3314 \end{array} \right\} \tag{6.12}$$

6.3.2 3-point designs

If in Figure 6.1 (a), $A = B = E$, then the design pattern is similar to the design pattern in Figure 5.8 (a), and hence an equally weighted 3-point trapezium related D-optimal design exists if $0 \le -\beta_0 \le 1.5434$ and $1.5434 \le \min\{\beta_0 + \beta_2 b, \beta_0 + \beta_1 b\}$, which is equivalent to

$$0 \le -\beta_0 \le 1.5434 \le \min\{\beta_0 + \beta_1 b, \beta_0 + \beta_2 b\}. \tag{6.13}$$

As model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ has 3 parameters, for given values of the parameters a 3-point design puts weight $\frac{1}{3}$ at each of its support points (Silvey (1980, p.42). In this case, the D-optimal design is of the form

$$\xi^* = \left\{ \begin{array}{ccc} (0, 0) & \left(0, \frac{u^* - \beta_0}{\beta_2}\right) & \left(\frac{u^* - \beta_0}{\beta_1}, 0\right) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array} \right\} \tag{6.14}$$

where u^* is the solution for u to the equation (5.38) for a given value of β_0. D-optimality of the design was proved in Theorem 5.4.

If in Figure 6.1 (b), $A = B = G$, then the parameters β_0, β_1 and β_2 satisfy the condition

$$0 \leq \beta_0 + (\beta_1 + \beta_2)b \leq 1.5434 \leq \min\{-\beta_0 - \beta_2 b, -\beta_0 - \beta_1 b\}. \tag{6.15}$$

Hence, as for the case of the 4-point D-optimal trapezium design, the 3-point D-optimal design is obtained from the D-optimal design (6.14) using the transformations (6.9) to give the D-optimal design

$$\xi^* = \left\{ \begin{array}{ccc} \left(b, \frac{-u^*-\beta_0-\beta_1 b}{\beta_2}\right) & \left(\frac{-u^*-\beta_0-\beta_2 b}{\beta_1}, b\right) & (b, b) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array} \right\} \tag{6.16}$$

The following example illustrates cases where parameter values lead to a 3-point design on the design space $[0, 2] \times [0, 2]$.

Example 6.2. Consider the two-variable binary logistic model without interaction $u = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, 2] \times [0, 2]$ and with $\beta_1 = 3$, $\beta_2 = 2$ and (a) $\beta_0 = -1$ and (b) $\beta_0 = -9$.

In part (a)

$$0 < -\beta_0 = 1 < 1.5434 < \min\{\beta_0 + \beta_1 b = 5, \beta_0 + \beta_2 b = 3\} = 3.$$

Therefore condition (6.13) is satisfied so that a 3-point design of the form (6.14) is appropriate to this example. Table 5.3 gives $u^* = 1.7960$ for $\beta_0 = -1$. Thus, the design (6.14) is

$$\xi^* = \left\{ \begin{array}{ccc} (0, 0) & (0, 1.398) & (0.932, 0) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array} \right\}. \tag{6.17}$$

In part (b), consider following transformation (6.10) by writing the parameters of the model in part (a) as $\beta_0' = -1$, $\beta_1' = 3$ and $\beta_2' = 2$. The parameters $\beta_0 = -9$, $\beta_1 = 3$ and $\beta_2 = 2$ lead to a 3-point design since they satisfy condition (6.15). In fact,

$$0 < \beta_0 + \beta_1 b + \beta_2 b = 1 < 1.5434 < \min\{-\beta_0 - \beta_1 b = 3, -\beta_0 - \beta_2 b = 5\} = 3.$$

Hence, the 3-point D-optimal design for this case can be obtained using the transformation (6.9) on the 3-point D-optimal design (6.17) to give

$$\xi^* = \left\{ \begin{array}{ccc} (2, 0.602) & (1.068, 2) & (2, 2) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array} \right\}. \tag{6.18}$$

6.4. Some non-trapezium candidate D-optimal designs

The trapezium designs described in Section 6.3 have support points which either are all above or all below the main diagonal of the design space $[0, b] \times [0, b]$ (see Figures 6.1 (a) and (b)). For some values of the vector $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^T$ and the bound b, the conditions for the existence of the D-optimal designs with these patterns may not always be satisfied, i.e. D-optimal trapezium designs with design patterns in Figures 6.1 (a) and (b) may not exist. In such cases, the alternative designs to be sought are the non-trapezium D-optimal designs which are discussed in this section. A few design patterns for the 3-point and 4-point non-trapezium designs are considered for illustration purposes.

6.4.1 4-point non-trapezium designs

Case of Figure 6.2 (a) pattern

Consider the design pattern of the candidate 4-point D-optimal design displayed in Figure 6.2 (a). The design pattern of Figure 6.2 (a) verify necessary conditions of a trapezium design on the design space $[0, b] \times [0, b]$ since $\beta_0 + \beta_1 b < 0$ and $\beta_0 + \beta_2 b < 0$. However, this condition is not sufficient to obtain a trapezium design because the logit line JD in Figure 6.2 does not lie completely below the diagonal line FH. Without any restrictions on x_2 the support points of the design can be A, I, J and D, i.e. point J is outside the design space $[0, b] \times [0, b]$. The restriction $0 \leq x_2 \leq b$ moves point J to F, and possibly changes the positions of A and D in the x_1 direction, and the position of I in the x_2 direction. This results in a non-trapezium design with support points A, I, F, D, and for which point I is not necessarily on line AB. Furthermore, the distribution of the weights among the support points is not expected to be systematic as in the case of trapezium designs. This makes the analytic construction of non-trapezium D-optimal designs difficult. However, the D-optimal designs can be constructed numerically for known values of the parameters β_0, β_1 and β_2 in the model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$. In addition, the condition for design of Figure 6.2 (a) to exist is easy to establish. In fact, if the design pattern of Figure 6.2 (a) is possible, then the following can be concluded. The vertex $(0, 0)$ is not a support point of the candidate D-optimal design and $-\infty < \beta_0 < -1.5434$ (see Theorem 5.2). Since point $(0, b)$ is above the logit line $u = 0$ and point J is above the main diagonal of the design space $[0, b] \times [0, b]$, then $0 < \beta_0 + \beta_2 b \leq 1.5434 \leq \beta_0 + \beta_1 b$. Hence, overall, the condition for the 4-point candidate

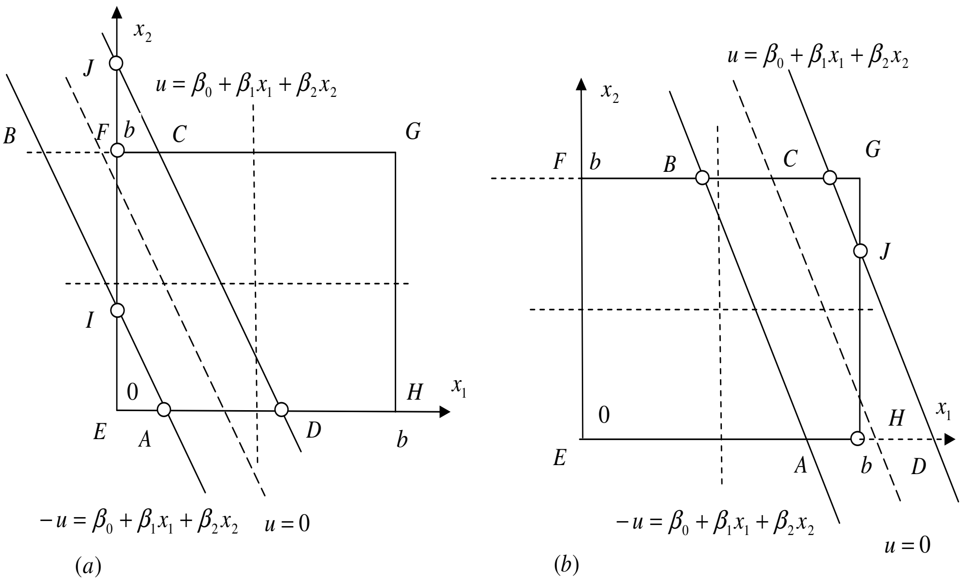


Figure 6.2: Design patterns for candidate 4-point D-optimal non-trapezium designs for the two-variable binary logistic model without interaction $u=\beta_0+\beta_1x_1+\beta_2x_2$ on the design space $[0,b]\times[0,b]$ with (a) $0\le\beta_0+\beta_2b\le 1.5434\le\min\{-\beta_0,\beta_0+\beta_1b\}$ and (b) $0\le-\beta_0-\beta_1b\le 1.5434\le\min\{-\beta_0-\beta_2b,\beta_0+(\beta_1+\beta_2)b\}$.

non-trapezium D-optimal design of Figure 6.2 (a) to exist is

$$0<\beta_0+\beta_2b\le 1.5434\le\min\{-\beta_0,\beta_0+\beta_1b\}. \tag{6.19}$$

Note that one of the optimal logits is at point F, i.e. $u=\beta_0+\beta_2b$. The following example illustrates a case where parameter values lead to the non-trapezium design of Figure 6.2 (a).

Example 6.3. Consider the two-variable binary logistic model $u=\text{logit}(p)=\beta_0+\beta_1x_1+\beta_2x_2$ on the design space $[0,2]\times[0,2]$ where $\beta_0=-3$, $\beta_1=3$ and $\beta_2=2$.

In this case $0<\beta_0+\beta_2b=1<1.5434<\min\{-\beta_0=3,\beta_0+\beta_1b=3\}=3$. Thus, condition (6.19) is satisfied and the D-optimal design exists with support points as in Figure 6.2 (a). One of the optimal logits is $u=\beta_0+b\beta_2=1$. The other optimal logits have to be found numerically. Numerical optimization gives the following design

$$\xi^*=\left\{\begin{array}{c}(x_1,x_2)\\ \lambda\\ u\end{array}\right\}=\left\{\begin{array}{cccc}(0.507,0) & (0,0.767) & (0,2) & (1.493,0)\\ 0.2653 & 0.0815 & 0.3262 & 0.3270\\ -1.478 & -1.466 & 1 & 1.478\end{array}\right\}. \tag{6.20}$$

This design highlights the fact that $(0.507,0)$ and $(0,0.767)$ are not on the same logit line and thus do not have equal weights. Clearly, the support points of the design (6.20) displayed in

Figure 6.3 (a) are not vertices of a trapezium. The standardized variance function, $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$,

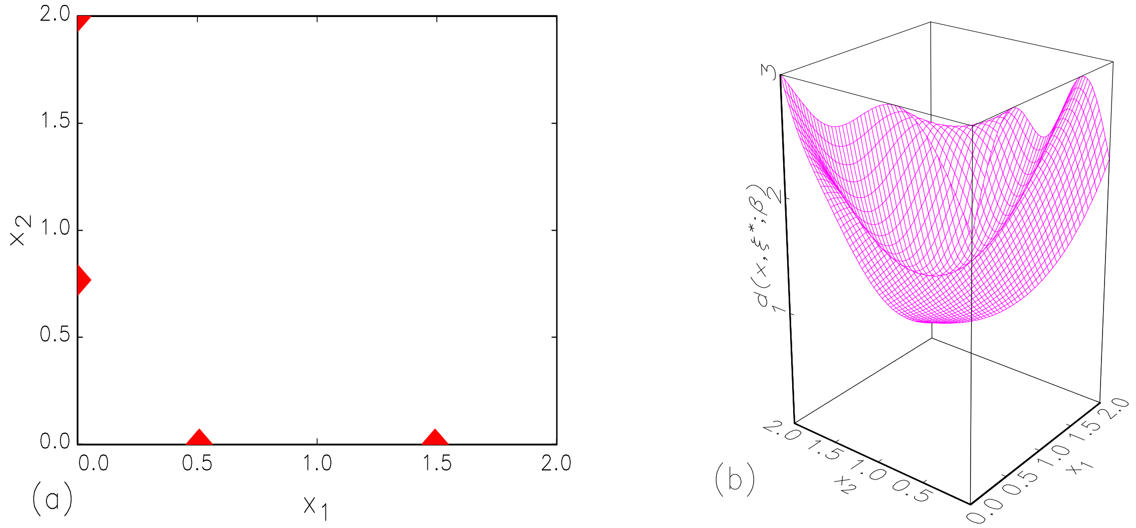


Figure 6.3: (a) Support points of the 4-point non-trapezium D-optimal design and (b) standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, 2] \times [0, 2]$ and with parameter vector $\boldsymbol{\beta} = (-3, 3, 2)^T$.

is plotted in Figure 6.3 (b). The graph indicates that at least numerically $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) = 3$ at the 4 support points of the design (6.20) and that $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) < 3$ everywhere else on the design space $[0, 2] \times [0, 2]$.

Case of Figure 6.2 (b) pattern

Now consider Figure 6.2 (b). The displayed design pattern is a reflection of the design pattern of Figure 6.2 (a) about the diagonal line FH. Hence the comments are similar to those introduced for the design pattern of Figure 6.2 (a), but the signs of $\beta_0 + \beta_1 b$ and $\beta_0 + \beta_2 b$ have changed from positive to negative. The support points of the candidate D-optimal design have locations similar to vertices of the quadrilateral $BCJH$ but J is not necessarily on the line CD. Also, one of the optimal logits is $u = \beta_0 + b\beta_1$. The condition for the design of Figure 6.2 (b) to exist is

$$0 < -\beta_0 - \beta_1 b \leq 1.5434 \leq \min\{-\beta_0 - \beta_2 b, \beta_0 + (\beta_1 + \beta_2)b\}. \tag{6.21}$$

The following example is a case where parameter values lead to a non-trapezium design of the form presented in Figure 6.2 (b).

Example 6.4. Consider the two-variable binary logistic model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, 2] \times [0, 2]$ where $\beta_0 = -7$, $\beta_1 = 3$ and $\beta_2 = 2$.

Condition (6.21) holds since

$$0 < -\beta_0 - \beta_1 b = 1 < 1.5434 < \min\{-\beta_0 - \beta_2 b = 3, \beta_0 + (\beta_1 + \beta_2)b = 3\} = 3.$$

Therefore, $\beta_0 + 2\beta_2 = -1$ is one of the optimal logits and the others can be calculated numerically to give the following candidate D-optimal design

$$\xi^* = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{cccc} (2,0) & (0.507, 2) & (1.493, 2) & (2, 1.233) \\ 0.3262 & 0.3270 & 0.2653 & 0.0815 \\ -1 & -1.478 & 1.478 & 1.466 \end{array} \right\}. \qquad (6.22)$$

Clearly, the support points of design (6.22) displayed in Figure 6.4 (a) are not vertices of a trapezium.

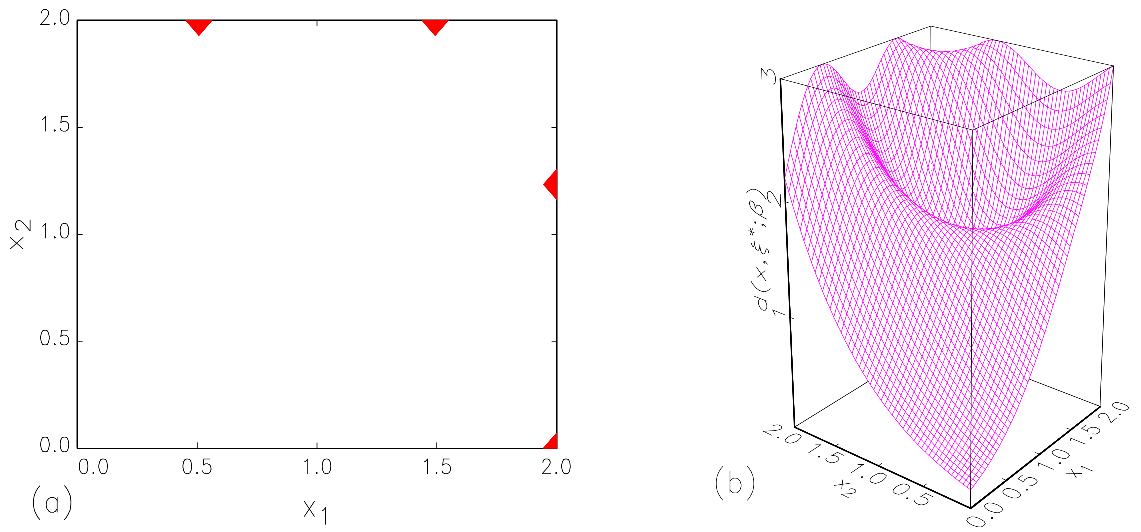


Figure 6.4: (a) Support points of the 4-point non-trapezium D-optimal design and (b) standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, 2] \times [0, 2]$ and with parameter vector $\boldsymbol{\beta} = (-7, 3, 2)^T$.

The standardized variance function, $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$, is plotted in Figure 6.4 (b), and the graph suggests that design (6.22) is D-optimal.

6.4.2 3-point non-trapezium related designs

The construction of the 3-point D-optimal design on the design spaces $[0, \infty) \times [0, \infty)$ and $[0, b] \times [0, b]$ was discussed in Section 5.3.3 and Section 6.3, respectively. There exists other cases of 3-point D-optimal designs on $[0, b] \times [0, b]$, and these are briefly discussed below.

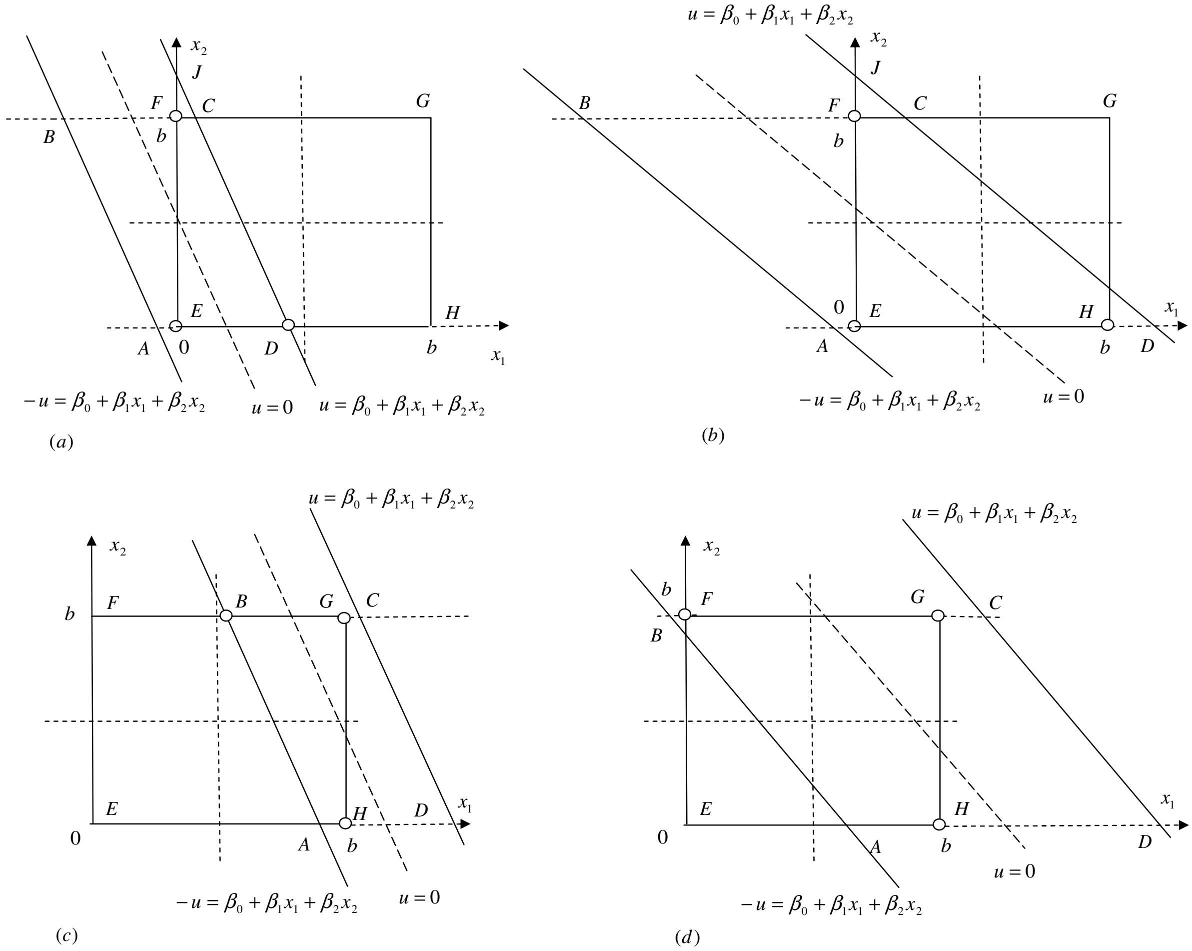


Figure 6.5: Sample patterns of 3-point D-optimal designs for the two-variable binary logistic model without interaction $u = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, b] \times [0, b]$.

Case of Figure 6.5 (a) pattern

Consider the design pattern displayed in Figures 6.5 (a). Clearly, the location of the logit $u = 0$ and the points $F(0, b)$ and $H(b, 0)$ in Figure 6.5 (a) imply that $\beta_0 + \beta_1 b > 0$ and $\beta_0 + \beta_2 b > 0$ so that the necessary conditions of a trapezium design on the design space $[0, b] \times [0, b]$ are satisfied. Without the restrictions $x_1 \geq 0$ and $x_2 \geq 0$, supports A and B can be outside the design space $[0, b] \times [0, b]$. The restriction of the design space $[0, b] \times [0, b]$ and the location of logit $u = 0$ moves point J to F and line AB degenerates to point A. Hence, the candidate D-optimal design has support points E, D and F. The conditions for the existence of the candidate D-optimal design are derived as follows. As point E at $(0, 0)$ is a support of the candidate D-optimal design, then $-1.5434 \leq \beta_0 \leq 0$ (see Theorem 5.4). Furthermore, since point F at $(0, b)$ is above logit $u = 0$ and point J is above the main diagonal of the design space $[0,b]\times[0,b]$, then $0<\beta_0+\beta_2 b\le 1.5434\le\beta_0+\beta_1 b$. Hence, overall the design of Figure 6.5 (a) must satisfy the condition

$$0<\ \max\{-\beta_0,\beta_0+b\beta_2\}\le 1.5434\le\beta_0+\beta_1 b. \tag{6.23}$$

The optimal logits at support points E and F are $u=\beta_0$ and $u=\beta_0+\beta_2 b$, respectively. Hence, the construction of the D-optimal design is reduced to searching for the optimal logit at point D. The form of the candidate D-optimal design is

$$\xi^*=\left\{\begin{array}{c}(x_1,x_2)\\ \lambda\\ u^*\end{array}\right\}=\left\{\begin{array}{ccc}(0,0) & (0,b) & (u,0)\\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3}\\ \beta_0 & \beta_0+\beta_2 b & u^*\end{array}\right\} \tag{6.24}$$

where u^* is the optimal logit. The information matrix for $\boldsymbol{\beta}=(\beta_0,\beta_1,\beta_2)^T$ (4.19) evaluated at design (6.24) is

$$M(\xi;\boldsymbol{\beta})=\frac{1}{3}\begin{bmatrix}\frac{e^{\beta_0+b\beta_2}}{\left(1+e^{\beta_0+b\beta_2}\right)^2}+\frac{e^u}{(1+e^u)^2}+\frac{e^{\beta_0}}{\left(1+e^{\beta_0}\right)^2} & \frac{e^u(u-\beta_0)}{3\beta_1(1+e^u)^2} & \frac{be^{\beta_0+b\beta_2}}{3\left(1+e^{\beta_0+b\beta_2}\right)^2}\\ \frac{e^u(u-\beta_0)}{3\beta_1(1+e^u)^2} & \frac{e^u(\beta_0-u)^2}{3\beta_1^2(1+e^u)^2} & 0\\ \frac{ae^{\beta_0+b\beta_2}}{3\left(1+e^{\beta_0+b\beta_2}\right)^2} & 0 & \frac{b^2e^{\beta_0+b\beta_2}}{3\left(1+e^{\beta_0+b\beta_2}\right)^2}\end{bmatrix} \tag{6.25}$$

and its determinant is

$$D=\frac{b^2(u-\beta_0)^2e^{u+2\beta_0+b\beta_2}}{27\beta_1^2(1+e^{\beta_0})^2(1+e^{\beta_0+b\beta_2})^2(1+e^u)^2}. \tag{6.26}$$

The determinant (6.26) is proportional to the square root of determinant (5.36) obtained for the case of the 3-point design on the design space $[0,\infty]\times[0,\infty)$. Differentiating (6.26) with respect to u gives

$$\frac{\partial D}{\partial u}=\frac{b^2e^{2\beta_0+b\beta_2+u}(u-\beta_0)(2-\beta_0+u+2e^u+\beta_0e^u-ue^u)}{27\beta_1^2(1+e^{\beta_0})^2(1+e^{\beta_0+b\beta_2})^2(1+e^u)^3}. \tag{6.27}$$

Equating (6.27) to zero and solving the equation for u gives $u=\beta_0$ or u which satisfies the equation

$$2-\beta_0+u+2e^u+\beta_0e^u-ue^u=0. \tag{6.28}$$

Equation (6.28) is the same as equation (5.38). It was shown in Section 5.3.3 that the equation (5.38) has a unique solution for u on $[0,\infty)$, but it is also easy to show that it has a unique solution on $[\beta_0,\infty)$ where $\beta_0<0$ and $u\ge\beta_0$. In fact, consider the continuous function $f(u)=2-\beta_0+u+2e^u+\beta_0e^u-ue^u$ on $[\beta_0,\infty)$. Clearly, $f(\beta_0)=2+2^{\beta_0}>0$ and $\lim_{u\to\infty}f(u)=-\lim_{u\to\infty}ue^u=-\infty$, i.e. $f(u)=0$ has at least one solution on $[\beta_0,\infty)$. In addition,

$f''(u) = (\beta_0 - u)e^u < 0$ since $u \geq 0$ and $\beta_0 < 0$. Hence, $f(u)$ is concave on $[0, \infty)$ and thus, $f(u) = 0$ or equation (6.28) has a unique solution for u on $[\beta_0, \infty)$.

Hence, the proof of D-optimality of design (6.24) follows from Theorem 5.4. The following is an illustration of parameter values which lead to a design of the form (6.24).

Example 6.5. Consider the two-variable binary logistic model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, 2] \times [0, 2]$ where $\beta_0 = -1$, $\beta_1 = 3$ and $\beta_2 = 1$.

Condition (6.23) holds since

$$0 < \max\{-\beta_0, \beta_0 + b\beta_2\} = 1 \leq 1.5434 \leq \beta_0 + \beta_1 b = 5.$$

Hence, the candidate D-optimal design is (6.24) where u^* is the solution to equation (6.28). Since $\beta_0 = -1$, solving equation (6.28), or using Table 5.3, gives $u^* = 1.796$, and hence $\dfrac{u^* - \beta_0}{\beta_1} = 0.932$. The two other optimal logits are $\beta_0 = -1$ and $\beta_0 + b\beta_2 = 1$. Hence, the candidate D-optimal design (6.24) is

$$\xi^* = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{ccc} (0,0) & (0,2) & (0.932, 0) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ -1 & 1 & 1.796 \end{array} \right\}. \tag{6.29}$$

The support points of the design (6.29) are displayed in Figure 6.6 (a). The standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ is plotted in Figure 6.26 (b), and the graph suggests that $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) \leq 3$ with equality holding only at each of the support points of the design (6.29). Thus, the design (6.29) is shown numerically to be D-optimal.

If the design space $[-1, 1] \times [-1, 1]$ was used instead of $[0, 2] \times [0, 2]$, transforming the parameter vector $\boldsymbol{\beta} = (-1, 3, 1)^T$ using (6.4) gives $\boldsymbol{\beta}' = (3, 3, 1)^T$ and hence, the D-optimal design (6.29) becomes

$$\xi^* = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{ccc} (-1,-1) & (-1,1) & (-0.068, 0) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ -1 & 1 & 1.796 \end{array} \right\}. \tag{6.30}$$

Atkinson and Haines (1996) used a numerical approach with design space $[-1, 1] \times [-1, 1]$ and parameter vector $\boldsymbol{\beta} = (3, 3, 1)^T$ and found the same D-optimal design as (6.30).

Case of Figure 6.5 (b) design pattern

Now consider Figure 6.5 (b). As in the case of the design pattern of Figure 6.5 (a), $\beta_0 + \beta_1 b > 0$ and $\beta_0 + \beta_2 b > 0$ so that the necessary conditions of a trapezium design on the design space

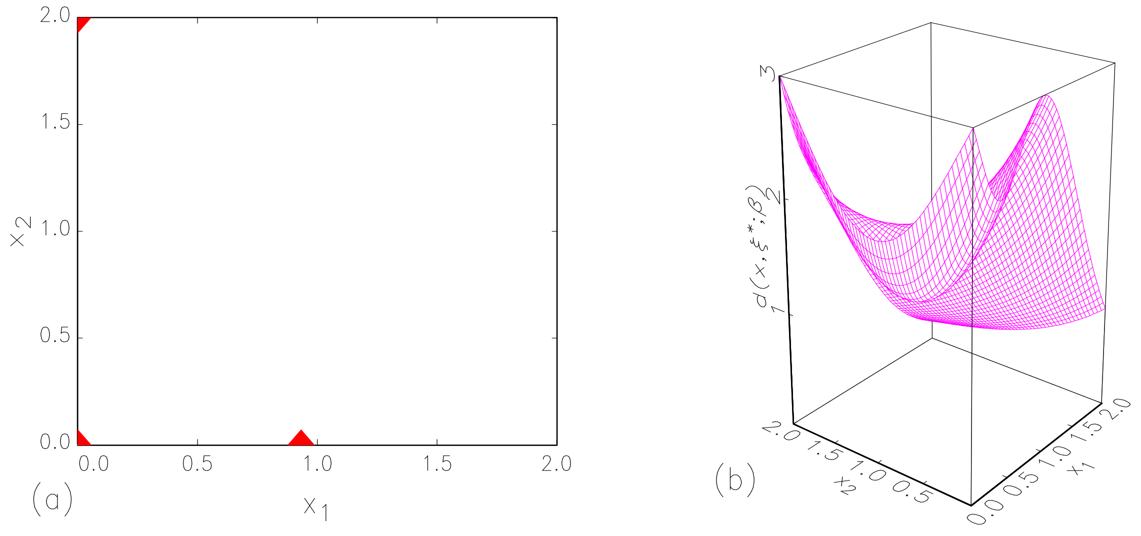


Figure 6.6: (a) Support points of the 3-point non-trapezium D-optimal design and (b) standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, 2] \times [0, 2]$ and with parameter vector $\boldsymbol{\beta} = (-1, 3, 1)^T$. Design pattern of Figure 6.5 (a).

$[0, b] \times [0, b]$ are satisfied. If the design space was $[0, \infty) \times [0, \infty)$, then following the same arguments as in the preceding case, the D-optimal design could be a 3-point design with support points at E, D and J (see Section 5.3.3). Since D and J are located outside the design space $[0, b] \times [0, b]$ they must move to H and F respectively, hence resulting in a 3-point candidate D-optimal design with support points at E, D and F. As $E(0, 0)$ is a support point, then $-1.5434 \leq \beta_0 \leq 0$ (see Theorem 5.4), and as logit line u is above the main diagonal of the design space $[0, b] \times [0, b]$ and below the vertex G at (b, b), then

$$\max\{\beta_0 + \beta_1 b, \beta_0 + \beta_2 b\} \leq 1.5434 \leq \beta_0 + (\beta_1 + \beta_2) b.$$

Hence, the design of Figure 6.5 (b) can exists provided that

$$0 < \max\{-\beta_0, \beta_0 + \beta_1 b, \beta_0 + \beta_2 b\} \leq 1.5434 \leq \beta_0 + (\beta_1 + \beta_2) b. \tag{6.31}$$

Clearly, the optimal logit lines are $u = \beta_0$, $u = \beta_0 + \beta_1 b$ and $u = \beta_0 + \beta_2 b$. The following example illustrates this design case.

Example 6.6. Consider the two-variable binary logistic model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, 2] \times [0, 2]$ where $\beta_0 = -1.4$, $\beta_1 = 1.2$ and $\beta_2 = 1$.

Condition (6.31) is satisfied since

$$0 < \max\{-\beta_0 = 1.4, \beta_0 + \beta_1 b = 1, \beta_0 + \beta_2 b = 0.6\} = 1.4 < 1.5434 < \beta_0 + (\beta_1 + \beta_2) b = 3.$$

Hence the candidate D-optimal design is given by

$$\xi^* = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{ccc} (0,0) & (0,2) & (2,0) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ -1.4 & 0.6 & 1 \end{array} \right\}. \tag{6.32}$$

To save space the plots of the support points and the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ on the design space $[0,2] \times [0,2]$ are not reported here, but it can be demonstrated numerically that $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) \leq 3$ with equality holding only at the three support points.

Case of Figure 6.5 (c) design pattern

Consider Figure 6.5 (c). The design of Figure 6.5 (c) is the symmetric reflection of the design of Figure 6.5 (a) with respect to the centre of symmetry of the design space $[0,b] \times [0,b]$. Let (y_1, y_2) and (x_1, x_2) be corresponding support points in Figure 6.5 (a) and Figure 6.5 (c) respectively. Then,

$$(x_1, x_2) = (b - y_1, b - y_2) \tag{6.33}$$

Clearly, if u is a logit line in Figure 6.5 (c), the corresponding logit line in Figure 6.5 (a) is $\widetilde{u} = -u$. Hence, the vector of parameters $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^T$ and $\boldsymbol{\beta}' = (\beta_0', \beta_1', \beta_2')^T$ in Figure 6.5 (c) and Figure 6.5 (a) respectively are linked by

$$\begin{cases} \beta_0' = -\beta_0 - (\beta_1 + \beta_2)b \\ \beta_1' = \beta_1 \\ \beta_2' = \beta_2 \end{cases} \tag{6.34}$$

In fact, assume that $u = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ and $\widetilde{u} = \beta_0' + \beta_1' x_1' + \beta_2' x_2'$. It follows from the transformation (6.33) that

$$\begin{aligned} \beta_0' + \beta_1' y_1 + \beta_2' y_2 = \widetilde{u} &= -u \\ &= -\beta_0 - \beta_1 x_1 - \beta_2 x_2 \\ &= -\beta_0 - \beta_1 (b - y_1) - \beta_2 (b - y_2) \\ &= -\beta_0 - (\beta_1 + \beta_2) b - \beta_1 y_1 - \beta_2 y_2 \end{aligned}$$

which implies the relations (6.34). The following example is a case where the design pattern of Figure 6.5 (c) is derived from the design pattern of Figure 6.5 (a).

Example 6.7. Consider the two-variable binary logistic model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0,2] \times [0,2]$ where $\beta_0 = -7$, $\beta_1 = 3$ and $\beta_2 = 1$. The parameters $\beta_0 = -7$, $\beta_1 = 3$ and $\beta_2 = 1$ of this example are linked to the parameters of Example 6.5, here written as $\beta_0' = -1$, $\beta_1' = 3$ and $\beta_2' = 1$, by the relations (6.34). Hence, the following 3-point D-optimal design with the pattern as in Figure 6.5 (c) is derived from design (6.29) using the transformations (6.33) and is given by

$$\xi^* = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{ccc} (2,2) & (2,0) & (1.068, 2) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 1 & -1 & -1.796 \end{array} \right\}.$$

Case of Figure 6.5 (d) design pattern

The design of Figure 6.5 (d) is the symmetric reflection of the design of Figure 6.5 (b) with respect to the centre of symmetry of the design space $[0, b] \times [0, b]$. Using the same argument as in the previous case, the support points (x_1, x_2) of the D-optimal design can be obtained from support points (y_1, y_2) of Figure 6.5 (b) using the transformation (6.33). Parameters are linked by the relations (6.34). The following example is a case where the design pattern of Figure 6.5 (d) is derived from the design pattern of Figure 6.5 (b).

Example 6.8. Consider the two-variable binary logistic model $u = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, 2] \times [0, 2]$ where $\beta_0 = -3$, $\beta_1 = 1.2$ and $\beta_2 = 1$. The parameters of this example are linked to the parameters $\beta_0' = -1.4$, $\beta_1' = 1.2$ and $\beta_2' = 1$ of Example 6.6 by the relations (6.34). Hence the 3-point D-optimal design with pattern as in Figure 6.5 (d) follows as

$$\xi^* = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{ccc} (2,2) & (2,0) & (0,2) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 1.4 & -0.6 & -1 \end{array} \right\}.$$

6.5. General 4-point D-optimal parallelogram designs

6.5.1 Building block of 4-point parallelogram D-optimal designs

The assumption made in the construction of the 4-point D-optimal parallelogram designs by Sitter and Torsney (1995a) and Jia and Myers (2001) in Section 5.2 was that one pair of the support points lies on the logit line $-u$ and the other pair on the logit line u with u unrestricted in $\mathbb{R}$. The optimal designs found under this assumption were equally weighted 4-point D-optimal parallelogram designs with support points lying on the optimal logit lines $\pm u^* = 1.22291$ only. This section discusses the construction of 4-point D-optimal parallelogram designs for model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the rectangular design space $[0, b] \times [0, b]$, and hence for $\beta_1 > 0$ and $\beta_2 > 0$, the logit u is restricted by $\beta_0 \le u \le \beta_0 + (\beta_1 + \beta_2)b$. It will be conjectured and verified numerically that the support points of the candidate 4-point D-optimal designs are not necessarily on logit lines of the form $\pm u$. The following results are used to establish the conditions for the existence of these designs.

Consider Figure 6.7, and in particular the quadrilateral design pattern with support points A, I, C and J. The quadrilateral $AICJ$ is a parallelogram if the slope (m_1) of the line AJ is equal to the slope (m_2) of the line IC since both lines AI and JC have slope $-\dfrac{\beta_1}{\beta_2}$. The coordinates

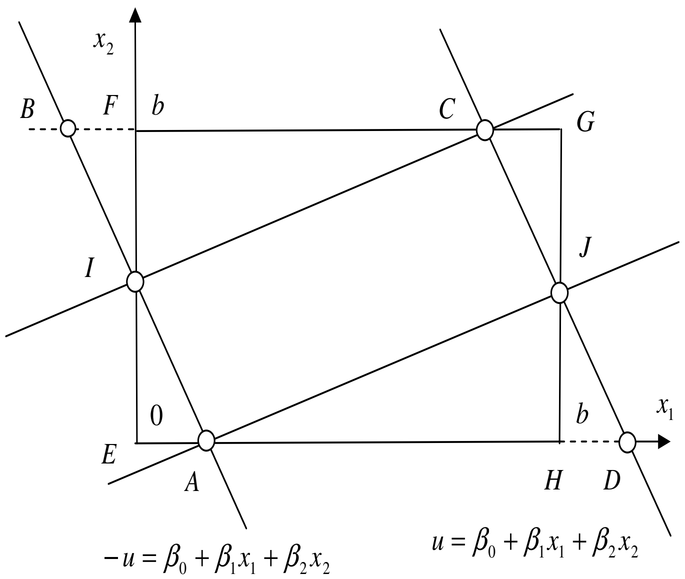


Figure 6.7: Building blocks for a parallelogram design for the two-variable binary logistic model with interaction $u = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, b] \times [0, b]$.

of the support points are: $A(\frac{-u-\beta_0}{\beta_1}, 0)$, $I(0, \frac{-u-\beta_0}{\beta_2})$, $C(\frac{u-\beta_0-\beta_2 b}{\beta_1}, b)$ and $J(b, \frac{u-\beta_0-\beta_1 b}{\beta_2})$. Hence,

$$m_1 = \frac{\frac{u-\beta_0-\beta_1 b}{\beta_2} - 0}{b + \frac{u+\beta_0}{\beta_1}} = \frac{\beta_1(u - \beta_0 - \beta_1 b)}{\beta_2(u + \beta_0 + \beta_1 b)} \quad \text{and} \quad m_2 = \frac{b + \frac{u+\beta_0}{\beta_2}}{\frac{u-\beta_0-\beta_2 b}{\beta_1} - 0} = \frac{\beta_1(u + \beta_0 + \beta_2 b)}{\beta_2(u - \beta_0 - \beta_2 b)}.$$

It thus follows that $m_1 = m_2$ implies that $\dfrac{u - \beta_0 - \beta_1 b}{u + \beta_0 + \beta_1 b} = \dfrac{u + \beta_0 + \beta_2 b}{u - \beta_0 - \beta_2 b}$. Simple algebra leads to $2\beta_0 + (\beta_1 + \beta_2)b = 0$. Thus, the condition for $AICJ$ to be a parallelogram is

$$2\beta_0 + (\beta_1 + \beta_2)b = 0 \text{ or, equivalently, } \beta_0 + \frac{1}{2}\beta_1 b + \frac{1}{2}\beta_2 b = 0. \tag{6.35}$$

Note that condition (6.35) implies that the logit line $\beta_0 + \beta_1 x_1 + \beta_2 x_2 = 0$ passes through the point $\left(\dfrac{b}{2}, \dfrac{b}{2}\right)$ which is the centre of symmetry of the square design space $[0, b] \times [0, b]$.

Equation (6.35) also implies that $\beta_0 + \beta_1 b = -(\beta_0 + \beta_2 b)$ which means that the logit values at points $F(0, b)$ and $H(b, 0)$ in Figure 6.7 have opposite signs, i.e. either $\beta_0 + \beta_1 b > 0$ and $\beta_0 + \beta_2 b < 0$ or $\beta_0 + \beta_2 b > 0$ and $\beta_0 + \beta_1 b < 0$. This fact was not the case for the trapezium designs and related designs discussed in Sections 6.3 and 6.4 where $\beta_0 + \beta_1 b$ and $\beta_0 + \beta_2 b$ had the same sign. Equation (6.35) has an infinite number of solutions for β_0, β_1 and β_2 which implies that there are an infinite number of parallelogram design patterns that satisfy condition (6.35). However, these parallelogram design patterns can be classified into a few broad categories described below.

6.5.2 4-point equally weighted designs

Consider the design patterns in Figures 6.8 (a) and 6.8 (b) where the candidate support points A, B, C and D are represented by circles. In Figure 6.8 (a) $\beta_1 > \beta_2$ and in Figure 6.8 (b) $\beta_1 < \beta_2$.

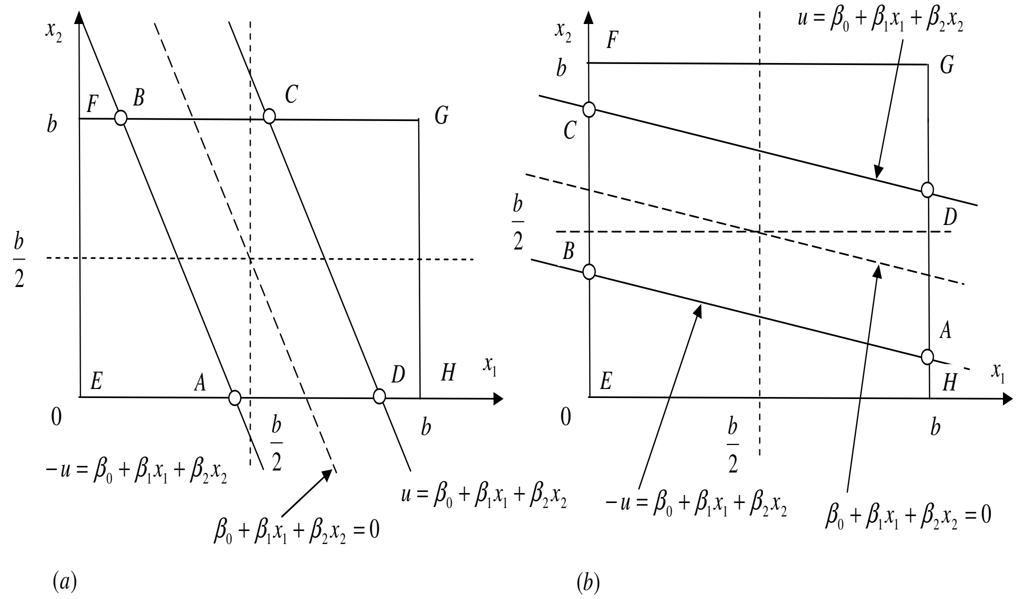


Figure 6.8: Parallelogram design patterns for the two-variable binary logistic model without interaction on the design space $[0, b] \times [0, b]$ and with parameters satisfying conditions $\beta_0 = -\frac{b}{2}(\beta_1 + \beta_2)$ and (a) $0 < \frac{2.44582}{b} \leq \beta_1 - \beta_2$, (b) $0 < \frac{2.44582}{b} \leq \beta_2 - \beta_1$.

In both figures, the 4 support points are vertices of a parallelogram with either $0 \leq x_1 \leq b$ and

$x_2 \in \{0, b\}$ or $x_1 \in \{0, b\}$ and $0 \leq x_2 \leq b$. Furthermore, the support points of the candidate 4-point D-optimal design are located on two parallel logit lines $\pm u$ which intersect only the two horizontal sides of the rectangular design space. Thus the D-optimal designs with design patterns in Figure 6.8 (a) and Figure 6.8 (b) are similar to those of Sitter and Torsney (1995a) with design locus G_w and of Jia and Myers (2001), and are a special case of the designs of Atkinson and Haines (1996). These designs were fully discussed in Section 5.2, and it was proved in Theorem 5.1 that the D-optimal design is an equally weighted 4-point design with support points on the logit lines $u^* = \pm 1.22291$. Note that as in the case of Sitter and Torsney (1995a) and of Jia and Myers (2001), discussed in Section 4.5 and Section 5.2, no candidate support point is on the logit $u = 0$. However, as it will be observed later the logit $u = 0$ serves as a line of reference in deriving the conditions for existence of various design patterns. The conditions on the parameters and the design space for which the designs of Figure 6.8 (a) and Figure 6.8 (b) are possible are given below.

Consider the design pattern in Figure 6.8 (a). The pattern is possible if the abscissa of point B is greater than or equal to that of F and the abscissa of D is less than or equal to that of H. Since the respective coordinates of F and H are $(0, b)$ on logit line $-u$ and $(b, 0)$ on logit line u, the design pattern of Figure 6.8 (a) is possible if the condition (6.35) is satisfied, $\beta_0 + \beta_2 b \leq -u < 0$ and $0 < u \leq \beta_0 + \beta_1 b$, or, equivalently if the condition (6.35) is satisfied, and $0 < u \leq -\beta_0 - \beta_2 b$ and $0 < u \leq \beta_0 + \beta_1 b$ which simplifies to

$$\beta_0 = -\frac{b}{2}(\beta_1 + \beta_2) \text{ and } 0 < u \leq \beta_0 + \beta_1 b = \frac{b}{2}(\beta_1 - \beta_2) \tag{6.36}$$

since condition (6.35) implies that $-\beta_0 - \beta_2 b = \beta_0 + \beta_1 b = \frac{b}{2}(\beta_1 - \beta_2)$. Since the optimal value of u is $u^* = 1.22291$, it follows from (6.36) that the conditions that an equally weighted 4-point parallelogram exists with $\beta_1 > \beta_2$ are

$$\beta_0 = -\frac{b}{2}(\beta_1 + \beta_2) \text{ and } 0 < \frac{2u^*}{b} \leq \beta_1 - \beta_2 \tag{6.37}$$

where $u^* = 1.22291$. Thus, if $b > 0$ and $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^T$ satisfy condition (6.37), then the candidate D-optimal design with the pattern in Figure 6.8 (a) in the (x_1, x_2)-space is

$$\xi^* = \left\{ \begin{array}{cccc} (\frac{-2u^*+b(\beta_1+\beta_2)}{2\beta_1}, 0) & (\frac{-2u^*+b(\beta_1-\beta_2)}{2\beta_1}, b) & (\frac{2u^*+b(\beta_1-\beta_2)}{2\beta_1}, b) & (\frac{2u^*+b(\beta_1+\beta_2)}{2\beta_1}, 0) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array} \right\} \tag{6.38}$$

where $u^* = 1.22291$, or, equivalently, in the (u_1, u_2) space, where $u_1 = u$ and $u_2 = x_2$, is

$$\xi_u^* = \left\{ \begin{array}{cccc} (-1.22291, 0) & (-1.22291, b) & (1.22291, b), & (1.22291, 0) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array} \right\}. \tag{6.39}$$

Note that the design ξ_u^* given in (6.39) can be deduced from the D-optimal design (5.3) by transforming each point (u, x_2) of design (5.3) to $(u, \frac{1}{2}(x_2+b))$. Then, it follows from Theorem 5.1 that the design ξ_u^* given in (6.39) is D-optimal, or, equivalently, the design ξ^* given in (6.38) is D-optimal.

In a similar way, the design pattern in Figure 6.8 (b) is possible if $2\beta_0+b(\beta_1+\beta_2) = 0$, $\beta_0+\beta_1 b \leq -u < 0$ and $0 < u \leq \beta_0 + \beta_2 b$, or, equivalently, if $\beta_0 = -\dfrac{b}{2}(\beta_1 + \beta_2)$ and $0 < \dfrac{2u}{b} \leq \beta_2 - \beta_1$. Hence, the conditions that an equally weighted 4-point parallelogram exists with $\beta_2 > \beta_1$ are

$$\beta_0 = -\frac{b}{2}(\beta_1 + \beta_2) \text{ and } 0 < \frac{2u^*}{b} \leq \beta_2 - \beta_1 \tag{6.40}$$

where $u^* = 1.22291$.

Thus, if $b > 0$ and $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^T$ satisfy the conditions (6.40), then the (x_1, x_2)-space is

$$\xi^* = \left\{ \begin{array}{cccc} (0, \frac{-2u^*+b(\beta_1+\beta_2)}{2\beta_2}) & (b, \frac{-2u^*+b(\beta_2-\beta_1)}{2\beta_2}) & (b, \frac{2u^*+b(\beta_2-\beta_1)}{2\beta_2}) & (0, \frac{2u^*+b(\beta_1+\beta_2)}{2\beta_2}) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array} \right\} \tag{6.41}$$

where $u^* = 1.22291$, or, equivalently, in the (u, x_2) space is

$$\xi_u^* = \left\{ \begin{array}{cccc} (0, -1.22291) & (b, -1.22291) & (b, 1.22291) & (0, 1.22291) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array} \right\}. \tag{6.42}$$

Note that the conditions (6.40) are obtained from (6.37) by exchanging β_1 and β_2 in model $u = \beta_0 + \beta_1 x_1 + \beta_2 x_2$. In addition, design (6.41) is obtained from design (6.38) by exchanging the first and second coordinates of each support point of design (6.38). Hence, it also follows from Theorem 5.1 that the design ξ_u^* given in (6.42) is D-optimal, or, equivalently, the design ξ^* given in (6.41) is D-optimal.

Finally, it can be deduced from conditions (6.37) and (6.40) that the conditions of existence of an equally weighted parallelogram D-optimal design on the design space $[0, b] \times [0, b]$ are

$$\beta_0 = -\frac{b}{2}(\beta_1 + \beta_2) \text{ and } 0 < \frac{2u^*}{b} \leq |\beta_1 - \beta_2| \tag{6.43}$$

where $u^* = 1.22291$ and $|\beta_1 - \beta_2|$ is the absolute value of $\beta_1 - \beta_2$. The following example illustrates a situation where the parameter values and the design space lead to equally weighted 4-point D-optimal designs given in (6.38) and in (6.41).

Example 6.9. Consider the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ where (a) $\beta_0 = -4$, $\beta_1 = 3$ and $\beta_2 = 1$ and (b) $\beta_0 = -4$,

$\beta_1 = 1$ and $\beta_2 = 3$ on the design space $[0,2] \times [0,2]$. The parameter values of part (a) satisfy condition (6.37) for design (6.38) since

$$\beta_0 = -\frac{b}{2}(\beta_1 + \beta_2) = -4 \text{ and } \frac{2u^*}{b} = 1.22291 < \beta_1 - \beta_2 = 2$$

Hence, the D-optimal design (6.38) is

$$\xi^* = \left\{ \begin{array}{cccc} (0.926, 0) & (0.259, 2) & (1.074, 2) & (1.741, 0) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array} \right\}. \tag{6.44}$$

The transformation of this design to the design space $[-1,1] \times [-1,1]$ is as follows. The transformations (6.3) and (6.4) can be written respectively as

$$\left\{ \begin{array}{l} x_{1new} = \frac{(1+1)x_1 - (0+2)}{2-0} = x_1 - 1 \\ x_{2new} = \frac{(1+1)x_2 - (0+2)}{2-0} = x_2 - 1 \end{array} \right.$$

and

$$\left\{ \begin{array}{l} \beta_{1new} = \frac{2-0}{1+1} \times 3 = 3 \\ \beta_{2new} = \frac{2-0}{1+1} \times 1 = 1 \\ \beta_{0new} = -4 + \frac{0+2}{1+1} \times 3 + \frac{0+2}{1+1} \times 1 = -4 + 3 + 1 = 0. \end{array} \right.$$

Therefore, the design (6.44) becomes

$$\xi^* = \left\{ \begin{array}{cccc} (-0.074, -1) & (-0.741, 1) & (0.074, 1) & (0.741, -1) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array} \right\}$$

which is the same as the D-optimal design found numerically by Atkinson and Haines (1996) for $\boldsymbol{\beta} = (0, 3, 1)^T$.

The parameter values of part (b) of the example satisfy condition (6.40) for design (6.41) since

$$\beta_0 = -\frac{b}{2}(\beta_1 + \beta_2) = -4 \text{ and } \frac{2u^*}{b} = 1.22291 < \beta_2 - \beta_1 = 2.$$

Hence, the D-optimal design (6.41) is

$$\xi^* = \left\{ \begin{array}{cccc} (0, 0.926) & (2, 0.259) & (2, 1.074) & (0, 1.741) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array} \right\}.$$

As expected this design can be obtained from design (6.44) by exchanging the first and the second coordinate of each support point.

6.5.3 4-point design with two equally weighted pairs of support points

Consider the designs of Figure 6.9 (a). Here, it is assumed that $\beta_1 > \beta_2$. Note that the location of the logit line $u = 0$ in Figure 6.9 (a) implies that $\beta_0 + \beta_1 b > 0$ and $\beta_0 + \beta_2 b < 0$, and hence the necessary conditions of a candidate 4-point parallelogram designs are satisfied. As the logit line $u = 0$ passes through the origin $\left(\frac{b}{2}, \frac{b}{2}\right)$ of the design space $[0, b] \times [0, b]$, then the building block condition (6.35) for a parallelogram design is satisfied. As discussed in Chapter

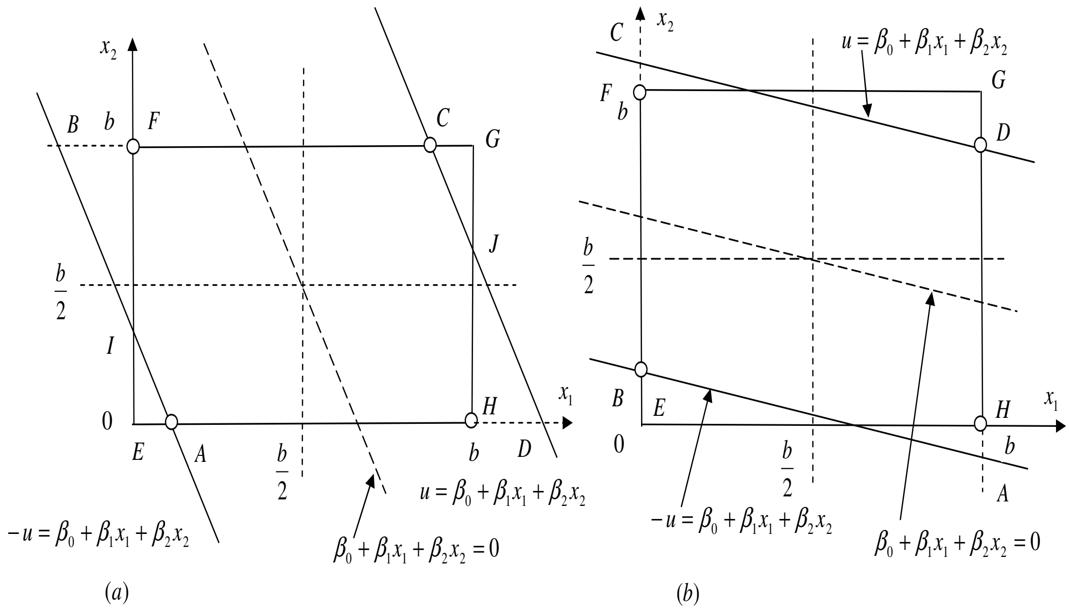


Figure 6.9: Parallelogram design patterns for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, b] \times [0, b]$ and with parameters satisfying conditions where $\beta_0 = -\frac{b}{2}(\beta_1 + \beta_2)$ and (a) $0 < \beta_1 - \beta_2 \leq \frac{2.44582}{b} < \beta_1 + \beta_2$, (b) $0 < \beta_2 - \beta_1 \leq \frac{2.44582}{b} < \beta_1 + \beta_2$.

4, Sections 4.5.1 and 4.5.4, if the restriction $0 \leq x_1 \leq b$ is not imposed in Figure 6.9 (a), then the support points of the candidate D-optimal design are A, B, C and D. However with the restriction $0 \leq x_1 \leq b$ the candidate D-optimal design has support points A, F, C and H. Note that A and F are not on the same logit line and therefore are not expected to have equal design weights. Similarly, points C and H are not expected to have equal weights. However, points A and C are on the respective logits $-u$ and u, with $u \geq 0$, and are also symmetric about the centre $\left(\frac{b}{2}, \frac{b}{2}\right)$ of the design space. Thus A and C are expected to have equal weights, say $\frac{\lambda}{2}$, in a D-optimal design. In the same figure, points F and H are symmetric with respect to the centre of symmetry of the design space $[0, b] \times [0, b]$. Hence, the two candidate support points F and H are expected to have equal weights, say $\frac{1-\lambda}{2}$, in a D-optimal design. Condition (6.35) implies that logit values at F and H are respectively $-a$ and a with $a = \beta_0 + \beta_2 b$ such that $0 < a = \beta_0 + \beta_1 b = -\beta_0 - \beta_2 b = \frac{b}{2}(\beta_1 - \beta_2) \leq u < \beta_0 + b(\beta_1 + \beta_2) = \frac{b}{2}(\beta_1 + \beta_2)$. Thus, the conditions for the design pattern of Figure 6.9 (a) to hold are

$$\beta_0 = -\frac{b}{2}(\beta_1 + \beta_2) \text{ and } 0 < \beta_1 - \beta_2 \leq \frac{2u^*}{b} < \beta_1 + \beta_2 \tag{6.45}$$

where $u^* = 1.22291$ is the optimal value of u when if the restriction was imposed on u. The candidate D-optimal design, in terms of (u, x_2) space, is of the form

$$\xi_u = \left\{ \begin{array}{cccc} (-u, 0) & (-a, b) & (u, b) & (a, 0) \\ \frac{\lambda}{2} & \frac{1-\lambda}{2} & \frac{\lambda}{2} & \frac{1-\lambda}{2} \end{array} \right\} \tag{6.46}$$

where $u \geq a = \beta_0 + \beta_1 b > 0$, and the optimal values of u and λ are not necessarily equal to 1.22291 and 0.5 respectively. Clearly $\pm a$ with $a \in (0, 1.22291)$ are the logit values at the corner points $(0, b)$ and $(b, 0)$. The purpose of using (u, x_2) space here, and not the (x_1, x_2)-space is to obtain relatively readable expressions of the information matrix and its determinant.

Now, consider deriving the optimal values of u and λ in design (6.46) algebraically. The information matrix for the model parameters is

$$M(\xi_u; \boldsymbol{\beta}) = \frac{(1-\lambda)e^a}{2(1+e^a)^2} \begin{bmatrix} 2 & 0 & b \\ 0 & 2a^2 & -ab \\ b & -ab & b^2 \end{bmatrix} + \frac{\lambda e^u}{2(1+e^u)^2} \begin{bmatrix} 2 & 0 & b \\ 0 & 2u^2 & -ub \\ b & -ub & b^2 \end{bmatrix} \tag{6.47}$$

and its determinant is

$$D = \frac{b^2\lambda(1-\lambda)e^{a+u}(a+u)^2 \left[2e^{a+u} + (1-\lambda)e^a(1+e^{2u}) + \lambda e^u(1+e^{2a})\right]}{4(1+e^a)^4(1+e^u)^4}. \tag{6.48}$$

The respective derivatives of D with respect to λ and u are

$$\frac{\partial D}{\partial \lambda} = \frac{b^2(a+u)^2 e^{a+u} \left[\lambda(2-3\lambda)e^u(1+e^{2a}) + 2(1-2\lambda)e^{a+u} + (1-4\lambda+3\lambda^2)e^a(1+e^{2u})\right]}{4(1+e^a)^4(1+e^u)^4}$$

and

$$\frac{\partial D}{\partial u} = \frac{b^2\lambda(1-\lambda)(a+u)e^{a+u}}{4(1+e^a)^4(1+e^u)^5} \Big\{ (a+u)\left[(1-\lambda)e^a(1-e^{3u}) + e^u(1-e^u)(2\lambda(1+e^{2a}) + (1+3\lambda)e^a)\right] + \\ 2(1-\lambda)e^a(1+e^{3u}) + 2e^u(1+e^u)\left[(3-\lambda)e^a + \lambda(1+e^{2a})\right] \Big\}.$$

Hence, the optimal values of λ and u are the solutions for (λ, u) to the simultaneous equations

$$\lambda(2-3\lambda)e^{u}(1+e^{2a})+2(1-2\lambda)e^{a+u}+(1-4\lambda+3\lambda^2)e^{a}(1+e^{2u})=0 \tag{6.49}$$

and

$$\begin{aligned}(a+u)&\left[(1-\lambda)e^{a}(1-e^{3u})+e^{u}(1-e^{u})(2\lambda(1+e^{2a})+(1+3\lambda)e^{a})\right]\\ &+2(1-\lambda)e^{a}(1+e^{3u})+2e^{u}(1+e^{u})\left[(3-\lambda)e^{a}+\lambda(1+e^{2a})\right]=0.\end{aligned} \tag{6.50}$$

An analytical closed form solution for u is not possible since neither of the equations is a linear function of u. However, for a given u, an analytical closed form solution for λ can be found from equation (6.49) which is a quadratic polynomial in λ. The two possible solutions of (6.49) for λ are

$$\lambda^{*}=\frac{S_1}{S_2+\sqrt{S_3}} \tag{6.51}$$

and

$$\lambda^{**}=\frac{S_1}{S_2-\sqrt{S_3}} \tag{6.52}$$

where $S_1=e^{a}(1+e^{u})^2$, $S_2=e^{a}(1+e^{u})^2+(e^{u}-e^{a})(e^{a+u}-1)$ and $S_3=[e^{a}(1+e^{u})^2-(e^{u}-e^{a})(e^{a+u}-1)]^2+e^{a}(1+e^{u})^2(e^{u}-e^{a})(e^{a+u}-1)$. A feasible solution must be positive and less than 1.

Consider λ^* as defined in (6.51). Since, $u \geq a > 0$, then $\lambda^* > 0$ because $S_1 > 0$, $S_2 > 0$ and $S_3 > 0$. Furthermore $\lambda^* < 1$ because $S_2+\sqrt{S_3}-S_1=(e^{u}-e^{a})(e^{a+u}-1)+S_3>0$. Hence, λ^* is a feasible value of λ since $0<\lambda^*<1$.

Now, consider λ^{**} as defined in (6.52). Clearly, $\lambda^{**} > 1$. This is achieved by showing that $S_1 > S_2-\sqrt{S_3}$. In fact, $S_1-(S_2-\sqrt{S_3})=-(e^{u}-e^{a})(e^{a+u}-1)+\sqrt{S_3}$. Note that $(e^{u}-e^{a})(e^{a+u}-1)>0$, $S_3>0$ and $S_3-[e^{u}-e^{a})(e^{a+u}-1)]^2=e^{a+u}(1+e^{a})^2(1+e^{u})^2>0$ which implies that $\sqrt{S_3}>(e^{u}-e^{a})(e^{a+u}-1)$. Hence $S_1 > S_2-\sqrt{S_3}$, and therefore λ^{**} is not a feasible value of λ.

Solving the equation (6.50) numerically, evaluated at a given value $a=\beta_0+\beta_1 b$ of the logit at the corner point $(b,0)$ and at weight $\lambda=\lambda^*$, for u gives u^* as the candidate optimal logit. For example, Table 6.1 gives values of u^* and λ^* for selected values of a. The values $a=0$ and $a=1.22291$ are included only to indicate boundary optimal values of λ and u. Figure 6.10 displays the scatter plot of u^* (circles) and λ^* (squares) versus a in the interval $[0, 1.22291]$. Both Table 6.1 and Figure 6.10 indicate that λ^* is an increasing function and u^* is a decreasing function of a. A scatter plot of u^* versus a is displayed by "circles" in Figure 6.10. Clearly

Table 6.1: Values of u^* and λ^* for selected values of $a = \beta_0 + \beta_1 b$: 4-point D-optimal parallelogram design for model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ with two equally weighted pairs of support points.

a	u^*	λ^*
0	2.1150	0.4017
0.1	2.0340	0.4064
0.125	2.0139	0.4077
0.25	1.9144	0.4146
0.375	1.8165	0.4223
0.5	1.7205	0.4314
0.625	1.6267	0.4410
0.75	1.5355	0.4518
0.875	1.4474	0.4636
1	1.3629	0.4762
1.15	1.2671	0.4971
1.22291	1.22291	0.5

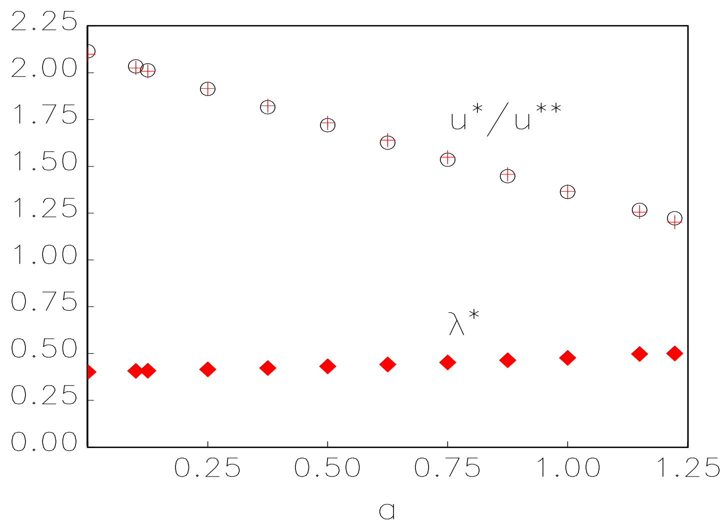


Figure 6.10: The scatter plot of λ^* (solid squares) and u^* (open circles) and u^{**} (plus) versus a. Values of λ^* and u^* are respective solutions for λ and u in the simultaneous equations (6.49 or 6.51) and (6.50) respectively, and u^{**} is given by (6.53).

a straight line with a negative slope can fit the data. The regression line for u^* as a simple linear function of a on $[0, 1.22291]$ is

$$u^{**} = \widehat{u} = 2.0994 - 0.733a \tag{6.53}$$

with $R^2 = 0.999$. The scatter plot of u^{**} versus a is represented by "plus" symbols in Figure 6.10. Clearly, the points (a, u^{**}) are very close to the points (a, u^*), and hence are good fit. If u^{**} is used instead of u^*, it can easily be checked that each D-efficiency, given by expression (3.38) with $p = 3$, of the design (6.46) with $u = u^{**}$ relative to the design (6.46) with $u = u^*$ is practically equal to 1.

For a candidate D-optimal design ξ_u^*, D-optimality can be checked by plotting the standardized variance function $d(u, \xi_u^*) = \mathrm{tr}[M^{-1}(\xi_u^*)M(\boldsymbol{u})]$ versus $u \geq \beta_0$. In this case

$$d(u, \xi_u^*) = \frac{e^{u-u^*-a}(1+e^a)^4(1+e^{u^*})^4(s_1+s_2)}{b^2\lambda^*(1-\lambda^*)[2e^{a+u^*} + (1-\lambda^*)e^a(1+e^{2u^*}) + \lambda^* e^{u^*}(1+e^{2a})](a+u^*)^2(1+e^u)^2} \tag{6.54}$$

where

$$s_1 = b^2\left\{-\left[\frac{(\lambda^*-1)ae^a}{(1+e^a)^2} + \frac{\lambda^* u^* e^{u^*}}{(1+e^{u^*})^2}\right]^2 + \left[\frac{(1-\lambda^*)e^a}{(1+e^a)^2} + \frac{\lambda^* e^{u^*}}{(1+e^{u^*})^2}\right]^2 u^2\right\}$$

and

$$s_2 = 2\left\{\frac{(1-\lambda^*)e^a}{(1+e^a)^2} + \frac{\lambda^* e^{u^*}}{(1+e^{u^*})^2}\right\}\left\{\left[\frac{(1-\lambda^*)ae^a(u-a)}{(1+e^a)^2} - \frac{\lambda^* u^* e^{u^*}(u+u^*)}{(1+e^{u^*})^2}\right]b(2x_2-b) + \right.$$
$$\left. 2x_2^2\left[\frac{(1-\lambda^*)a^2e^a}{(1+e^a)^2} + \frac{\lambda^* u^{*2} e^{u^*}}{(1+e^{u^*})^2}\right]\right\}.$$

Setting $u = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ in (6.54) transforms $d(u, \xi_u^*)$ with $u \geq \beta_0$ to $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ with $\boldsymbol{x} = (x_1, x_2)^T \in [0, b] \times [0, b]$. The following example illustrates the case of Figure 6.9 (a).

Example 6.10. Consider the two-variable binary logistic model without interaction $u = \mathrm{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, 2] \times [0, 2]$ where $\beta_0 = -2$, $\beta_1 = \frac{3}{2}$ and $\beta_2 = \frac{1}{2}$. Here, $b = 2$ and $2\beta_0 + (\beta_1 + \beta_2)b = 0$. In addition, $\beta_1 - \beta_2 = 1 < 1.22291 < -\beta_0 = 2$. Therefore condition (6.45) is satisfied. Thus, a 4-point design of the form (6.46) is expected. Moreover, $a = \beta_0 + \beta_1 b = 1$. Then, from Table 6.1, the optimal values of u and λ are $u^* = 1.3629$ and $\lambda^* = 0.4762$. Hence, $\dfrac{\lambda^*}{2} = 0.2381$ and $\dfrac{1-\lambda^*}{2} = 0.2619$, and the design (6.46) is given by

$$\xi_u^* = \left\{\begin{array}{c} (u, x_2) \\ \lambda \end{array}\right\} = \left\{\begin{array}{cccc} (-1.363, 0) & (-1, 2) & (1.363, 2) & (1, 0) \\ 0.2381 & 0.2619 & 0.2381 & 0.2619 \end{array}\right\}$$

or in the (x_1, x_2)-space by

$$\xi^* = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{cccc} (0.425, 0) & (0, 2) & (1.575, 2) & (2, 0) \\ 0.2381 & 0.2619 & 0.2381 & 0.2619 \\ -1.363 & -1 & 1.363 & 1 \end{array} \right\}. \tag{6.55}$$

The standardized variance function (6.54) associated with design (6.55) is

$$d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) = \frac{(27.117 - 24.187x_1 - 18.951x_2 + 9.129x_1^2 + 6.511x_2^2 + 5.930x_1x_2)e^u}{(1 + e^u)^2} \tag{6.56}$$

where $u = -2 + 1.5x_1 + 0.5x_2$. The support points of the candidate D-optimal design (6.55) are represented by symbols in Figure 6.11 (a) and are vertices of a parallelogram. Furthermore, the function (6.56) represented in Figure 6.11 (b) confirms numerically that $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) = 3$ at the four support points of the design (6.55) and $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) < 3$ elsewhere in $[0, 2] \times [0, 2]$ and thus the figure suggests global optimality of the design (6.55).

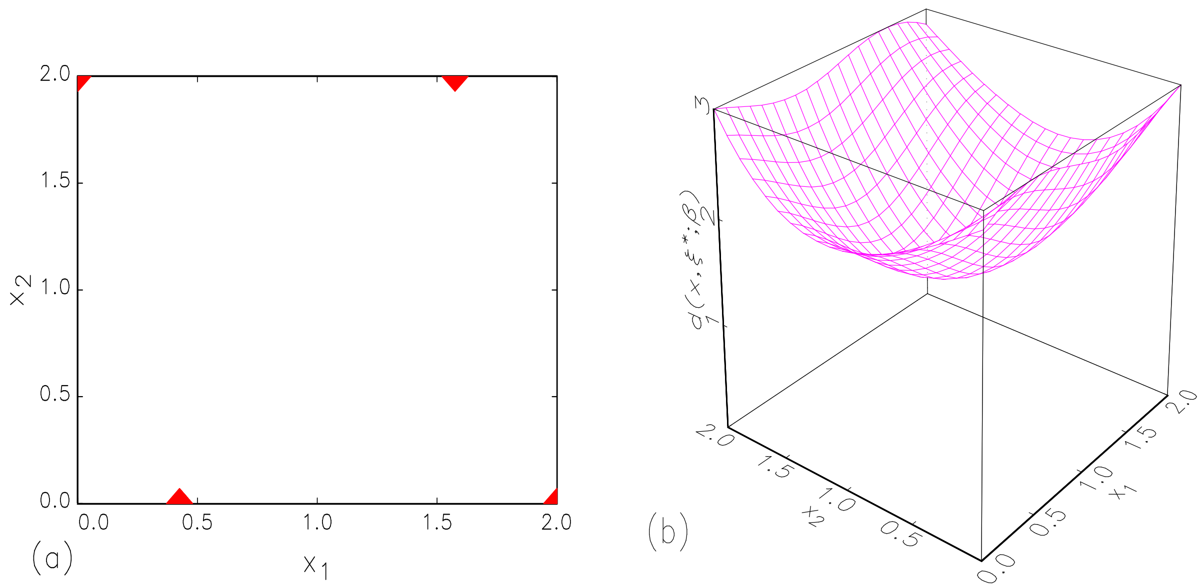


Figure 6.11: (a) Support points and (b) the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1x_1 + \beta_2x_2$ on the design space $[0, 2] \times [0, 2]$ and with parameter vector $\boldsymbol{\beta} = (-2, \frac{3}{2}, \frac{1}{2})^T$.

If the design space $[-1, 1] \times [-1, 1]$ was used instead of the design space $[0, 2] \times [0, 2]$, the parameter vector corresponding to $\left(-2, \frac{3}{2}, \frac{1}{2}\right)^T$ is $\left(0, \frac{3}{2}, \frac{1}{2}\right)^T$ and the new variables are

$x_{1\text{new}} = x_1 - 1$ and $x_{2\text{new}} = x_2 - 1$ so that the design (6.55) becomes

$$\xi^* = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{cccc} (-0.575, -1) & (-1, 1) & (0.575, 1) & (1, -1) \\ 0.2381 & 0.2619 & 0.2381 & 0.2619 \\ -1.363 & -1 & 1.363 & 1 \end{array} \right\}. \tag{6.57}$$

The design (6.57) corresponds to that constructed numerically by Atkinson and Haines (1996) for the parameter vector $\boldsymbol{\beta} = (0, \frac{3}{2}, \frac{1}{2})^T$ using the design space $[-1, 1] \times [-1, 1]$.

Now, consider the designs of Figure 6.9 (b) where it is assumed that $\beta_1 < \beta_2$ and $2\beta_0 + b(\beta_1 + \beta_2) = 0$. Reasoning in a similar way as for the design pattern of Figure 6.9 (a), the D-optimal design corresponding to Figure 6.9 (b) has support points at H, B, F and D, and the design pattern exists if the following conditions are satisfied

$$\beta_0 = -\frac{b}{2}(\beta_1 + \beta_2) \text{ and } 0 < \beta_2 - \beta_1 \leq \frac{2u^*}{b} < \beta_1 + \beta_2. \tag{6.58}$$

In this case the candidate D-optimal design, with support points expressed in term of (u, x_1), is

$$\xi_u = \left\{ \begin{array}{cccc} (-a, b) & (-u, 0) & (a, 0) & (u, b) \\ \frac{1-\lambda}{2} & \frac{\lambda}{2} & \frac{1-\lambda}{2} & \frac{\lambda}{2} \end{array} \right\} \tag{6.59}$$

where $0 < -a = -\beta_0 - \beta_1 b = \frac{b}{2}(\beta_2 - \beta_1) \leq 1.22291 < \frac{b}{2}(\beta_1 + \beta_2)$ and the optimal values of u and λ have to be calculated.

To summarize, it can be deduced from the conditions (6.45) and (6.58) that the conditions of existence of a 4-point parallelogram D-optimal design with two equally weighted pairs of support points on the design space $[0, b] \times [0, b]$ are

$$\beta_0 = -\frac{b}{2}(\beta_1 + \beta_2) \text{ and } 0 < |\beta_1 - \beta_2| \leq \frac{2u^*}{b} < \beta_1 + \beta_2 \tag{6.60}$$

where $u^* = 1.22291$ and $|\beta_1 - \beta_2|$ is the absolute value of $\beta_1 - \beta_2$.

6.5.4 4-point design with support points at the vertices of the design space

Consider the design patterns in Figures 6.12 (a) and Figure 6.12 (b) with support points E, F, G and H.

In each of the two figures, the logit line $u = 0$ passes through the centre of the design space $[0, b] \times [0, b]$, but the logit lines $\pm u$ pass outside the design space. Then, the conditions for

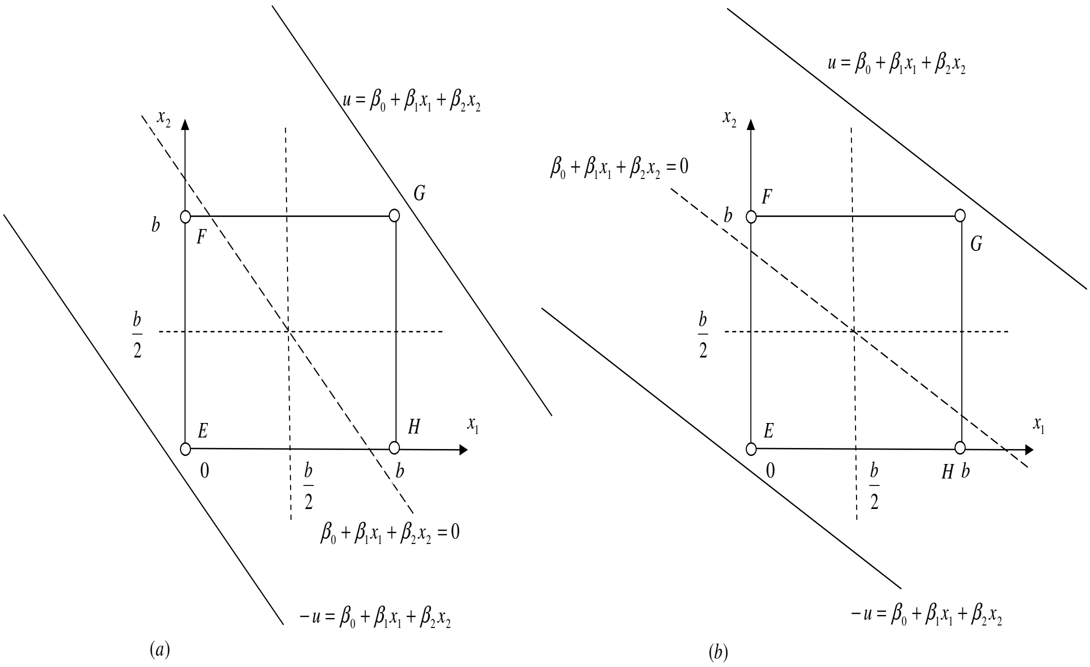


Figure 6.12: Parallelogram design for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ with $2\beta_0 + (\beta_1 + \beta_2)b = 0$ and support points at the vertices of the design space $[0, b] \times [0, b]$.

the design patterns of Figure 6.12 (a) to hold are $\beta_0 = -\frac{b}{2}(\beta_1 + \beta_2)$, $-u^* \le \beta_0 < \beta_0 + \beta_2 < 0$, and $0 < \beta_0 + \beta_1 b < \beta_0 + b(\beta_1 + \beta_2) \le u^*$ which can summarized as

$$\beta_0 = -\frac{b}{2}(\beta_1 + \beta_2) \text{ and } 0 < \beta_1 - \beta_2 < \beta_1 + \beta_2 \le \frac{2u^*}{b} \tag{6.61}$$

where $u^* = 1.22291$. In a similar way, the conditions for the design patterns of Figure 6.12 (b) to hold are

$$\beta_0 = -\frac{b}{2}(\beta_1 + \beta_2) \text{ and } 0 < \beta_2 - \beta_1 < \beta_1 + \beta_2 \le \frac{2u^*}{b} \tag{6.62}$$

where $u^* = 1.22291$. It follows from the conditions (6.61) and (6.62) that the conditions of existence of a 4-point parallelogram D-optimal design with support points at the vertices of the design space $[0, b] \times [0, b]$ for model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ with $\beta_1 \neq \beta_2$ and $2\beta_0 + b(\beta_1 + \beta_2) = 0$ are

$$\beta_0 = -\frac{b}{2}(\beta_1 + \beta_2) \text{ and } 0 < |\beta_1 - \beta_2| < \beta_1 + \beta_2 \le \frac{2u^*}{b} \tag{6.63}$$

where $u^* = 1.22291$ and $|\beta_1 - \beta_2|$ is the absolute value of $\beta_1 - \beta_2$. Since the feasible optimal logit lines must intersect the design space, the candidate D-optimal designs corresponding to

Figure 6.12 (a) and Figure 6.12 (b) are the same as designs (6.46) and (6.59) with u replaced by $-\beta_0$. For example, the design corresponding to the design pattern in Figure 6.12 (a) in terms of (x_1, x_2) coordinates is of the form

$$\xi = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{cccc} (0,0) & (0,b) & (b,b) & (b,0) \\ \frac{\lambda}{2} & \frac{1-\lambda}{2} & \frac{\lambda}{2} & \frac{1-\lambda}{2} \\ \beta_0 & -a & -\beta_0 & a \end{array} \right\} \tag{6.64}$$

where $a = \dfrac{b}{2}(\beta_1 - \beta_2) > 0$ and only the optimal value of λ needs to be calculated. Similarly, the design corresponding to the design pattern in Figure 6.12 (b) in terms of (x_1, x_2) coordinates is of the form

$$\xi = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{cccc} (0,0) & (0,b) & (b,b) & (b,0) \\ \frac{\lambda}{2} & \frac{1-\lambda}{2} & \frac{\lambda}{2} & \frac{1-\lambda}{2} \\ \beta_0 & a & -\beta_0 & -a \end{array} \right\} \tag{6.65}$$

where $a = \dfrac{b}{2}(\beta_2 - \beta_1) > 0$ and only the optimal value of λ needs to be calculated. The optimal values of λ and u are still those satisfying the simultaneous equations (6.49) and (6.50) with $u = u^* = -\beta_0$. Hence, the optimal weight is $\lambda = \lambda^*$ given by (6.51) with $u = -\beta_0$. The standardized variance function is also (6.54) with $u = -\beta_0$. The following example is an illustration of parameter values and the design space for a 4-point D-optimal design with support points at the vertices of the design space.

Example 6.11. Consider the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0,2] \times [0,2]$ where $\beta_0 = -1$, $\beta_1 = 0.6$ and $\beta_2 = 0.4$. Clearly the parameters satisfy conditions (6.61) since

$$\beta_0 = -\frac{b}{2}(\beta_1 + \beta_2) = -1 \text{ and } 0 < \beta_1 - \beta_2 = 0.2 < \beta_1 + \beta_2 = 1 < \frac{2u^*}{b} = 1.22291.$$

Hence, a 4-point D-optimal design with support points at the vertices of the design space $[0,2] \times [0,2]$ is conjectured. The weight (6.51) is $\lambda^* = 0.4716$ obtained for $u = -\beta_0 = 1$ and $a = \beta_0 + 2\beta_1 = 0.2$. The resultant D-optimal design of the form (6.64) is therefore given by

$$\xi^* = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{cccc} (0,0) & (0,2) & (2,2) & (2,0) \\ 0.2358 & 0.2642 & 0.2358 & 0.2642 \\ -1 & -0.2 & 1 & 0.2 \end{array} \right\}.$$

6.6. D-optimal designs for the case when $\beta_1 = \beta_2$ and $\beta_0 + b\beta_1 = 0$

The parallelogram D-optimal designs in Section 6.5 were constructed on the design space $[0, b] \times [0, b]$ under the assumption that $\beta_1 \neq \beta_2$ and $2\beta_0 + b(\beta_1 + \beta_2) = 0$ (condition (6.35)).

Now, consider the case when $\beta_1 = \beta_2$ and $2\beta_0 + b(\beta_1 + \beta_2) = 0$ which imply that $\beta_1 = \beta_2$ and $\beta_0 + b\beta_1 = 0$. These conditions imply that $\beta_0 + b\beta_1 = \beta_0 + b\beta_2 = 0$ which is equivalent to the logit line $u = 0$ passing through the vertices $(b, 0)$ and $(0, b)$ of the design space $[0, b] \times [0, b]$. Figures 6.13 (a) and (b) display the patterns of the candidate D-optimal designs in this case. Figure 6.13 (a) displays the pattern of a candidate 4-point D-optimal design with support points E, F, G and H, while Figure 6.13 (b) displays the pattern of a 6-point D-optimal design with support points A, I, F, C, J and H. It will be conjectured later that the latter design is a mixture of the 4-point D-optimal designs with support points A, F, J and H, and the 4-point D-optimal design with support points I, H, C and F.

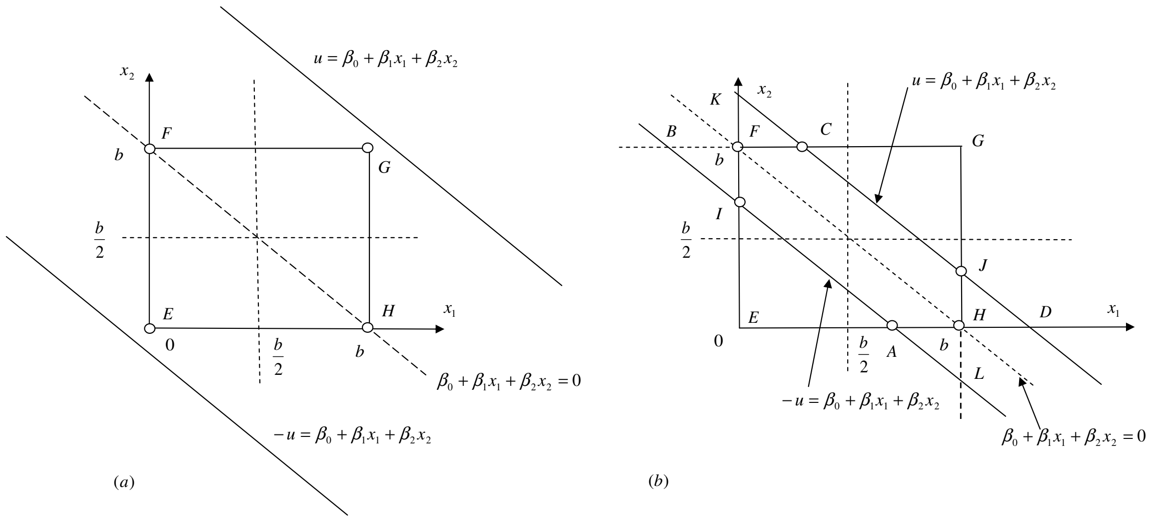


Figure 6.13: 6- and 4-point designs for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, b] \times [0, b]$ with $2\beta_0 + (\beta_1 + \beta_2)b = 0$.

Case of Figure 6.13 (a) design patterns

Consider the design pattern in Figure 6.13 (a) with support points at E, F, G and H. This design pattern is the same as the designs in Figures 6.12 (a) and (b) when the logit $u = 0$ passes through the vertices $(b, 0)$ and $(0, b)$ of the design space which implies that $a = 0$. It was found in Section 6.5 that for $a \in [0, 1.22291]$, the value at the boundary $a = 0$ is $u^* = 2.1150$, and that u^* is a decreasing function of a (see Table 6.1 and equation (6.53) or Figure 6.10). Hence, if $\beta_1 = \beta_2$ and $\beta_0 = -\dfrac{b}{2}(\beta_1 + \beta_2)$, a candidate 4-point D-optimal design with support points at the vertices of the design space exists if the logit lines $u^* = \pm 2.1150$ passe at $(0, 0)$ and (b, b) or outside the design space $[0, b] \times [0, b]$, i.e. if $\beta_0 + b(\beta_1 + \beta_2) = b\beta_1 \leq 2.1150$ and

$-2.1150 \leq \beta_0 = -b\beta_1$, or, equivalently, if

$$\beta_0 = -b\beta_1 \text{ and } 0 < \beta_1 = \beta_2 \leq \frac{2.1150}{b}. \tag{6.66}$$

The optimal logits are $u = \beta_0$ passing through point E, $u = \beta_0 + b\beta_2$ passing through point F, $u = \beta_0 + b(\beta_1 + \beta_2) = -\beta_0$ passing through point G, and $u = \beta_0 + b\beta_1$ passing through point H in Figure 6.13 (a). The design weights are still calculated using λ^* given in (6.51) with $a = 0$ and $u = -\beta_0$ to give

$$\lambda^* = \frac{(1 + e^{\beta_0})^2}{2(1 + e^{2\beta_0}) + \sqrt{1 + 14e^{2\beta_0} + e^{4\beta_0}}}. \tag{6.67}$$

Hence, the D-optimal design in this case has the same pattern as design (6.64) where $a = 0$, i.e.

$$\xi = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{cccc} (0,0) & (0,b) & (b,b) & (b,0) \\ \frac{\lambda}{2} & \frac{1-\lambda}{2} & \frac{\lambda}{2} & \frac{1-\lambda}{2} \\ \beta_0 & 0 & -\beta_0 & 0 \end{array} \right\}. \tag{6.68}$$

The following example illustrates the discussion.

Example 6.12. Consider the two-variable logistic model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0,2] \times [0,2]$ where the parameter vector is $\boldsymbol{\beta} = (-2, 1, 1)^T$. In this example, $a = \beta_0 + 2\beta_1 = 0$, $\beta_0 + b\beta_1 = 0$, $\beta_1 = \beta_2 = 1 < \frac{2.1150}{b} = 1.0575$. Hence, a candidate D-optimal design is the design in (6.68) with $\beta_0 = -2$. Further, the optimal logits are $a = 0$, $\beta_0 = -2$ and $-\beta_0 = 2$ and $\pm u = -\beta_0 = 2$. The optimal weight is (6.67) where $\beta_0 = -2$ to give $\lambda^* = 0.4082$. Hence, $\frac{\lambda^*}{2} = 0.2041$ and $\frac{1 - \lambda^*}{2} = 0.2959$. Therefore, the candidate D-optimal design of the form (6.68) is

$$\xi^* = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{cccc} (0,0) & (0,2) & (2,2) & (2,0) \\ 0.2041 & 0.2959 & 0.2041 & 0.2959 \\ -2 & 0 & 2 & 0 \end{array} \right\}. \tag{6.69}$$

The standardized variance function (6.54) is

$$d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) = \frac{(28.926 - 23.711x_1 - 23.711x_2 + 7.599x_1^2 + 7.599x_2^2 + 8.513x_1x_2)e^u}{(1 + e^u)^2} \tag{6.70}$$

where $u = -2 + x_1 + x_2$ and $(x_1, x_2) \in [0,2] \times [0,2]$.

The support points of the design (6.69) are represented by symbols in Figure 6.14 (a). The standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ illustrated in Figure 6.14 (b) suggests that the design (6.69) is D-optimal.

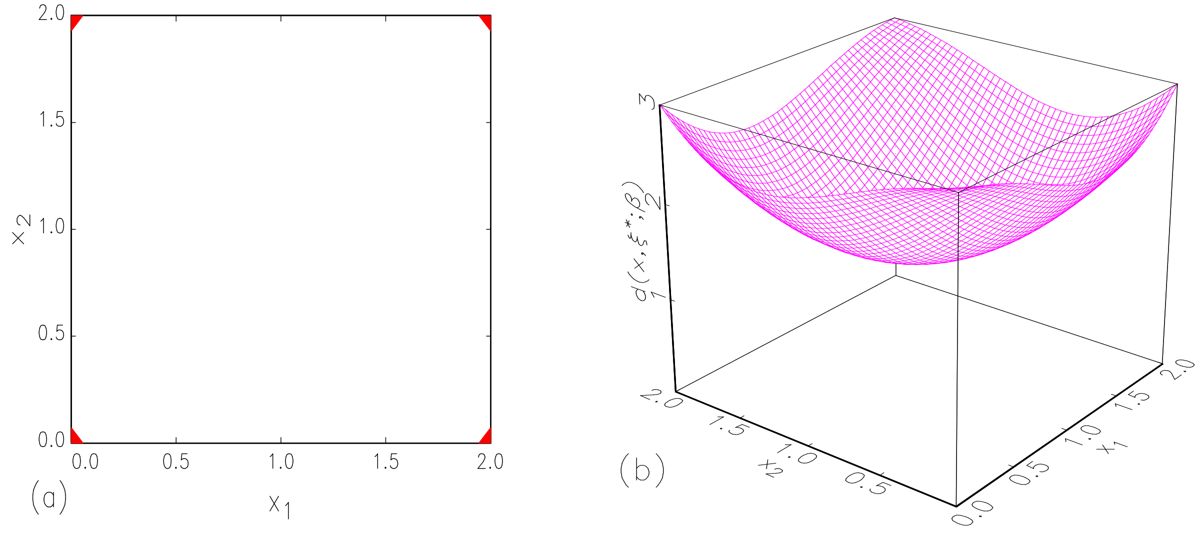


Figure 6.14: (a) Support points of the D-optimal design and (b) the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, 2] \times [0, 2]$ and with parameter vector $\boldsymbol{\beta} = (-2, 1, 1)^T$.

Adding -1 to each coordinate in design (6.69) leads to the same result as that found numerically by Atkinson and Haines (1996), and Atkinson et al. (2007, p. 403) using the design space $[-1, 1] \times [-1, 1]$ and parameter vector $\boldsymbol{\beta} = (0, 1, 1)^T$.

Comment

A slight departure from the conditions $\beta_0 + b\beta_1 = 0$ and/or $0 < \beta_1 = \beta_2 < \dfrac{2.1150}{b}$ on the design space $[0, b] \times [0, b]$ can still lead to a 4-point D-optimal design with support at the vertices of the design space. However, the weights at the support points of the candidate D-optimal designs are not expected to be equal and they cannot be calculated using result (6.67). The following example is a numerical illustration.

Example 6.13. Consider the two-variable logistic model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, 2] \times [0, 2]$ where the parameter vector is $\boldsymbol{\beta} = (-2.114, 1.05, 1.04)^T$. In this case $b = 2$ and condition $\beta_0 + b\beta_1 = 0$ is not satisfied since $\beta_0 + b\beta_1 = -0.014$. Here, the logit line $u = 0$ passes just below the centre $(0, 0)$ of the design space $[0, 2] \times [0, 2]$. Furthermore, the logits at $(0, 2)$ and $(2, 0)$ depart slightly from $a = 0$ since $\beta_0 + b\beta_1 = -0.014$ and $\beta_0 + b\beta_2 = -0.034$. As the logits at $(0, 2)$ and $(2, 0)$ are not of the form $\pm a$, then result (6.51) cannot be used in the calculation of weights at support points of the candidate D-optimal design. The candidate D-optimal design constructed numerically for this example is

$$\xi_u^* = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{cccc} (0,0) & (0,2) & (2,2) & (2,0) \\ 0.1944 & 0.2985 & 0.2086 & 0.2985 \\ -2.114 & -0.034 & 2.066 & -0.014 \end{array} \right\}. \tag{6.71}$$

Note that no pairs of logits are exactly equal. Also, the weights at points $(0,0)$ and $(2,2)$ are not equal. The support points of the design (6.71) are represented by triangle symbols in Figure 6.15 (a). The standardized variance function $d(\boldsymbol{x}, \xi^*; \beta)$ illustrated in Figure 6.15 (b) suggests that the design (6.71) is D-optimal.

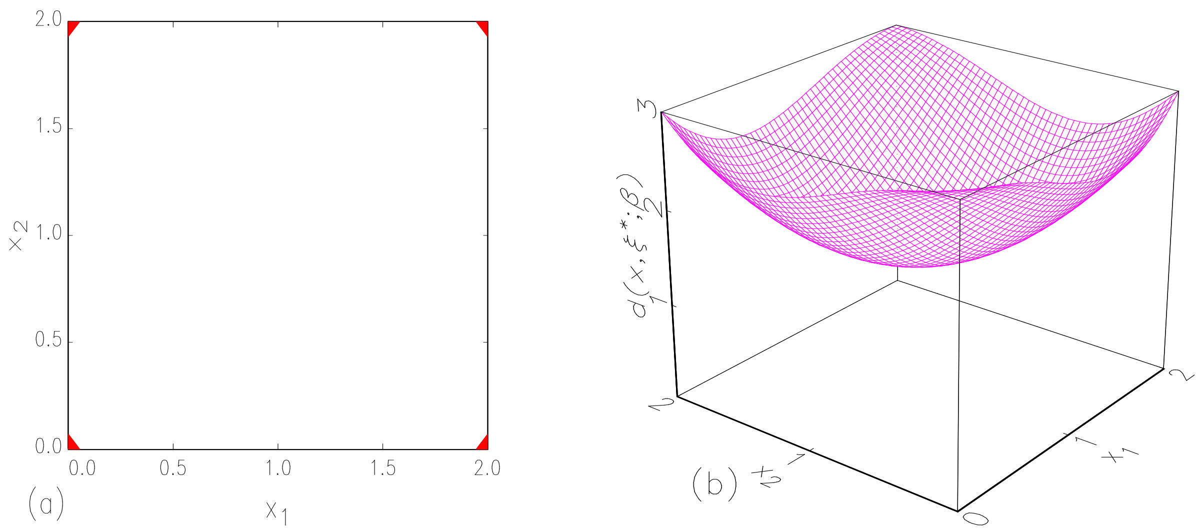


Figure 6.15: (a) Support points of the D-optimal design and (b) the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0,2] \times [0,2]$ and with parameter vector $\boldsymbol{\beta} = (-2.114, 1.05, 1.04)^T$.

Case of Figure 6.13 (b) design patterns

Consider Figure 6.13 (b). The design pattern of Figure 6.13 (b) is generated by the design pattern of 6.13 (a) by moving the logit lines $\pm u$ from the vertices E and G to the interior of the design space. Recall from the above subsection that the conditions for the existence of the design pattern of Figure 6.13 (a) is that $\beta_0 = -b\beta_1$ and $0 < \beta_1 = \beta_2 \leq \dfrac{2.1150}{b}$. Thus, it is clear that the condition for the existence of the design patterns of Figure 6.13 (b) is that $\beta_0 = -b\beta_1$ and $\beta_1 = \beta_2 > \dfrac{2.1150}{b}$. The candidate D-optimal designs can be found as follows.

Without the restriction $0 \leq x_1 \leq b$, and with the assumption that $\beta_1 > \beta_2$, the candidate 4-point D-optimal design would have support points A, B, C and D in Figure 6.13 (b). As discussed in Section 6.5 (see design pattern of Figure 6.9 (a)), the restriction $0 \leq x_1 \leq b$ would force the candidate 4-point D-optimal design to have support points A, F, C and H in 6.13 (b). Similarly, without the restriction $0 \leq x_2 \leq b$ and with the assumption that $\beta_1 < \beta_2$, the candidate 4-point D-optimal design would have support points I, K, J and L. Then, the restriction $0 \leq x_2 \leq b$ would force the candidate 4-point D-optimal design to have support points I, F, J and H. However, as the equality $\beta_1 = \beta_2$ is equivalent to the simultaneous inequalities $\beta_1 \geq \beta_2$ and $\beta_1 \leq \beta_2$, it can be conjectured that none of the suggested designs would be D-optimal on its own. Intuitively, the symmetry in parameters β_1 and β_2 generated by the equality $\beta_1 = \beta_2$ indicates that the candidate D-optimal designs are those with support points symmetric with respect to the logit line $u = 0$. Hence, the candidate D-optimal designs are the 4-point designs with support points A, F, J and H, or I, H, C and F, and the 6-point design with support points A, I, F, C, J and H. The candidate D-optimal designs for these 3 cases are discussed below.

Design with support points A, F, J and H

Consider Figure 6.13 (b). The line AJ is perpendicular to the logit line $u = 0$, line FH. In fact, it was found in Section 6.5 (see Figure 6.7) that the slope m_1 of line AJ is $m_1 = \dfrac{\beta_1(u - \beta_0 - \beta_1 b)}{\beta_2(u + \beta_0 + \beta_1 b)}$. Since $\beta_1 = \beta_2$ and $\beta_0 = -b\beta_1 = -b\beta_2$, then $m_1 = 1$. Clearly, the slope of the line FH is $m_2 = -1$. As $m_1 \times m_2 = -1$, then the line AJ is perpendicular to the line FH. Furthermore, the triangle AHJ is isosceles in H since the distances from A and J to H are both equal to $\dfrac{u}{\beta_1}$. Hence, points A and J are equidistant from the logit line $u = 0$, line FH, and it is reasonable to assume that the support points A and J will be equally weighted with weight $\dfrac{\lambda_1}{2}$. The design pattern $AFJH$ is not a parallelogram, and hence the support points F and H are expected to be unequally weighted. If λ_2 is the weight of support point F, then the weight of support point H must be $\lambda_3 = 1 - \lambda_1 - \lambda_2$. Thus, under the conditions $\beta_0 = -b\beta_1$

and $\beta_1 = \beta_2$, the candidate D-optimal design with support points A, F, J and H has the form

$$\xi = \left\{ \begin{array}{cccc} (b + \frac{bu}{\beta_0}, 0) & (0, b) & (b, -\frac{bu}{\beta_0}) & (b, 0) \\ \frac{\lambda_1}{2} & \lambda_2 & \frac{\lambda_1}{2} & 1 - \lambda_1 - \lambda_2 \end{array} \right\}. \tag{6.72}$$

The information matrix for the parameters β_0 and $\beta_1 = \beta_2$ evaluated at design (6.72) is

$$\begin{bmatrix} \frac{(1+e^u)^2 - \lambda_1(1-e^u)^2}{4(1+e^u)^2} & \frac{b\{\beta_0(1-\lambda_1-\lambda_2)(1+e^u)^2 + 2\lambda_1 e^u(u+2\beta_0)\}}{4\beta_0(1+e^u)^2} & \frac{b[\beta_0\lambda_2(1+e^u)^2 - 2\lambda_1 u e^u]}{4\beta_0(1+e^u)^2} \\ \frac{b\{\beta_0(1-\lambda_1-\lambda_2)(1+e^u)^2 + 2\lambda_1 e^u(u+2\beta_0)\}}{4\beta_0(1+e^u)^2} & \frac{b^2\{\beta_0^2(1-\lambda_1-\lambda_2)(1+e^u)^2 + 2\lambda_1 e^u[\beta_0^2 + (u+\beta_0)^2]\}}{4\beta_0^2(1+e^u)^2} & -\frac{\lambda_1 b^2 u e^u}{2\beta_0(1+e^u)^2} \\ \frac{b[\beta_0\lambda_2(1+e^u)^2 - 2\lambda_1 u e^u]}{4\beta_0(1+e^u)^2} & -\frac{\lambda_1 b^2 u e^u}{2\beta_0(1+e^u)^2} & \frac{b^2[\beta_0^2\lambda_2(1+e^u)^2 + 2\lambda_1 u^2 e^u]}{4\beta_0^2(1+e^u)^2} \end{bmatrix} \tag{6.73}$$

where $u > 0$ and $0 < \lambda_i < 1$ for $i = 1, 2$, and the determinant of the information matrix (6.73) is

$$D = \frac{b^4 \lambda_1 u^2 e^u \left\{\beta_0^2 \lambda_2 \left[(1 - \lambda_1 - \lambda_2)(1 + e^u)^2 + \lambda_1 e^u [u^2(1 - \lambda) + 4(1 + \lambda_2 \beta_0 u)]\right]\right\}}{16\beta_0^4(1 + e^u)^4}. \tag{6.74}$$

It is difficult to find the optimal values of u, and $\lambda_1, \lambda_2 \in (0, 1)$ which maximizes D given by (6.74) analytically. However, the following example indicates that the 4-point D-optimal design (6.72) can be obtained numerically.

Example 6.14. Consider the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, 2] \times [0, 2]$ where the parameter vector is $\boldsymbol{\beta} = (-4, 2, 2)^T$. Here $b = 2$. Since $\beta_1 = \beta_2 = 2 > \dfrac{2.1150}{b} = 1.0575$ and $\beta_0 + b\beta_1 = 0$, a 4-point D-design of the form (6.72) is expected. The Mathematica directive:

$$\texttt{NMaximize}[\{\texttt{D}, 0 < \lambda_1 < 1, 0 < \lambda_2 < 1, \texttt{u} > 0\}, \{\lambda_1, \lambda_2, \texttt{u}\}]$$

gives $D = 0.001317$ or $-\ln|M(\xi^*; \boldsymbol{\beta})| = 6.632$, $\lambda_1 = 0.4808$, $\lambda_2 = 0.3267$ and $u = 1.7644$. Hence, the candidate 4-point D-optimal design (6.72) is given by

$$\xi^* = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{cccc} (1.118, 0) & (0, 2) & (2, 0.882) & (2, 0) \\ 0.2404 & 0.3267 & 0.2404 & 0.1925 \\ -1.764 & 0 & 1.764 & 0 \end{array} \right\}. \tag{6.75}$$

The corresponding standardized variance function is

$$d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) = \frac{(90.879 - 85.611x_1 - 85.611x_2 + 23.086x_1^2 + 23.086x_2^2 + 39.439x_1x_2)e^u}{(1 + e^u)^2} \tag{6.76}$$

where $u = -4 + 2x_1 + 2x_2$. The support points of design (6.75) are displayed in Figure 6.16 (a). The standardized variance function (6.76) is plotted in Figure 6.17 (b). Clearly, $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) \leq 3$ with equality holding at 6 support points among which the 4 support points of design (6.75) form a subset.

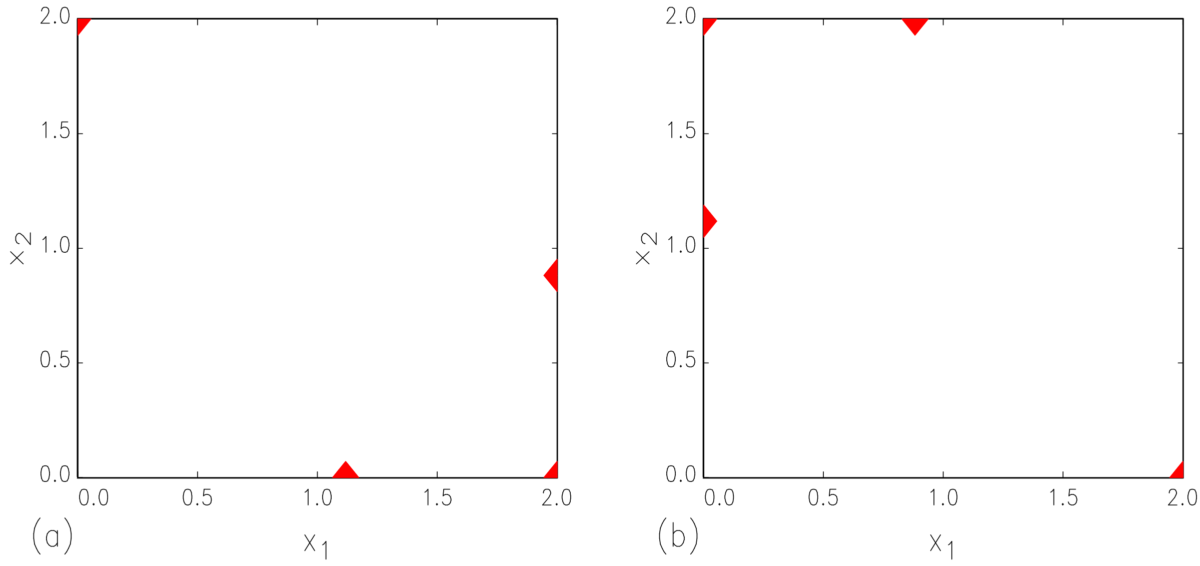


Figure 6.16: Support points of two equivalent 4-point D-optimal designs for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0,2] \times [0,2]$ and with parameter vector $\boldsymbol{\beta} = (-4, 2, 2)^T$.

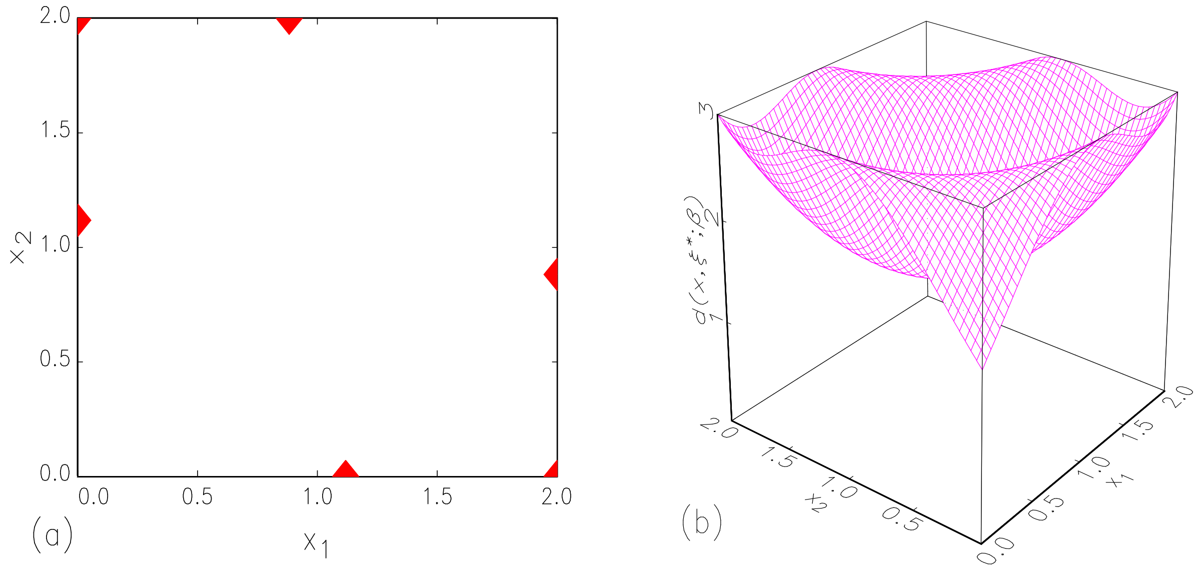


Figure 6.17: (a) Support points of the 6-point D-optimal design and (b) the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ given in (6.76) for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0,2] \times [0,2]$ and with parameter vector $\boldsymbol{\beta} = (-4, 2, 2)^T$.

Design with support points I, H, C and F

The same arguments which yielded design (6.72) can be used to construct a candidate D-optimal design with support points I and C each with weight $\frac{\lambda_1}{2}$, and support points H and F with respective weights λ_2 and $1-\lambda_1-\lambda_2$. Under the conditions $\beta_0 = -b\beta_1$ and $\beta_1 = \beta_2 > \dfrac{2.1150}{b}$, the form of the candidate D-optimal design is

$$\xi = \left\{ \begin{array}{cccc} (0, b+\frac{bu}{\beta_0}) & (b,0) & (-\frac{bu}{\beta_0}, b) & (0,b) \\ \frac{\lambda_1}{2} & \lambda_2 & \frac{\lambda_1}{2} & 1-\lambda_1-\lambda_2 \end{array} \right\} \tag{6.77}$$

which is obtained from design (6.72) by interchanging the coordinate values of the support points. Hence, the determinant of the information matrix for the parameters β_0 and $\beta_1 = \beta_2$ evaluated at design (6.77) is also (6.74). Then, the D-optimal values for $u > 0$ and $\lambda_1, \lambda_2 \in (0,1)$ for design (6.72) and the present design (6.77) are the same.

Example 6.15. (**Example 6.14 continued**). Using the results found in Example 6.14, the 4-point D-optimal design (6.77) is given by

$$\xi^* = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{cccc} (0, 1.118) & (0,2) & (0.882, 2) & (2,0) \\ 0.2404 & 0.1925 & 0.2404 & 0.3267 \\ -1.764 & 0 & 1.764 & 0 \end{array} \right\} \tag{6.78}$$

with the same criterion value $-\ln|M(\xi^*;\boldsymbol{\beta})| = 6.632$ as in Example 6.14. The support points of design (6.78) are displayed in Figure 6.16 (b), and the standardized variance function is again plotted in Figure 6.17 (b) and suggesting that the design (6.78) is D-optimal.

Design with support points A, I, F, C, J and H.

An example of a 6-point D-optimal design with support points A, I, F, C, J and H was constructed numerically by Atkinson and Haines (1996), and is reported in Atkinson et al. (2007, p. 403) to be a convex combination of two 4-point D-optimal designs. In the context of the design pattern of Figure 6.13 (b), the 6-point design is the convex combination

$$\text{Design } AIFCJH = \alpha(\text{Design } AFJH) + (1-\alpha)(\text{Design } \; IHCF)$$

where the mixing constant α is such that $0 \le \alpha \le 1$. As the analytic derivations were not given by Atkinson et al. (2007, p. 403), and comprehensive derivations are difficult, the aim of the following discussion is to sketch the rationale of the 6-point D-optimal design patterns and then provide evidence supporting the theory by a numerical example.

Consider Figure 6.13 (b). It was shown in the above paragraphs that the lines AJ and IC are perpendicular to the logit line $u = 0$, line FH, so that the support points A and J are equally weighted each with weight $\frac{\lambda_1}{2}$, and the support points I and C are equally weighted each with weight $\frac{\lambda_2}{2}$. If λ_3 is the weight of support point F, then the weight of support point H must be $\lambda_3 = 1-\lambda_1-\lambda_2$ so that the sum of the weights of the six support points equals 1. Hence, under the conditions $\beta_0 = -b\beta_1$ and $\beta_1 = \beta_2 > \frac{2.1150}{b}$, the candidate 6-point D-optimal design with support points A, I, F, C, J and H has the form

$$\xi = \left\{ \begin{array}{cccccc} (b+\frac{bu}{\beta_0}, 0) & (0, b+\frac{bu}{\beta_0}) & (0,b) & (-\frac{bu}{\beta_0}, b) & (b, -\frac{bu}{\beta_0}) & (b,0) \\ \frac{\lambda_1}{2} & \frac{\lambda_2}{2} & \lambda_3 & \frac{\lambda_2}{2} & \frac{\lambda_1}{2} & 1-\lambda_1-\lambda_2-\lambda_3 \end{array} \right\}. \tag{6.79}$$

The information matrix for the parameters β_0 and $\beta_1 = \beta_2$ evaluated at design (6.79) is given by

$$M(\xi;\boldsymbol{\beta}) = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \tag{6.80}$$

where

$$\begin{aligned}
m_{11} &= \frac{(1-\lambda_2)(1+e^u)^2-\lambda_1(1-e^u)^2}{4(1+e^u)^2} \\
m_{21} &= \frac{b\{\beta_0[(1-\lambda_2-\lambda_3)(1+e^u)^2-\lambda_2(1-e^u)^2]+2(\lambda_1-\lambda_2)ue^u\}}{4\beta_0(1+e^u)^2} = m_{12} \\
m_{22} &= \frac{b^2\{\beta_0^2[(1-\lambda_2-\lambda_3)(1+e^u)^2-\lambda_1(1-e^u)^2]+2ue^u[(\lambda_1+\lambda_2)u+2\lambda_1\beta_0]\}}{4\beta_0^2(1+e^u)^2} \\
m_{31} &= \frac{b[\lambda_3(1+e^u)^2+4\lambda_2e^u]-2ue^u(\lambda_1-\lambda_2)\}}{4\beta_0(1+e^u)^2} = m_{13} \\
m_{32} &= -\frac{b^2ue^u(\lambda_1+\lambda_2)}{2\beta_0(1+e^u)^2} = m_{23} \\
m_{33} &= \frac{b^2\{\beta_0^2[\lambda_3(1+e^u)^2]+2ue^u[(\lambda_1+\lambda_2)u+2\beta_0\lambda_2]\}}{4\beta_0^2(1+e^u)^2}
\end{aligned}$$

with $u > 0$ and $0 < \lambda_i < 1$ for $i = 1, 2, 3$.

The expression for the determinant of the information matrix (6.80) is a long completed expression and hence is not reported here. Furthermore, it is difficult to find the optimal values of $u > 0$, and λ_1, λ_2, $\lambda_3 \in (0,1)$ which maximize the determinant of the information matrix given by (6.80) analytically. However, the following numerical example suggests that the 6-point D-optimal design (6.79) can exist, and is a mixture of the 4-point D-optimal designs (6.72) and (6.77).

Example 6.16. (**Example 6.14 continued**). Consider Example 6.14 and suppose a 6-point D-optimal design has the form (6.79). The assumption of a 6-point design follows the fact that $\beta_0 = -b\beta_1 = -4$ and $\beta_1 = \beta_2 = 2 > \frac{2.1150}{b} = 1.0575$. Mathematica was used to maximize the determinant of the information (6.80) with respect to u and the three weights λ_1, λ_2 and λ_3

but failed to converge. However, numerical calculation using the Gauss program in Appendix B gives the following 6-point design

$$\xi^* = \left\{ \begin{array}{cccccc} (1.118, 0) & (0, 1.118) & (0, 2) & (0.882, 2) & (2, 0.882) & (2, 0) \\ 0.1306 & 0.1098 & 0.2654 & 0.1098 & 0.1306 & 0.2538 \\ -1.764 & -1.764 & 0 & 1.764 & 1.764 & 0 \end{array} \right\}. \tag{6.81}$$

Clearly, the patterns of the support points and corresponding weights agree with those conjectured theoretically. The weights, say $\lambda_i^{(3)}$, of the design given in (6.81) are the convex combinations

$$\lambda_i^{(3)} = \alpha\lambda_i^{(1)} + (1-\alpha)\lambda_i^{(2)} \tag{6.82}$$

where $\lambda_i^{(1)}$ are weights associated with the support of the design (6.75), and $\lambda_i^{(2)}$ are weights associated with the support of design given (6.78) for $i = 1, 2, \ldots, 6$, and α is the mixing constant. The expression on the right hand side of equation (6.82) is such $\lambda_i^{(j)} = 0$ if the ith support point in design (6.81) is not present in the jth design, for $i = 1, 2, \ldots, 6$ and $j = 1, 2$. A comparison of the weights in designs (6.75), (6.78) and (6.81) indicates that the mixing constant of the convex combination (6.82) is $\alpha = 0.5433$. For example, the support point $(1.118, 0)$ in design (6.81) is present in design (6.75) and absent in design (6.78), and thus the weight 0.1306 at the support point $(1.118, 0)$ in design (6.81) can be calculated as $\lambda_1^{(3)} = 0.5433 \times 0.2408 + (1 - 0.5433) \times 0 = 0.1306$. Similarly, the weight 0.1098 at the support point $(0, 1.118)$ in design (6.81) can be calculated as $\lambda_2^{(3)} = 0.5433 \times 0 + (1 - 0.5433) \times 0.2404 = 0.1098$. As the support point $(0, 2)$ in design (6.81) is present in both designs (6.75) and (6.78), the weight 0.2654 at $(0, 2)$ in design (6.81) can be calculated as $\lambda_3^{(3)} = 0.5433 \times 0.3267 + 0.4567 \times 0.1925 = 0.2654$. The weights at the remaining 3 support points in design (6.81) can also be derived from the weights of the same support points in designs (6.75) and (6.78). The support points of the design (6.81) are represented by symbols in Figure 6.17 (a). Note that superimposing one of Figure 6.16 (a) and Figure 6.16 (b) upon the other gives Figure 6.17 (a). The standardized variance function is still (6.76), and is again plotted in Figure 6.17 (b). The criterion value for the design (6.81) is $-\ln|M(\xi^*; \boldsymbol{\beta})| = 6.632$, i.e. the same as in the two 4-point D-optimal designs (6.75) and (6.78).

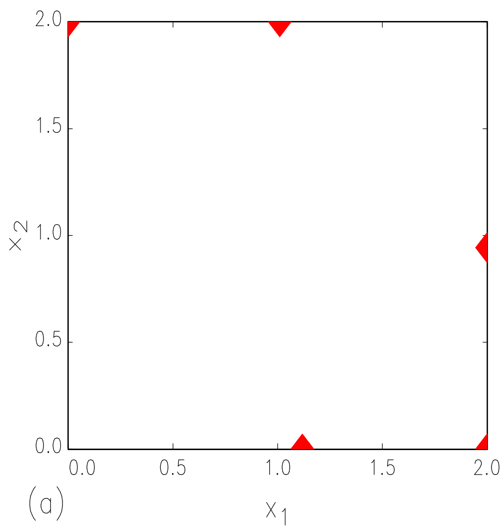
5-point D-optimal designs

Slight deviations from the condition $2\beta_0 + b(\beta_1 + \beta_2) = 0$ or $\beta_1 = \beta_2 > \frac{2.1150}{b}$ can still lead to a 6-point D-optimal design which is a mixture of two 4-point D-optimal designs, but, for brevity, examples are not discussed. For instance, the parameters $\beta_0 = -4$, $\beta_1 = 2.04$ and $\beta_2 = 2$ do not satisfy condition $2\beta_0 + b(\beta_1 + \beta_2) = 0$, but lead to a 6-point D-optimal design, but for brevity the example is not discussed. However, the departure from the conditions of a 6-point design can be such that no 6-point design is still D-optimal. The following example indicates that a small change of the parameter $\beta_2 = 2$ to $\beta_2 = 1.9$ leads to a 5-point D-optimal design.

Example 6.17. Consider the two-variable logistic model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0,2] \times [0,2]$ where the parameter vector is $\boldsymbol{\beta} = (-4, 2, 1.9)^T$. Here, $\beta = 2 > 1.0575$, and $\beta_2 = 1.9 > 1.0575$, but $2\beta_0 + b(\beta_1 + \beta_2) = -0.2 < 0$. In this case the logit line $\beta_0 + \beta_1 x_1 + \beta_2 x_2 = 0$ does not pass through the centre $(1,1)$ of the design space $[0,2] \times [0,2]$, but has been slightly translated upward. Numerical calculation gives the following 5-point design

$$\xi^* = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{ccccc} (1.118, 0) & (0, 2) & (1.010, 2) & (2, 0) & (2, 0.943) \\ 0.2023 & 0.3097 & 0.0873 & 0.2233 & 0.1774 \\ -1.763 & -0.2 & 1.820 & 0 & 1.791 \end{array} \right\}. \tag{6.83}$$

The support points of the design (6.83) are represented by symbols in Figure 6.18 (a). Figure 6.18 (b) gives a plot of the standardized variance function $d(\boldsymbol{x}, \xi^*; \beta)$. The plot suggests that the design is globally optimal.



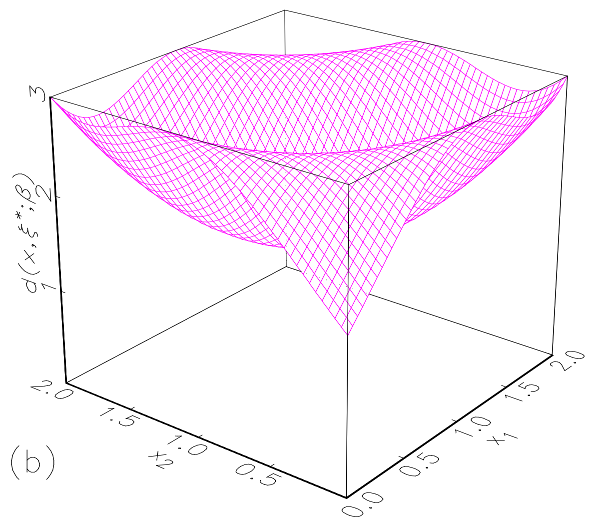


Figure 6.18: (a) Support points of the D-optimal design and (b) the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0,2] \times [0,2]$ and with parameter vector $\boldsymbol{\beta} = (-4, 2, 1.9)^T$.

6.7. Summary chart for the case when $2\beta_0 + b(\beta_1 + \beta_2) = 0$

The D-optimal parallelogram designs in Section 6.5 and the D-optimal designs in Section 6.6 share the condition that $2\beta_0 + (\beta_1 + \beta_2)b = 0$. Figure 6.19 displays a chart that summarizes cases when $2\beta_0 + (\beta_1 + \beta_2)b = 0$. The case of $b = 2$ was taken for illustration purpose and the pattern will be similar for any other value of b.

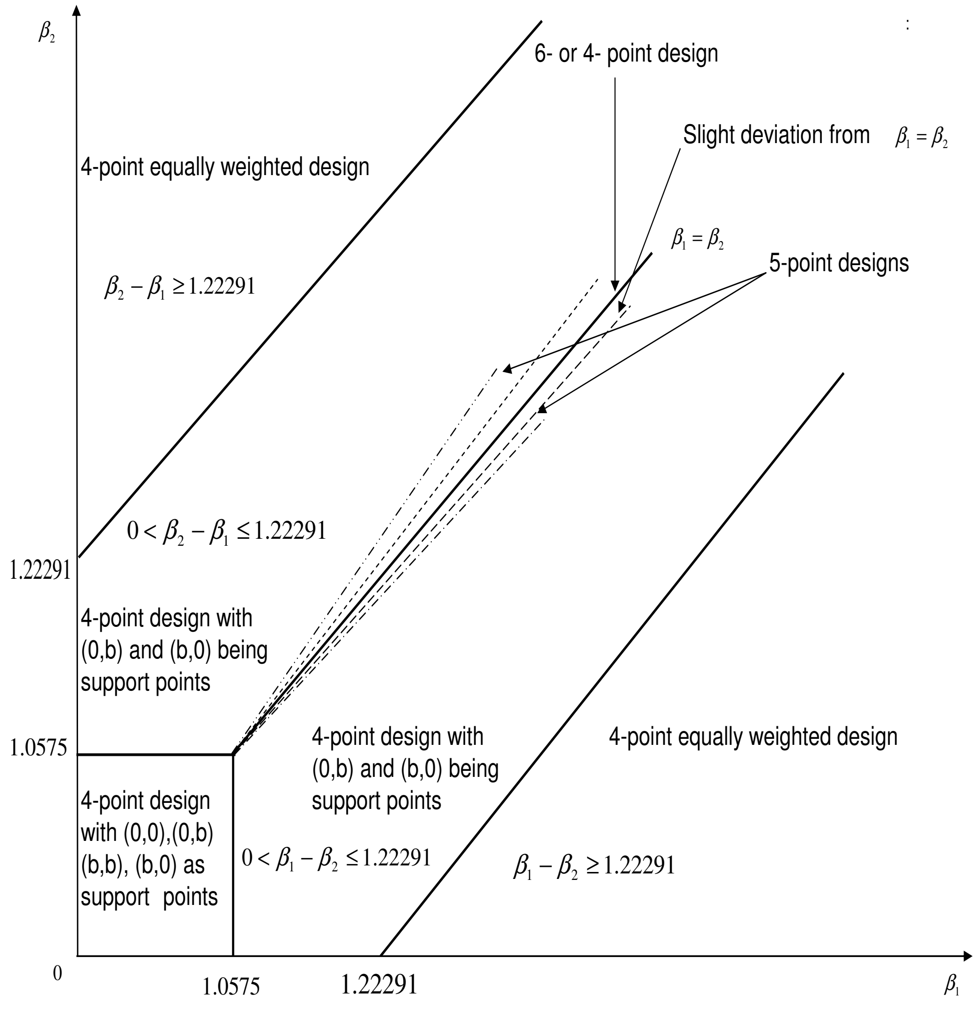


Figure 6.19: Pattern and number of support points of the D-optimal design for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, b] \times [0, b]$ with $b = 2$ and when $2\beta_0 + (\beta_1 + \beta_2)b = 0$.

6.8. Non-parallelogram designs

6.8.1 Non-parallelogram designs with 1 or 2 nonconsecutive support points at the vertices of the design space

In Section 6.5, it was shown that for the two-variable binary logistic model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ with $\beta_1 \neq \beta_2$, the necessary condition for the existence of a D-optimal parallelogram design is that $2\beta_0 + (\beta_1 + \beta_2)b = 0$, or equivalently, that the logit line $u = 0$ passes through the centre of the design space $[0, b] \times [0, b]$. If the condition does not hold, then the D-optimal designs, if they exist, are non-parallelogram designs. Examples of such designs are the trapezium and related designs discussed in Section 6.3. For these designs, the logit line $u = 0$ intersects two non parallel sides of the design space $[0, b] \times [0, b]$, or equivalently $\beta_0 + \beta_1 b$ and $\beta_0 + \beta_2 b$ have same sign. In addition, it was found in Section 6.6 that the D-optimal designs associated with conditions $2\beta_0 + (\beta_1 + \beta_2)b = 0$ and $\beta_1 = \beta_2 > 1.0575$ are not parallelogram designs, but have 5 or 6 support points.

In this section, non-parallelogram D-optimal designs for the case when the logit line $u = 0$ does not pass through the centre of the design space $[0, b] \times [0, b]$ but intersects two parallel sides of the design space are considered. For example, see Figures 6.20 (a) to (c) in which the logit line $u = 0$ intersects the horizontal sides of the design space $[0, b] \times [0, b]$. Similar design patterns, not presented here, are those in which the logit line $u = 0$ intersects the vertical sides of the design space.

The analytical construction of D-optimal designs with design patterns such as those displayed in Figures 6.20 (a) to (c) is expected to be more difficult than the construction of D-optimal parallelogram or trapezium designs because of the asymmetry of the designs induced by the logit $u = 0$ passing off centre of the design space. However, conditions for the existence of the D-optimal designs can be derived as in Section 6.4. This section considers the construction of 4-point D-optimal designs with patterns in Figures 6.20 (a) to (c) only. Four-point designs with 3 support points at the vertices of the design space are discussed in Section 6.8.2.

<u>Case of Figure 6.20 (a) design pattern</u>

Consider Figure 6.20 (a). Using similar arguments to those used in Sections 6.4 and 6.5, the candidate D-optimal design with the design pattern in Figure 6.20 (a) can only have support

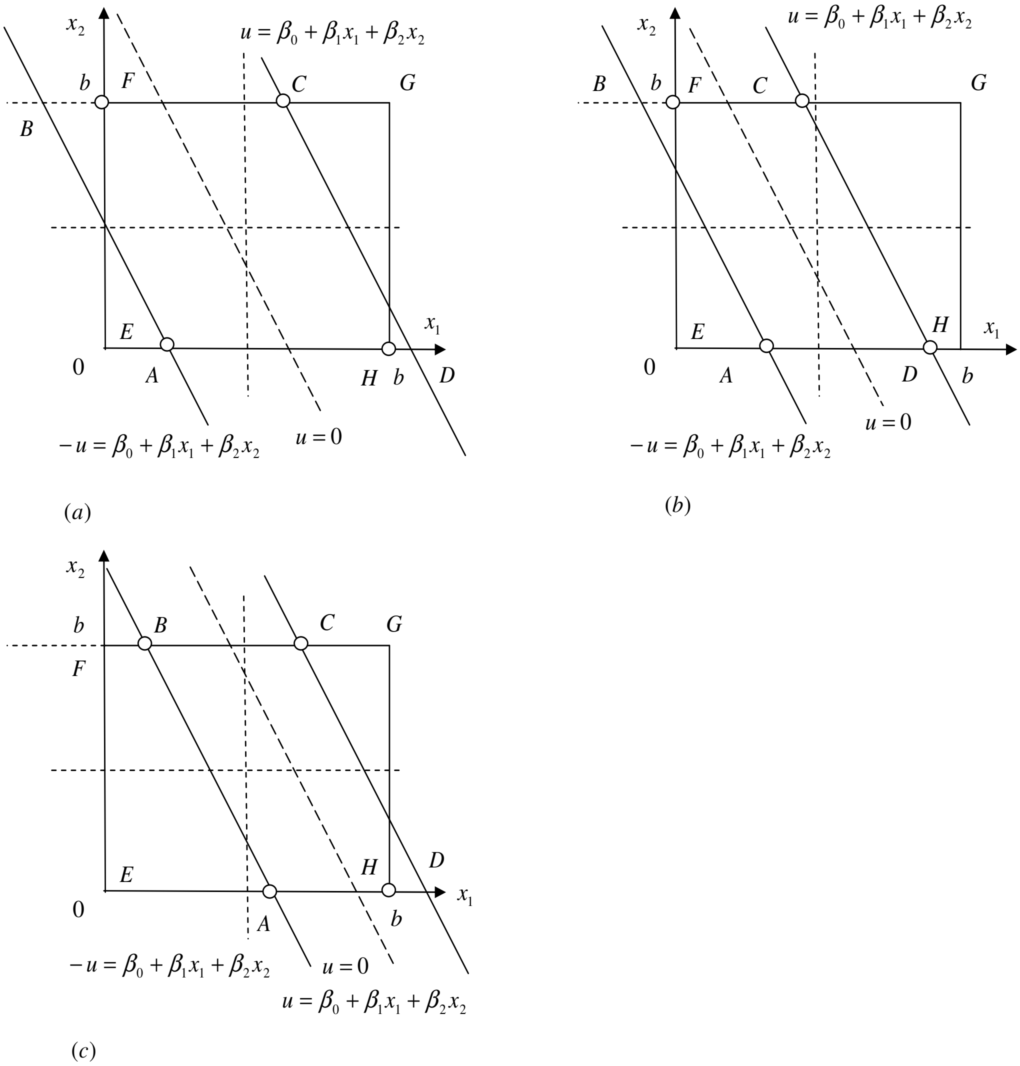


Figure 6.20: 4-point D-optimal designs for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, b] \times [0, b]$ with $2\beta_0 + (\beta_1 + \beta_2)b \neq 0$ and one or two nonadjacent vertices support points.

points indicated by circles. In this case the following conditions hold:

$$\beta_0 \leq -u^* \leq \beta_0 + \beta_2 b < 0 \text{ and } 0 < \beta_0 + \beta_1 b \leq u^* \leq \beta_0 + (\beta_1 + \beta_2)b$$

where $u^* = 1.22291$. The two conditions can be summarized as

$$0 < \max\{-\beta_0 - \beta_2 b, \beta_0 + \beta_1 b\} \leq 1.22291 \leq \min\{-\beta_0, \beta_0 + (\beta_1 + \beta_2)b\}. \tag{6.84}$$

Clearly, the respective candidate optimal logits at the support points F and H are $\beta_0 + b\beta_2$ and $\beta_0 + b\beta_1$. As the design pattern in Figure 6.20 (a) is the same as the design pattern in Figure 6.9 (a) but with condition (6.35) not satisfied, the candidate D-optimal design is expected to be an unequally weighted 4-point design of the form

$$\xi^* = \left\{ \begin{array}{cccc} (\frac{-u_1^*-\beta_0}{\beta_1}, 0) & (0, b) & (\frac{u_2^*-\beta_0-b\beta_2}{\beta_1}, b) & (b, 0) \\ \lambda_1 & \lambda_2 & \lambda_3 & 1-\lambda_1-\lambda_2-\lambda_3 \end{array} \right\} \tag{6.85}$$

where $u_1^* \neq u_2^*$, $u_1^* > 0$, $u_2^* > 0$, and the design weights λ_i with $0 < \lambda_i < 1$ $(i = 1, \ldots, 4)$ are not necessarily equal.

Example 6.18. Consider the two-variable binary logistic model $u = \text{logit}(p) = \beta_0+\beta_1 x_1+\beta_2 x_2$ on the design space $[0, 2] \times [0, 2]$ where the parameter vector is $\boldsymbol{\beta} = (-3, 2, 1.4)^T$. In this example, $2\beta_0 + 2(\beta_1 + \beta_2) = 0.8 \neq 0$. Therefore condition (6.35) is not satisfied. Furthermore

$$0 < \max\{-\beta_0 - 2\beta_2 = 0.2, \beta_0 + 2\beta_1 = 1\} = 1 < 1.22291 < \min\{-\beta_0 = 3, \beta_0 + 2(\beta_1 + \beta_2) = 4.6\} = 4.6.$$

Thus, condition (6.84) holds and an unequally weighted 4-point design with two vertices as support points $(0, 2)$ and $(2, 0)$ is expected. Numerical calculations give the following unequally weighted 4-point design as expected:

$$\xi^* = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{cccc} (0.665, 0) & (0, 2) & (0.850, 2) & (2, 0) \\ 0.2937 & 0.2774 & 0.1212 & 0.3077 \\ -1.671 & -0.2 & 1.498 & 1 \end{array} \right\}. \tag{6.86}$$

The support points of the design (6.86) are represented by symbols in Figure 6.21 (a). Clearly these support points follow the shape of the quadrilateral $AFCH$ in Figure 6.20 (a), and thus are not vertices of a parallelogram. The standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$, for all $\boldsymbol{x}^T = (x_1, x_2) \in [0, 2] \times [0, 2]$, represented in Figure 6.21 (b) suggests that the design (6.86) is D-optimal.

Case of Figure 6.20 (b) design pattern

Again using similar arguments to those used in Sections 6.4 and 6.5, the candidate D-optimal designs with the design pattern in Figure 6.20 (b) have support points indicated by circles. In this case the following conditions hold:

$$\beta_0 \leq -u^* \leq \beta_0 + \beta_2 b < 0 \text{ and } 0 < u^* \leq \beta_0 + \beta_1 b$$

where $u^* = 1.22291$. The two conditions can be summarized as

$$0 < -\beta_0 - \beta_2 b \leq 1.22291 \leq \min\{-\beta_0, \beta_0 + \beta_1 b\}. \tag{6.87}$$

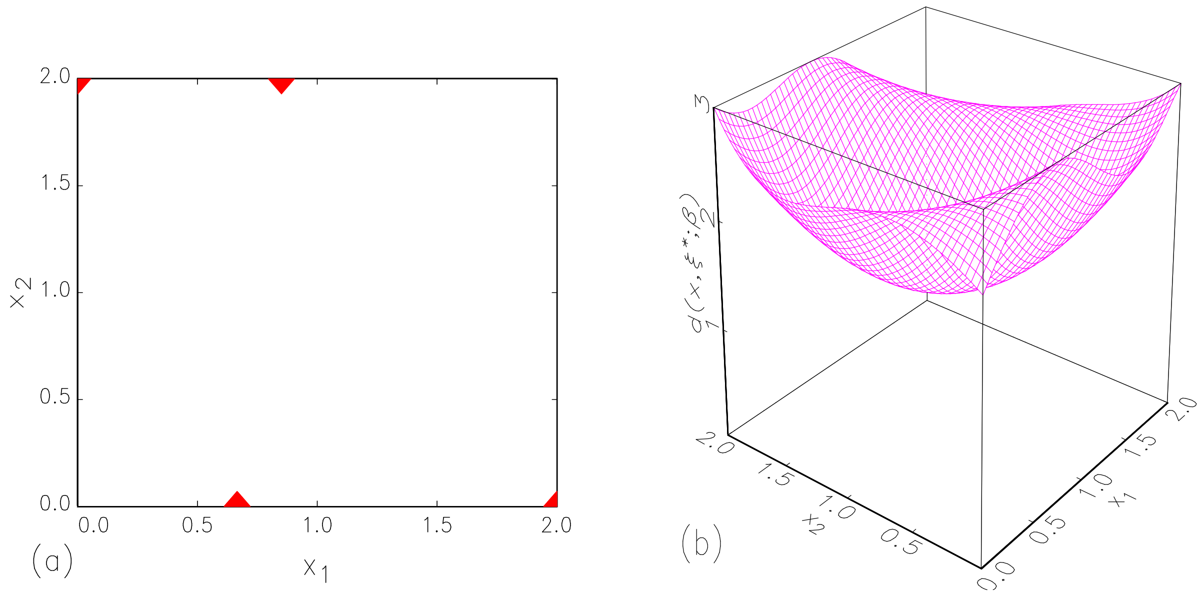


Figure 6.21: (a) Support points and (b) the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, 2] \times [0, 2]$ and with parameter vector $\boldsymbol{\beta} = (-3, 2, 1.4)^T$.

Clearly, the optimal logit at support point F is $\beta_0 + b\beta_2$. The support points A and D are symmetric with respect to $u = 0$, and hence are expected to have equal weights as in Figures 6.8 (a) and 6.8 (b). Hence, the expected candidate D-optimal design corresponding to the design pattern of Figure 6.20 (b) is

$$\xi^* = \left\{ \begin{array}{cccc} (\frac{-u_1^* - \beta_0}{\beta_1}, 0) & (0, b) & (\frac{u_2^* - \beta_0 - b\beta_2}{\beta_1}, b) & (\frac{u_2^* - \beta_0}{\beta_1}, 0) \\ \frac{\lambda_1}{2} & \lambda_2 & 1 - \lambda_2 & \frac{\lambda_1}{2} \end{array} \right\} \tag{6.88}$$

where $u_1^* \neq u_2^*$, $u_1^* > 0$, $u_2^* > 0$, and the design weights λ_i with $0 < \lambda_i < 1$ $(i = 1, \ldots, 2)$ are not necessarily equal.

Example 6.19. Consider the two-variable binary logistic model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, 2] \times [0, 2]$ where $\beta_0 = -3$, $\beta_1 = 3$ and $\beta_2 = 1$. Here, $b = 2$, $-\beta_0 = 3$, $-\beta_0 - \beta_2 b = 1$, and $\beta_0 + \beta_1 b = 3$ so that

$$0 < -\beta_0 - \beta_2 b < 1.22291 < \min\{-\beta_0, \beta_0 + \beta_1 b\}.$$

Therefore condition (6.87) is satisfied, and thus the suitable candidate D-optimal design is of the form (6.88). Hence one of the optimal logits is $\beta_0 + \beta_2 b = -1$. Numerical calculations give the following 4-point candidate D-optimal design

$$\xi^* = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{cccc} (0.574, 0) & (0, 2) & (0.755, 2) & (1.426, 0) \\ 0.2608 & 0.2497 & 0.2287 & 0.2608 \\ -1.279 & -1 & 1.265 & 1.279 \end{array} \right\}. \quad (6.89)$$

The support points of the candidate D-optimal design (6.89) are represented by symbols in Figure 6.22 (a). Observe that these support points are not vertices of a parallelogram and that points $(0.5737, 0)$ and $(1.4263, 0)$ have equal weights as expected. In addition, the candidate D-optimal design has only one vertex support point at $(0, 2)$. The standardized variance

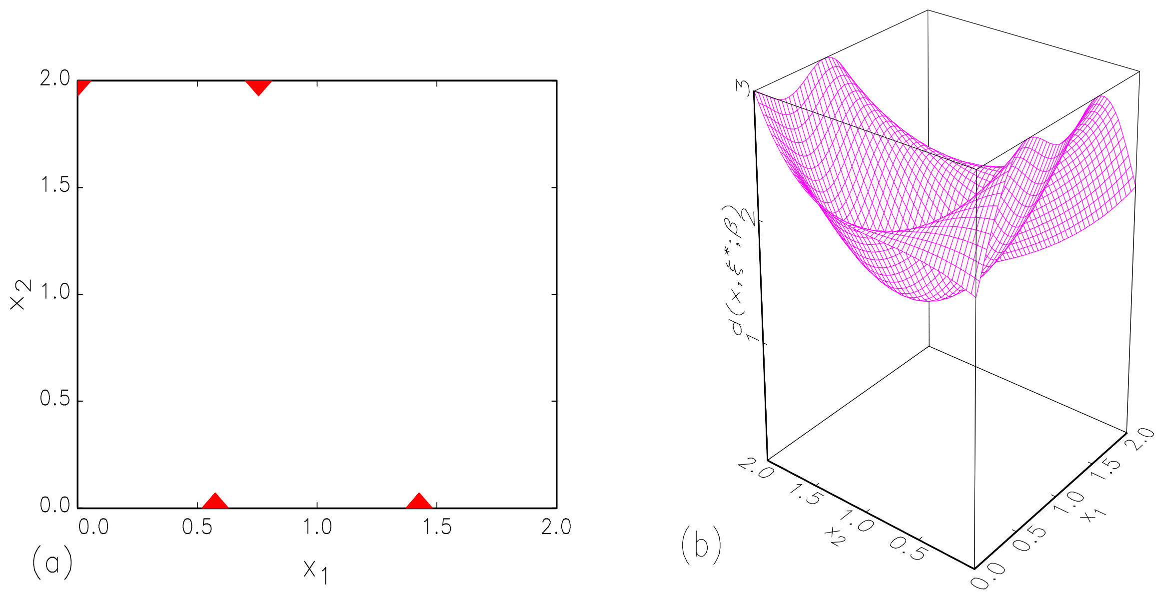


Figure 6.22: (a) Support points and (b) the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, 2] \times [0, 2]$ and with parameter vector $\boldsymbol{\beta} = (-3, 3, 1)^T$.

function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ represented in Figure 6.22 (b) has a value of 3 at the four support points and values less than 3 elsewhere in $[0, 2] \times [0, 2]$, and thus the design (6.89) is demonstrated to be D-optimal, at least numerically.

Case of Figure 6.20 (c) design pattern

The design pattern of the support points in Figure 6.20 (c) can be obtained from those of Figure 6.20 (b) by symmetric reflection about the centre of the design space $[0, b] \times [0, b]$. Using the same arguments as those developed for the previous case, the design of Figure 6.20

(c) holds provided the following condition is satisfied

$$0 < \beta_0 + \beta_1 b \le 1.22291 \le \min\{-\beta_0 - \beta_2 b, \beta_0 + (\beta_1 + \beta_2) b\}. \tag{6.90}$$

The expected candidate D-optimal design corresponding to the design pattern of Figure 6.20 (c) is thus given by

$$\xi^* = \left\{ \begin{array}{cccc} (\frac{-u_1^*-\beta_0}{\beta_1}, 0) & (\frac{-u_1^*-\beta_0-b\beta_2}{\beta_1}, b) & (\frac{u_2^*-\beta_0-b\beta_2}{\beta_1}, b) & (b, 0) \\ \lambda_1 & \frac{\lambda_2}{2} & 1\frac{\lambda_2}{2} & 1-\lambda_1-\lambda_2 \end{array} \right\} \tag{6.91}$$

where $u_1^* \neq u_2^*$, $u_1^* > 0$, $u_2^* > 0$, and the design weights λ_i with $0 < \lambda_i < 1$ $(i = 1, \ldots, 2)$ are not necessarily equal.

Example 6.20. Consider the two-variable binary logistic model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, 2] \times [0, 2]$ where $\beta_0 = -3$, $\beta_1 = 2$ and $\beta_2 = 0.5$. For this example,

$$0 < \beta_0 + \beta_1 b = 1 < 1.22291 < \min\{-\beta_0 - \beta_2 b, \beta_0 + (\beta_1 + \beta_2) b\} = 2.$$

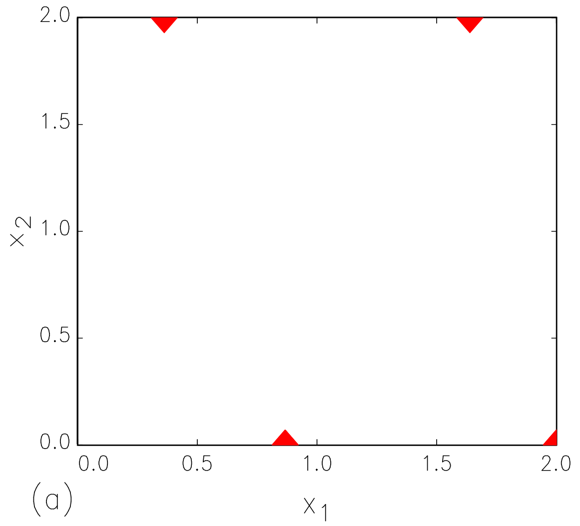
Therefore, condition (6.90) is satisfied. Hence, the expected candidate D-optimal design is of the form (6.91). One of the optimal logits is $\beta_0 + 2\beta_1 = 1$. Numerical calculations give the following 4-point candidate D-optimal design:

$$\xi^* = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{cccc} (0.867, 0) & (0.361, 2) & (1.639, 2) & (2, 0) \\ 0.2287 & 0.2608 & 0.2608 & 0.2497 \\ -1.265 & -1.279 & 1.279 & 1 \end{array} \right\}. \tag{6.92}$$

This design satisfies the conjectures in terms of support points and weights. The support points of design (6.92) are represented by symbols in Figure 6.23 (a) and clearly these points are not vertices of a parallelogram. The standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ represented by Figure 6.23 (d) suggests that the design (6.92) is D-optimal.

6.8.2 Non-parallelogram designs with 2 or 3 consecutive support points at the vertices of the design space

This section discusses the construction of 4-point non-parallelogram D-optimal designs with two or three adjacent support points at the vertices of the design space $[0, b] \times [0, b]$. The patterns of the designs is displayed in Figures 6.24 (a) to (d). Note that for all the four design patterns, the logit line $u = 0$ passes off the centre of the design space, but intersects the two horizontal sides. Furthermore, one of the logit lines $-u$ and u passes outside the design space.



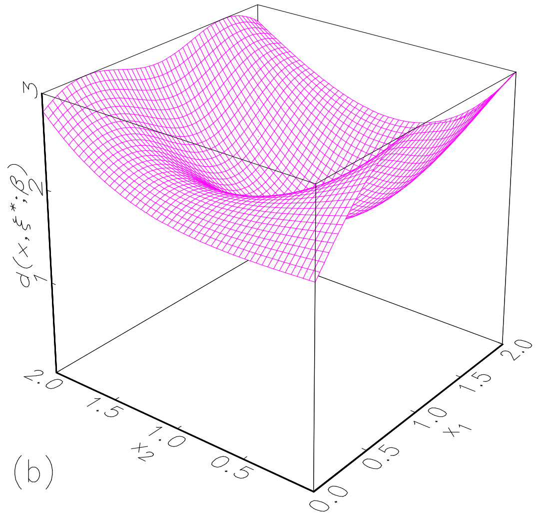


Figure 6.23: (a) Support points and (b) the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, 2] \times [0, 2]$ and with parameter vector $\boldsymbol{\beta} = (-3, 2, 0.5)^T$.

Case of Figure 6.24 (a) design pattern

Consider Figure 6.24 (a). On the design space $[0, \infty) \times [0, \infty)$ and under the condition that $0 \leq -\beta_0 \leq 1.5434$, the D-optimal design is known to be an equally weighted 3-point design (see Section 5.3.3). Following Sitter and Torsney (1995a), and Jia and Myers (2001), on the design space $(-\infty, \infty) \times [0, b]$, the support points of the candidate D-optimal design are A, B, C, and D, and lie on the optimal logit lines $u^* = \pm 1.22291$ (see Section 5.2). Now, on the design space $[0, b] \times [0, b]$, the points A, B, and D are outside the design space. Using the same arguments as in Section 6.8.1 for the design space $[0, b] \times [0, b]$, the candidate D-optimal designs has support points E, F, C, and H in Figure 6.24 (a). Hence, 3 of the 4 candidate optimal logits are $u = \beta_0$ passing through point E, $u = \beta_0 + \beta_1 b$ passing through point H, and $u = \beta_0 + \beta_2 b$ passing through point F. The candidate optimal logit passing through point C, and the unequal weights of the support points have to be calculated. The conditions for the existence of the design pattern of Figure 6.24 (a) are

$$-1.5434 \leq \beta_0 < \beta_0 + \beta_2 b \leq 0 \text{ and } 0 < \beta_0 + \beta_1 b \leq 1.5434 < \beta_0 + (\beta_1 + \beta_2) b$$

and the two conditions can be summarized as

$$0 < \max\{-\beta_0, \beta_0 + \beta_1 b\} \leq 1.5434 < \beta_0 + (\beta_1 + \beta_2) b. \tag{6.93}$$

Example 6.21. Consider the two-variable binary logistic model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$

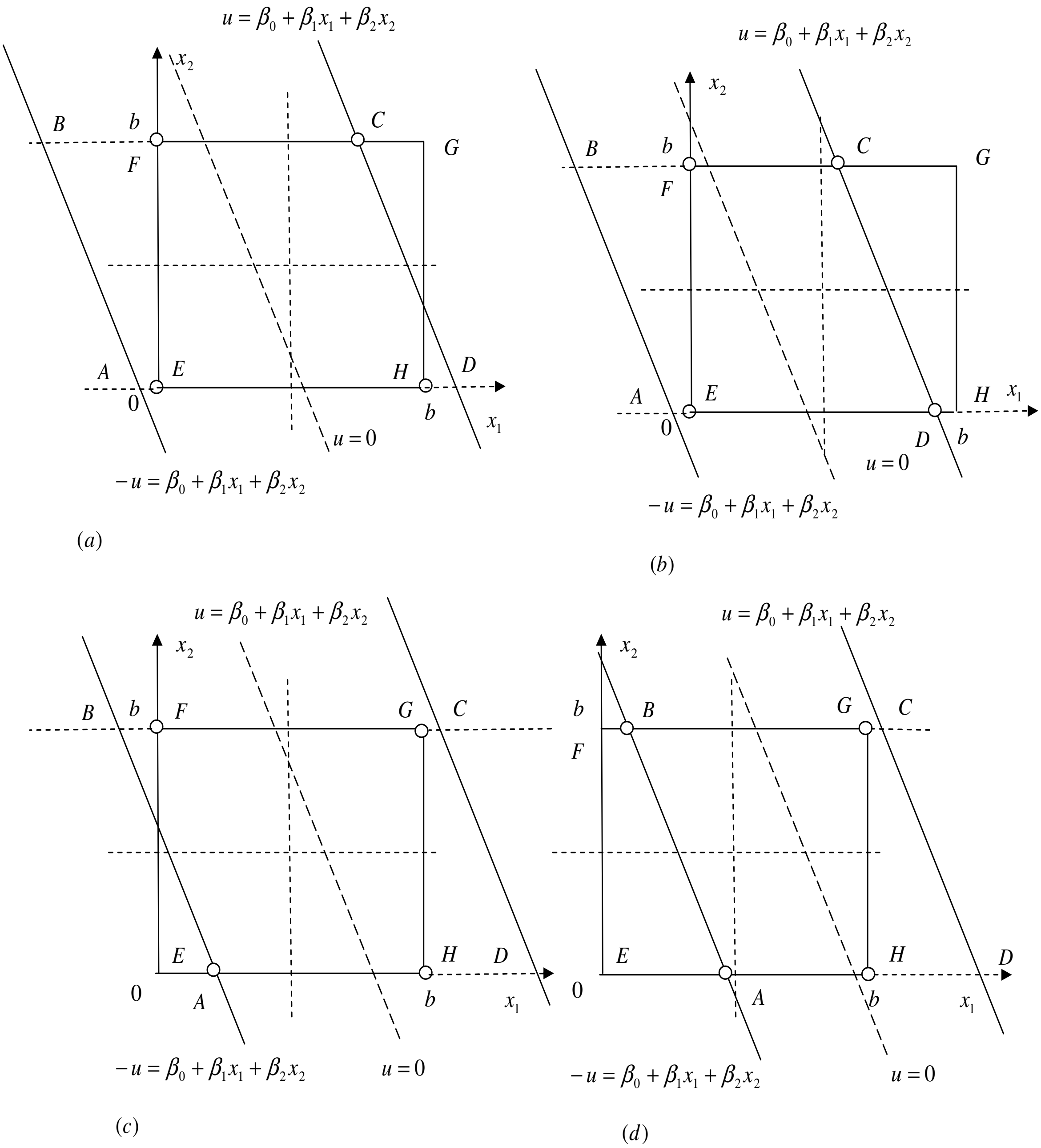


Figure 6.24: 4-point D-optimal designs for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, b] \times [0, b]$ with $2\beta_0 + (\beta_1 + \beta_2)b \neq 0$ and two or three adjacent vertices support points.

on the design space $[0, 2] \times [0, 2]$ where $\beta_0 = -1.4$, $\beta_1 = 1.2$ and $\beta_2 = 0.6$. In this case,

$$0 < \max\{-\beta_0 = 1.4, \beta_0 + \beta_1 b = 1\} = 1.4 < 1.5434 < \beta_0 + 2(\beta_1 + \beta_2) = 2.2.$$

Thus, condition (6.93) is satisfied and hence the design pattern of Figure 6.24 (a) is conjectured to hold. The D-optimal design corresponding to this case is

$$\xi^* = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{cccc} (0,0) & (0,2) & (1.516,2) & (2,0) \\ 0.2864 & 0.2797 & 0.1371 & 0.2968 \\ -1.4 & -0.2 & 1.6194 & 1 \end{array} \right\}. \quad (6.94)$$

The support points of the design (6.94) are represented by symbols in Figure 6.25 (a) and

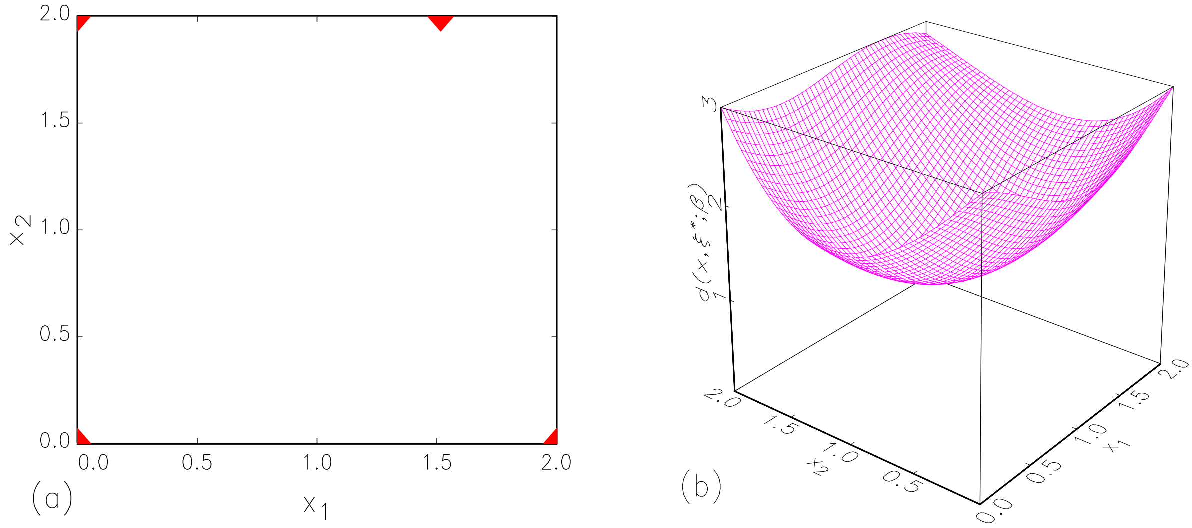


Figure 6.25: (a) Support points and (b) the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ of the D-optimal design for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, 2] \times [0, 2]$ and with parameter vector $\boldsymbol{\beta} = (-1.4, 1.2, 0.6)^T$.

clearly, are not vertices of a parallelogram. As expected no pair of weights are equal for the design (6.94). Furthermore, Figure 6.25 (b) demonstrates that the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) = 3$ at the 4 support points and $d(u, \xi^*; \boldsymbol{\beta}) < 3$ elsewhere as required for an optimal design. Hence, the design (6.94) is demonstrated to be D-optimal, at least numerically.

Case of Figure 6.24 (b) design pattern

Consider Figure 6.24 (b). Using similar arguments to those used in the case of Figure 6.24 (a), the support points of the candidate D-optimal design corresponding to Figure 6.24 (b) are E, F, C and D. Hence, 2 of the 4 candidate optimal logits are $u = \beta_0$ passing through point E and $u = \beta_0 + \beta_2 b$ passing through point F. The candidate optimal logit passing through points C and D, and the unequal weights of the support points have to be calculated numerically. The conditions for the existence of the design pattern of Figure 6.24 (b) are

$$-1.5434 \leq \beta_0 < \beta_0 + \beta_2 b \leq 0 \text{ and } 0 < 1.5434 \leq \beta_0 + \beta_1 b$$

and the two conditions can be summarized as

$$0 \leq -\beta_0 \leq 1.5434 \leq \beta_0 + \beta_1 b. \tag{6.95}$$

Example 6.22. Consider the two-variable binary logistic model $u = \text{logit}(p) = \beta_0+\beta_1 x_1+\beta_2 x_2$ on the design space $[0, 2] \times [0, 2]$ where $\beta_0 = -1.4$, $\beta_1 = 1.8$ and $\beta_2 = 0.6$. This example is the same as Example 6.21 except that the value of β_1 changes from 1.2 to 1.8. This implies that $0 < -\beta_0 = 1.4 < 1.5434 < \beta_0 + \beta_1 b = 2.2$ and thus condition (6.95) holds, and hence the candidate D-optimal design is that with the pattern displayed in Figure 6.24 (b). Calculations gives the following candidate D-optimal design

$$\xi^* = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{cccc} (0, 0) & (0, 2) & (0.840, 2) & (1.618, 0) \\ 0.3117 & 0.2749 & 0.1034 & 0.3100 \\ -1.4 & -0.2 & 1.312 & 1.512 \end{array} \right\}. \tag{6.96}$$

Observe that no pairs of the 4 support points have equal design weights. The support points of the design (6.96) are represented by symbols in Figure 6.26 (a). As expected these support points are not vertices of a parallelogram. Figure 6.26 (b) suggests that the design (6.96) is D-optimal since the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) = 3$ at each of the 4 support points of the design (6.96), and $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) < 3$ elsewhere in $[0, 2] \times [0, 2]$. Thus, the design (6.96) is demonstrated to be D-optimal, at least numerically.

Case 3 of Figure 6.24 (c) design pattern

Consider Figure 6.24 (c). The design pattern of Figure 6.24 (c) is the symmetric reflection of the design pattern of Figure 6.24 (a) with respect to the centre of the design space $[0, b] \times [0, b]$. Hence, the support points of the candidate D-optimal design are A, F, G, and H. Reasoning in a similar way as for the design pattern of Figure 6.24 (a), the condition for the existence of the design pattern of Figure 6.24 (c) is

$$0 < \max\{-\beta_0 - \beta_2 b, \beta_0 + (\beta_1 + \beta_2)b\} \leq 1.5434 < -\beta_0. \tag{6.97}$$

Example 6.23. Consider the two-variable logistic model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, 2] \times [0, 2]$ where $\beta_0 = -2$, $\beta_1 = 1.2$ and $\beta_2 = 0.5$. For this example, condition (6.97) holds since

$$0 < \max\{-\beta_0 - \beta_2 b = 1, \beta_0 + (\beta_1 + \beta_2)b = 1.4\} = 1.4 < 1.5434 < -\beta_0 = 2.$$

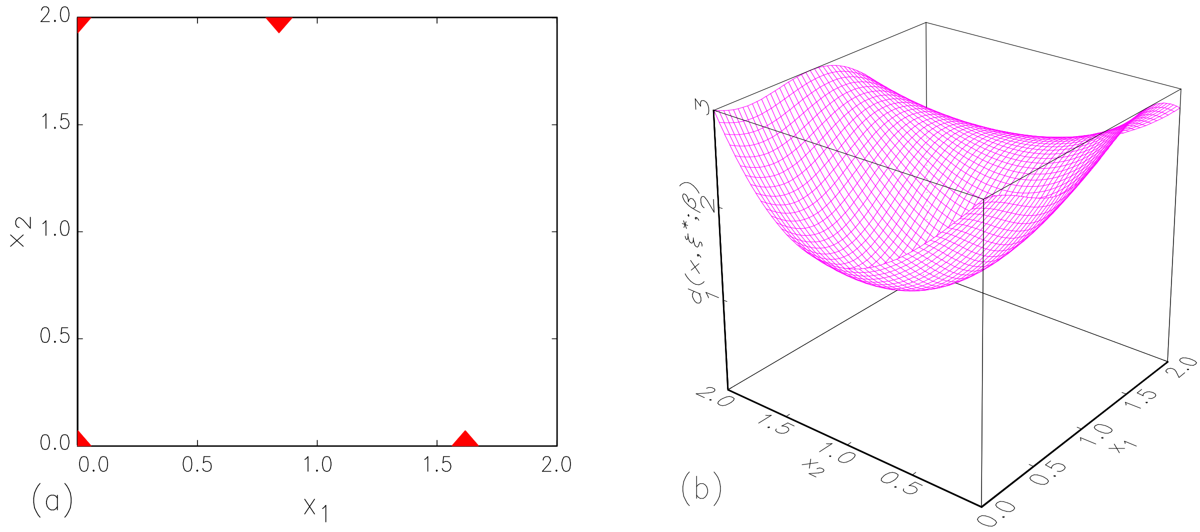


Figure 6.26: (a) Support points and (b) the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ of the D-optimal design for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, 2] \times [0, 2]$ and with parameter vector $\boldsymbol{\beta} = (-1.4, 1.8, 0.6)^T$.

Therefore, a design pattern of the form displayed in Figure 6.24 (c) is conjectured. The candidate D-optimal design for this case is

$$\xi^* = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{cccc} (0.362, 0) & (0, 2) & (2, 2) & (2, 0) \\ 0.1697 & 0.2860 & 0.2715 & 0.2728 \\ -1.566 & -1 & 1.4 & 0.4 \end{array} \right\}. \tag{6.98}$$

The support points of the design (6.98) are represented by symbols in Figure 6.27 (a) and the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ is represented in Figure 6.27 (b). Figure 6.27 (a) illustrates that the support points of the design (6.98) have the pattern of Figure 6.24 (c) and Figure 6.27 (b) demonstrates, at least numerically, that the design (6.98) is D-optimal.

Case of Figure 6.24 (d) design pattern

Consider Figure 6.24 (d). The design pattern of Figure 6.24 (d) is the symmetric reflection of the design pattern of Figure 6.24 (b) with respect to the centre of the design space $[0, b] \times [0, b]$. Hence the support points of the candidate D-optimal design are A, B, G and H. Reasoning in a similar way as for the design pattern of Figure 6.24 (b), the condition for the existence of the design pattern of Figure 6.24 (d) is

$$0 < \beta_0 + (\beta_1 + \beta_2)b \leq 1.5434 < -\beta_0 - \beta_2 b. \tag{6.99}$$

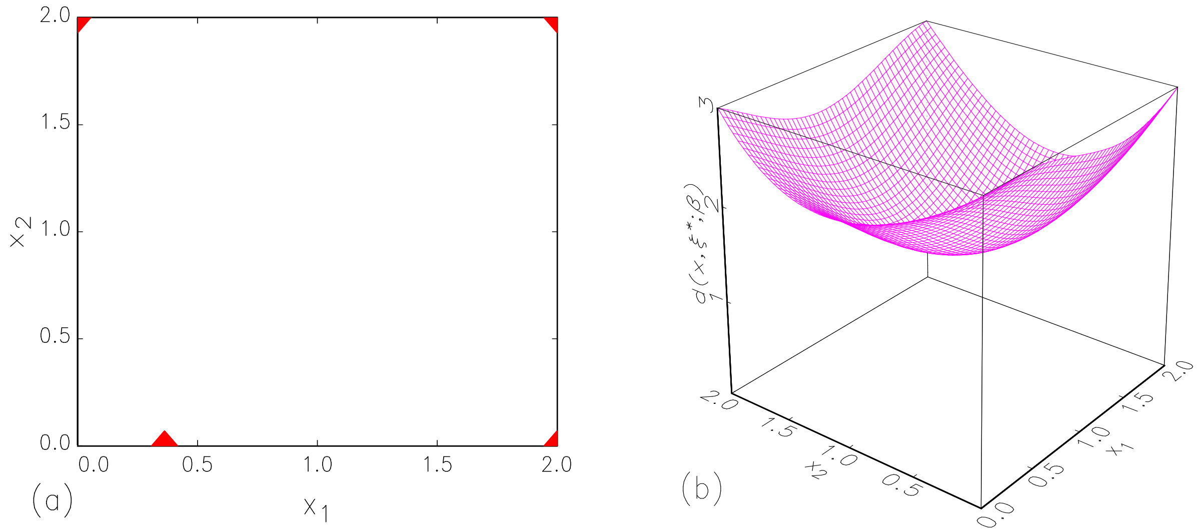


Figure 6.27: (a) Support points and (b) the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ of the D-optimal design for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ and with parameter vector is $\boldsymbol{\beta} = (-2, 1.2, 0.5)^T$ on the design space $[0, 2] \times [0, 2]$.

Example 6.24. Consider the two-variable logistic model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, 2] \times [0, 2]$ where $\beta_0 = -3$, $\beta_1 = 1.6$ and $\beta_2 = 0.5$. In this example,

$$0 < \beta_0 + (\beta_1 + \beta_2)b = 1.2 < 1.5434 < -\beta_0 - \beta_2 b = 2.$$

Therefore, condition (6.99) holds. Hence the appropriate design pattern is of the form indicated in Figure 6.24 (d). The corresponding candidate D-optimal design is conjectured to be

$$\xi^* = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{cccc} (1.002, 0) & (0.255, 2) & (2, 2) & (2, 0) \\ 0.111 & 0.303 & 0.310 & 0.276 \\ -1.397 & -1.591 & 1.2 & 0.2 \end{array} \right\}. \tag{6.100}$$

The support points of the design (6.100) are represented by symbols in Figure 6.28 (a) and the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ is represented in Figure 6.27 (b). Clearly, Figure 6.28 (a) illustrates that the design (6.100) has the design pattern of Figure 6.24 (d) and Figure 6.28 (b) suggests that the design (6.100) is D-optimal.

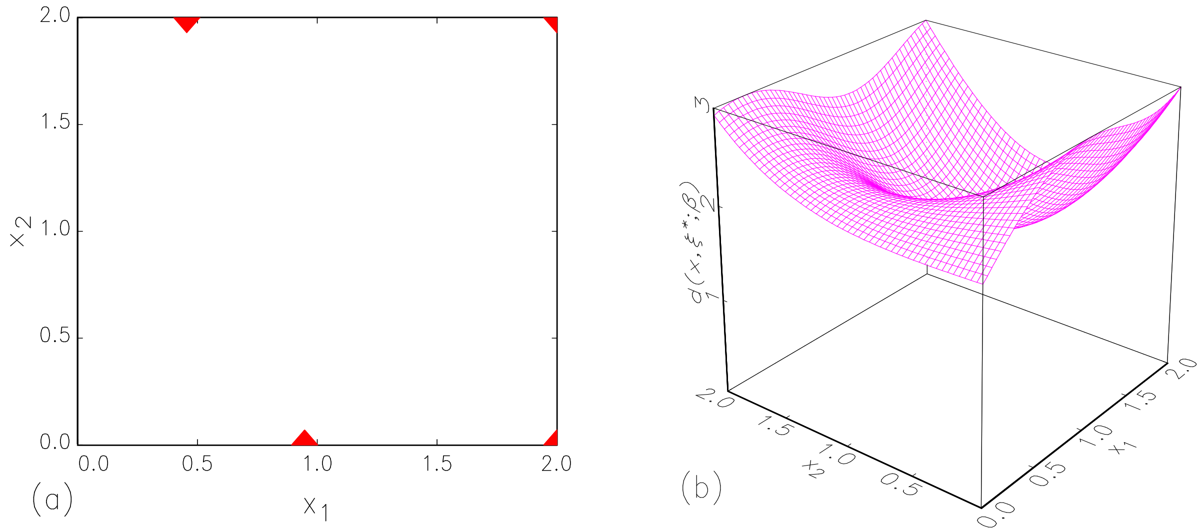


Figure 6.28: (a) Support points and (b) the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ of the D-optimal design for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, 2] \times [0, 2]$ and with parameter vector $\boldsymbol{\beta} = (-3, 1.6, 0.5)^T$.

6.9. Practical example

Consider the data analyzed by Martin (1942) and by Greco and Lawrence (1988) presented in Table A.1. The purpose of the experiment which generated the data was to examine the lethal effect of two insecticides, rotenone and deguelin, on chrysanthemum aphids. The explanatory variables, d_1 and d_2, were concentrations in milligrams per littre, mg/l, of the two insecticides. The design setting appears to be 16-point ray design on the design space $[0, 10.2] \times [0, 50.5]$ with 5 support point points on each of the rays $d_1 = 0$ and $d_2 = 0$, and 6 support points on the ray $d_2 \simeq 4d_1$. Greco and Lawrence (1988) fitted the binary response model (2.29) in Section 2.4.2. The results for the present data were that the interaction parameter was not significantly different from zero.

The aim of introducing this example here is first to show that the two-variable binary logistic model without interaction, discussed in this chapter, provides similar results to those reported by Greco and Lawrence (1988). Secondly, and more importantly, the example is chosen as a real life application of the trapezium design introduced in this thesis, and in order to show that the 4-point trapezium design is more efficient than the 16-point design used to generate the data. It has been found in Section 4.4.3, Example 4.1, that fitting the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 d_1 + \beta_2 d_2 + \beta_{12} d_1 d_2$ to the present data set gives the estimates $\widehat{\beta}_{12} = 0.0074$ (p-value $= 0.2237$). Since the interaction parameter is not significantly different from zero, fitting the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 d_1 + \beta_2 d_2$ gives parameter estimates $\widehat{\beta}_0 = -2.0297$, $\widehat{\beta}_1 = 0.4276$ and $\widehat{\beta}_2 = 0.1554$ all significantly different from zero. Transforming the rectangular design space $[0, 10.2]\times[0, 50.5]$ to the square design space $[0, 2]\times[0, 2]$ using the transformation (6.4) gives $\beta_{0new} = -2.0297$, $\beta_{1new} = 2.1808$ and $\beta_{2new} = 3.9239$. For this example

$$1.5434 < \min\{-\beta_{0new} = 2.0297, \beta_{0new} + \beta_{1new}b = 2.3319, \beta_{0new} + \beta_{2new}b = 5.8181\} = 2.0297.$$

Hence, condition (6.5) holds and thus the candidate D-optimal design is conjectured to be a 4-point trapezium design. The design for this example is therefore

$$\xi^* = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{cccc} (0.257, 0) & (0, 0.143) & (0, 0.892) & (1.605, 0) \\ 0.1688 & 0.1688 & 0.3312 & 0.3312 \\ -1.470 & -1.470 & 1.470 & 1.470 \end{array} \right\} \tag{6.101}$$

where $x_1 = 5.1d_1$ and $x_2 = 25.25d_2$ following transformations (6.3). The support points of the design (6.101) are represented by symbols in Figure 6.29 (a) and are vertices of a trapezium. The weights associated with points located on the same logit are equal as expected for a 4-point trapezium design. For all $\boldsymbol{x} = (x_1, x_2)^T$ with $(x_1, x_2) \in [0, 2] \times [0, 2]$, the standardized

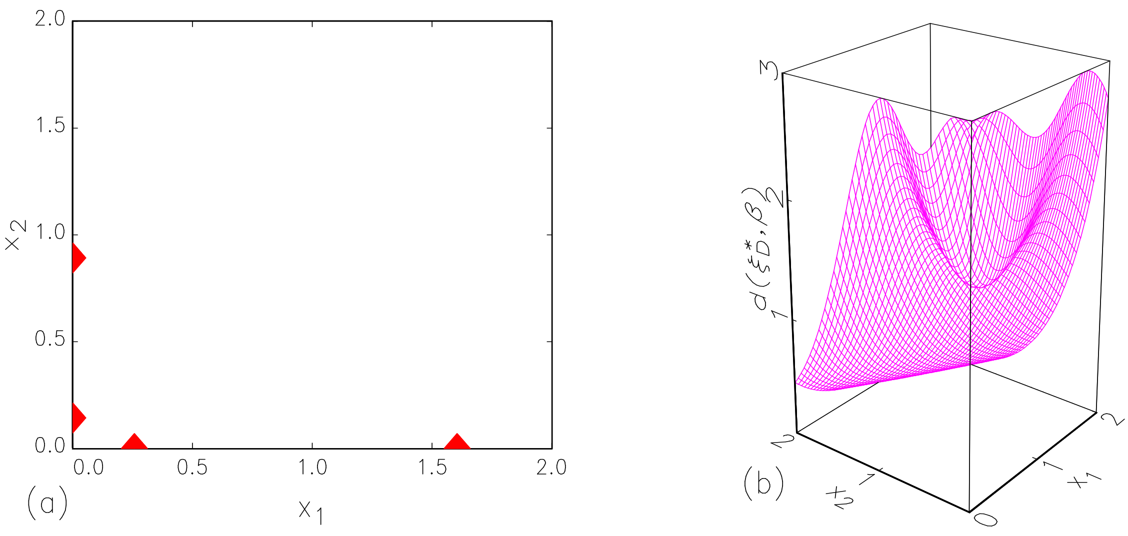


Figure 6.29: (a) Support points of the D-optimal design and (b) the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, 2] \times [0, 2]$ and with parameter vector $\boldsymbol{\beta} = (-2.0297, 2.1808, 3.9239)^T$ for Martin (1942) data.

variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ plotted in Figure 6.29 (b) suggests that the design is D-optimal since $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) \leq 3$ with equality at the support points of the design (6.101).

The D-optimal design for the original data can be obtained using the transformations (6.3) and is given by

$$\xi^* = \left\{ \begin{array}{c} (d_1, d_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{cccc} (1.311, 0) & (0, 3.611) & (0, 22.523) & (8.1855, 0) \\ 0.1688 & 0.1688 & 0.3312 & 0.3312 \\ -1.470 & -1.470 & 1.470 & 1.470 \end{array} \right\}. \qquad (6.102)$$

Now, consider calculating the efficiency of the 16-point design, ξ, used to generate the data of Table A.1 relative to the D-optimal design ξ^* given by (6.102). The determinant of the information for the parameter vector $\boldsymbol{\beta} = (-2.0297, 0.4276, 0.1554)^T$ evaluated at design ξ is $|M(\xi; \boldsymbol{\beta})| = 1.071$, while the determinant of the information matrix of the same parameter vector at design (6.102) is $|M(\xi^*; \boldsymbol{\beta})| = 3.106$. Hence, the D-efficiency (3.38) for the original design ξ relative to the D-optimal design (6.102) is

$$D_{\text{eff}} = \left\{ \frac{|M(\xi; \boldsymbol{\beta})|}{|M(\xi^*; \boldsymbol{\beta})|} \right\}^{1/3} = 0.701.$$

Thus, since $D_{\text{eff}} = 0.701 < 1$, then the 4-point D-optimal trapezium design (6.102) is more efficient than the 16-point ray design used to generate the data of Table A.1.

6.10. Conclusions

In this chapter, the construction of the D-optimal designs for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[a, b] \times [c, d]$, transformed to $[0, b] \times [0, b]$ for ease of calculations, was discussed. The candidate D-optimal were investigated semi-analytically and semi-numerically. Conditions for the existence of the design patterns of the candidate D-optimal designs were derived and the optimal designs were constructed numerically, and the optimality of the designs checked graphically. The number of support points of the D-optimal designs varied from 3 to 6. In cases when $\beta_0 + \beta_1 b > 0$ and $\beta_0 + \beta_2 b > 0$ and the logit line $u = 1.5434$ lies below the diagonal line joining points $(0, b)$ and $(b, 0)$, the D-optimal designs were found to be same as the 4-point trapezium and the 3-point designs discussed in Chapter 5. In addition cases when $\beta_0 + \beta_1 b < 0$ and $\beta_0 + \beta_2 b < 0$ and the logit line $u = -1.5434$ lies above the diagonal line joining points $(0, b)$ and $(b, 0)$, the D-optimal designs were also found to be 4-point trapezium D-optimal designs and the 3-point D-optimal designs. The term trapezium is taken in the sense that parallel sides are logit lines.

Non-trapezium D-optimal designs were found in cases when $\beta_0 + \beta_1 b$ and $\beta_0 + \beta_2 b$ have the same sign, but the logit line $u = 1.5434$ passes above at least one of the points $(0, b)$ and $(b, 0)$. In cases when the signs of $\beta_0 + \beta_1 b$ and $\beta_0 + \beta_2 b$ alternate but with $\beta_1 \neq \beta_2$ and the logit line $u = 0$ passes through the origin of the design space $[0, b] \times [0, b]$, i.e. $\beta_0 = -\dfrac{b(\beta_1 + \beta_2)}{2}$, the candidate D-optimal designs were found to be same as the 4-point parallelogram designs of Sitter and Torsney (1995a), and Jia and Myers (2001), and some cases of Atkinson and Haines (1996) discussed in 4 and in Section 5.2. D-optimal parallelogram designs for model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design space $[0, b] \times [0, b]$ were also found when $\beta_0 = -b\beta_1$ and $\beta_1 = \beta_2 \leq \dfrac{2.1150}{b}$. Candidate D-optimal designs with 5 or 6 support points were found in cases when $\beta_1 = \beta_2 > \dfrac{2.1150}{b}$ and $\beta_0 = -b\beta_1$. Cases of 4-point non-parallelogram designs were found when the signs of $\beta_0 + \beta_1 b$ and $\beta_0 + \beta_2 b$ alternate, and $\beta_1 \neq \beta_2$, but with $\beta_0 \neq -\frac{b(\beta_1+\beta_2)}{2}$. A practical example on the joint effect of two insecticides was used to show that optimal designs use less experimental runs but are more efficient than some well known experimental designs.

7

D-optimal Designs for the Two-Variable Binary Logistic Model with Interaction

7.1. Introduction

Drugs, insecticides, or other compounds may exhibit a synergistic effect when administered in combination. Some models for detecting an interaction between two drugs were reviewed in Chapter 2. This chapter will deal with the construction of D-optimal designs for the two-variable binary logistic model with interaction

$$u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 \tag{7.1}$$

discussed in Section 2.4.4. In the model (7.1), $x_1 \in \mathbb{R}$ and $x_2 \in \mathbb{R}$ can be taken as dose or log-dose concentrations of two drugs. The theory of statistical tests for drug interaction is well developed. For example, Hosmer and Lemeshow (2000) and Collett (2003) describe the applications to data that is described by the binary logistic regression model without and with interaction. However, very little work exists in literature on the construction of optimal designs for detecting drug interaction. D-optimal designs for the precise estimation of the parameters of the two-variable binary logistic model with interaction (7.1) was discussed by Jia and Myers (2001). The class of the designs discussed by Jia and Myers (2001) is narrow since the designs were restricted to equally weighted 4-point designs with support points being the vertices of a parallelogram whose two opposite sides lie on the logit lines $u = \pm 1.5434$ (see Section 4.6.2). In addition to these restrictions on the class of the designs, Jia and Myers (2001) did not prove the D-optimality of their 4-point design. These two issues are addressed in this chapter by further investigating the candidate D-optimal design patterns introduced in Jia and Myers (2001). In particular, the D-optimal designs do not necessarily need to have 4 support points, and the support points of the D-optimal designs do not necessarily have to be equally weighted. Kupchak (2000), as reviewed in Section 4.6.1, also discussed the construction of D-optimal designs for the two-variable binary logistic model with interaction (7.1), but his interest was mainly in constructing designs for the precise estimation of the interaction parameter β_{12} alone, and for the precise estimation of all the parameters assuming that the interaction $\beta_{12} = 0$. In this chapter, analytical and numerical constructions of D-optimal designs for the precise estimation of all the four parameters in model (7.1) are undertaken. The approach used to construct candidate D-optimal design patterns is that of Jia and Myers (2001) with some modifications and extensions.

This chapter contains the following sections. Section 7.2 gives a proof of the D-optimality of the design of Jia and Myers (2001) for the two-variable binary logistic model with interaction. Section 7.3 discusses a variant of the design approach of Jia and Myers (2001) on the design space $(-\infty, \infty) \times [c, d]$ where c and d are real numbers. Section 7.4 discusses the construction of the D-optimal designs for synergy using the two-variable binary logistic model (7.1) with parameters satisfying the conditions $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_{12} > 0$. Section 7.5 discusses the construction of the D-optimal designs for antagonism using model (7.1) with parameters satisfying the conditions $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_{12} < 0$. Section 7.6 briefly discusses the construction of D-optimal designs for synergy and antagonism for model (7.1) with respectively $\beta_0 > 0$, $\beta_1 < 0$, $\beta_2 < 0$ and $\beta_{12} < 0$, and $\beta_0 > 0$, $\beta_1 < 0$, $\beta_2 < 0$ and $\beta_{12} > 0$. Two practical examples of constructing D-optimal designs for synergy and antagonism are given in Section 7.7 and Section 7.8 concludes the chapter. Throughout the chapter, D-optimal designs for simple cases are constructed analytically using the Equivalence Theorem for D-optimality. In complex cases the patterns of the candidate D-optimal designs are conjectured under specific conditions, and then the D-optimal designs are constructed numerically using the Gauss program in Appendix C. The D-optimality of the designs is checked graphically by plotting the standardized variance function over the design space.

7.2. D-optimality of the design of Jia and Myers

In Section 4.4.2, the equation $u = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ given in model (7.1), with $(x_1, x_2) \in \mathbb{R}^2$, was described by Jia and Myers (2001) as an hyperbola with centre $(x_{01}, x_{02}) = \left(\frac{-\beta_2}{\beta_{12}}, \frac{-\beta_1}{\beta_{12}} \right)$ for a fixed value of u. Then model (7.1) was re-parameterized to model

$$u = \text{logit}(p) = \beta_0^* + \beta_{12} z_1 z_2 \tag{7.2}$$

where $\beta_0^* = \beta_0 - \dfrac{\beta_1\beta_2}{\beta_{12}}$ is the value of the logit u given in (7.1) with $(x_1, x_2) = (x_{01}, x_{02})$, and $(z_1, z_2) = \left(x_1 + \dfrac{\beta_2}{\beta_{12}}, x_2 + \dfrac{\beta_1}{\beta_{12}}\right) \in \mathbb{R}^2$ is the transformed value of an arbitrary (x_1, x_2) by translation which moves the centre $\left(\dfrac{-\beta_2}{\beta_{12}}, \dfrac{-\beta_1}{\beta_{12}}\right)$ to $(0, 0)$ (see Figures 4.7 to 4.9). Jia and Myers (2001) suggested for model (7.2) a 4-point candidate D-optimal design ξ with support points located on each of the two branches of the hyperbolae $u_1 = \beta_0^* + \beta_{12}z_1z_2$ and $u_2 = \beta_0^* + \beta_{12}z_1z_2$ with $u_1 < u_2$. The abscissa of the support points located on the two branches of the hyperbola u_1 were chosen as $z_{11} = \Delta_1$ and $z_{31} = -\Delta_1$, and those on the hyperbola u_2 as $z_{21} = t\Delta_1$ and $z_{41} = -t\Delta_1$ where $\Delta_1 > 0$ and $t \in (0, \infty)$. In addition, Jia and Myers (2001) assumed that the support points located on the same hyperbola, or logit line, have equal weights. Hence, the candidate D-optimal design proposed by Jia and Myers (2001) for model (7.2) on the (z_1, z_2) design space $\mathbb{R}^2$ is of the form

$$\xi = \left\{ \begin{array}{cccc} \left(\Delta_1, \dfrac{u_1 - \beta_0^*}{\beta_{12}\Delta_1}\right) & \left(t\Delta_1, \dfrac{u_2 - \beta_0^*}{t\beta_{12}\Delta_1}\right) & \left(-\Delta_1, -\dfrac{u_1 - \beta_0^*}{\beta_{12}\Delta_1}\right) & \left(-t\Delta_1, -\dfrac{u_2 - \beta_0^*}{t\beta_{12}\Delta_1}\right) \\ \frac{\lambda}{2} & \frac{1-\lambda}{2} & \frac{\lambda}{2} & \frac{1-\lambda}{2} \end{array} \right\}. \tag{7.3}$$

Jia and Myers (2001) set $t = 1$ so that the support points of the candidate D-optimal design (7.3) are vertices of a parallelogram and obtained an equally weighted 4-point design with support points on the two logit lines $u^* = \pm 1.5434$. Hence, the design (7.3) of Jia and Myers (2001) can be written as

$$\xi^* = \left\{ \begin{array}{cccc} \left(\Delta_1, \dfrac{-u^* - \beta_0^*}{\beta_{12}\Delta_1}\right) & \left(\Delta_1, \dfrac{u^* - \beta_0^*}{\beta_{12}\Delta_1}\right) & \left(-\Delta_1, \dfrac{u^* + \beta_0^*}{\beta_{12}\Delta_1}\right) & \left(-\Delta_1, \dfrac{-u^* + \beta_0^*}{\beta_{12}\Delta_1}\right) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array} \right\}$$

where $\Delta_1 > 0$ and $u^* = 1.5434$. Jia and Myers (2001) did not prove the D-optimality of the above 4-point design. However, a proof for D-optimality of the design ξ^* is easy by noting that model (7.2) is explicitly defined by only two parameters β_0^* and β_{12}. The following proposition is on the D-optimality of the design ξ^*.

Proposition 7.1. The design

$$\xi^* = \left\{ \begin{array}{cccc} \left(\Delta_1, \dfrac{-u^* - \beta_0^*}{\beta_{12}\Delta_1}\right) & \left(\Delta_1, \dfrac{u^* - \beta_0^*}{\beta_{12}\Delta_1}\right) & \left(-\Delta_1, \dfrac{u^* + \beta_0^*}{\beta_{12}\Delta_1}\right) & \left(-\Delta_1, \dfrac{-u^* + \beta_0^*}{\beta_{12}\Delta_1}\right) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array} \right\}, \tag{7.4}$$

for the two-variable binary logistic model $u = \text{logit}(p) = \beta_0^* + \beta_{12}z_1z_2$ on the design space $\mathbb{R}^2$, where $\Delta_1 > 0$ and $u^* = 1.5434$, is D-optimal.

Proof

The information matrix for $\boldsymbol{\beta} = (\beta_0^*, \beta_{12})^T$, in model (7.2), evaluated at a support point $\boldsymbol{z} = (z_1, z_2)^T$ is

$$M(\boldsymbol{z};\boldsymbol{\beta}) = \frac{e^u}{(1+e^u)^2}\begin{bmatrix} 1 \\ z_1 z_2 \end{bmatrix}[1 \quad z_1 z_2] = \frac{e^u}{(1+e^u)^2}\begin{bmatrix} 1 & z_1 z_2 \\ z_1 z_2 & z_1^2 z_2^2 \end{bmatrix}. \tag{7.5}$$

Furthermore, the information matrix for $\boldsymbol{\beta} = (\beta_0^*, \beta_{12})^T$ evaluated at the 4-point design (7.4) is

$$M(\xi^*;\boldsymbol{\beta}) = \frac{1}{4}\sum_{i=1}^{4} M(\boldsymbol{z}_i;\boldsymbol{\beta}) = \frac{e^{u^*}}{(1+e^{u^*})^2}\begin{bmatrix} 1 & -\frac{\beta_0^*}{\beta_{12}} \\ -\frac{\beta_0^*}{\beta_{12}} & \frac{(\beta_0^{*2}+u^{*2})}{\beta_{12}^2} \end{bmatrix}$$

where $u^* = 1.5434$. The D-optimality of design (7.4) is proved by showing that the standardized variance function $d(\boldsymbol{z}, \xi^*;\boldsymbol{\beta}) = \operatorname{tr}\{M^{-1}(\xi^*;\boldsymbol{\beta})M(\boldsymbol{z};\boldsymbol{\beta})\} \leq 2$ for all $\boldsymbol{z} = (z_1, z_2)^T \in \mathbb{R}^2$ with the equality holding at the support points of design ξ^*. Let $u = \beta_0^* + \beta_{12} z_1 z_2$ where $(z_1, z_2) \in \mathbb{R}^2$. Simple calculations combined with the results of Theorem 4.1 lead to

$$\begin{aligned} d(\boldsymbol{z}, \xi^*;\boldsymbol{\beta}) &= \frac{e^{u-u^*}(1+e^{u^*})^2(u^{*2} + \beta_0^{*2} + 2\beta_0^*\beta_{12}z_1z_2 + \beta_{12}^2 z_1^2 z_2^2)}{u^{*2}(1+e^u)^2} \\ &= \frac{e^{u-u^*}(1+e^{u^*})^2(u^{*2}+u^2)}{u^{*2}(1+e^u)^2} \leq 2 \end{aligned}$$

for all $u \in \mathbb{R}$ with the equality holding at $u = u^*$. ■

7.3. An alternative to the design of Jia and Myers

7.3.1 Equally weighted 4-point D-optimal designs on the design space $\mathbb{R} \times [c, d]$

In Section 4.6.2, it was argued that setting $t = 1$ leads to a narrow class of designs, particularly designs with support points which are vertices of a parallelogram. In addition, guessing the value of Δ_1 in design (7.4) further narrows the class of D-optimal designs. This section discusses an alternative approach for constructing a 4-point D-optimal design without an explicit knowledge of Δ_1 and without forcing the candidate 4-point D-optimal designs to have support points at the vertices of a parallelogram. In this regard, consider the two-variable binary logistic models (7.1) and (7.2). As in the case of the construction of D-optimal designs for the two-variable binary logistic model without interaction discussed in Chapters 5 and 6, a restriction on the design space $\mathbb{R}^2$ is needed for constructing D-optimal designs for the two-variable binary logistic model with interaction (7.1). To obtain a restriction on the design space

$\mathbb{R}^2$ for model (7.1), assume that $x_1 \in \mathbb{R}$ and $x_2 \in [c, d]$ where c and d are fixed real numbers. The restriction $c \leq x_2 \leq d$ is equivalent to $c^* \leq z_2 \leq d^*$ in model (7.2), where $c^* = c + \dfrac{\beta_1}{\beta_{12}}$ and $d^* = d + \dfrac{\beta_1}{\beta_{12}}$. Since $x_1 \in \mathbb{R}$, it follows from (7.1) that the logit u is unrestricted and hence u can theoretically be any real number. Also, as in Jia and Myers (2001), assume that the support points of the D-optimal design are located on two logits u_1 and u_2 with $u_1 < u_2$. As in Section 4.6.2, the following three design cases can be considered.

1. $u_1 < \beta_0^* < u_2$. Figure 7.1 (a) displays the design pattern for this case when $\beta_{12} < 0$, and Figure 7.1 (b) displays the design pattern for the case when $\beta_{12} > 0$. The support points of the candidate 4-point D-optimal design are A, B, C and D which are spread in the 4 quadrants of the (z_1, z_2) design space with $z_1 \in \mathbb{R}$ and $z_2 \in [c^*, d^*]$.

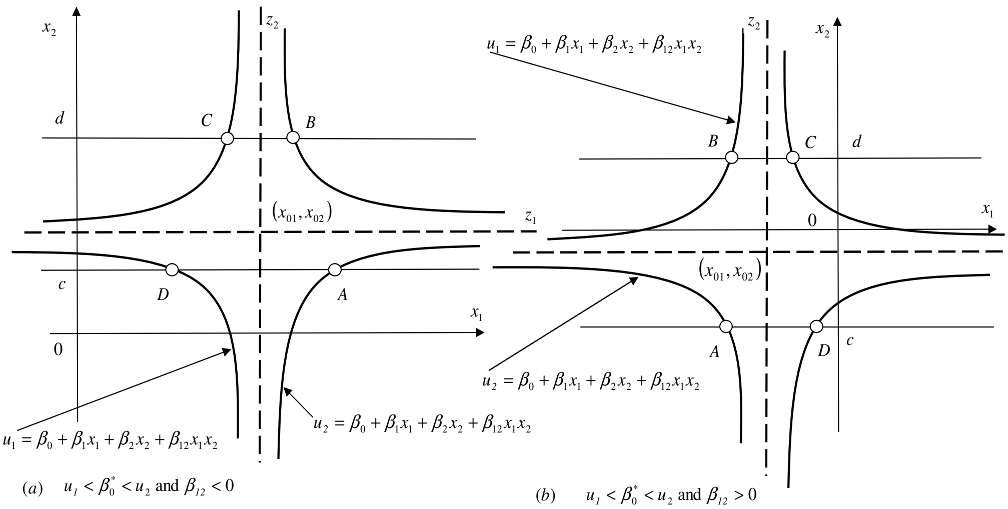


Figure 7.1: Design patterns of the candidate D-optimal design for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the (x_1, x_2) design space $\mathbb{R} \times [c, d]$ and for the case of $u_1 < \beta_0^* < u_2$: (a) $\beta_{12} < 0$; (b) $\beta_{12} > 0$. Circles are support points.

2. $\beta_0^* < u_1 < u_2$. Figures 7.2 (a) and (b) display the respective design patterns for this case when $\beta_{12} < 0$ and $\beta_{12} > 0$. The support points of the 4-point candidate D-optimal design are A, B, C and D which are in the following two quadrants of the (z_1, z_2) design space: (a) second and fourth quadrants when $\beta_{12} < 0$; and (b) first and third quadrants when $\beta_{12} > 0$.

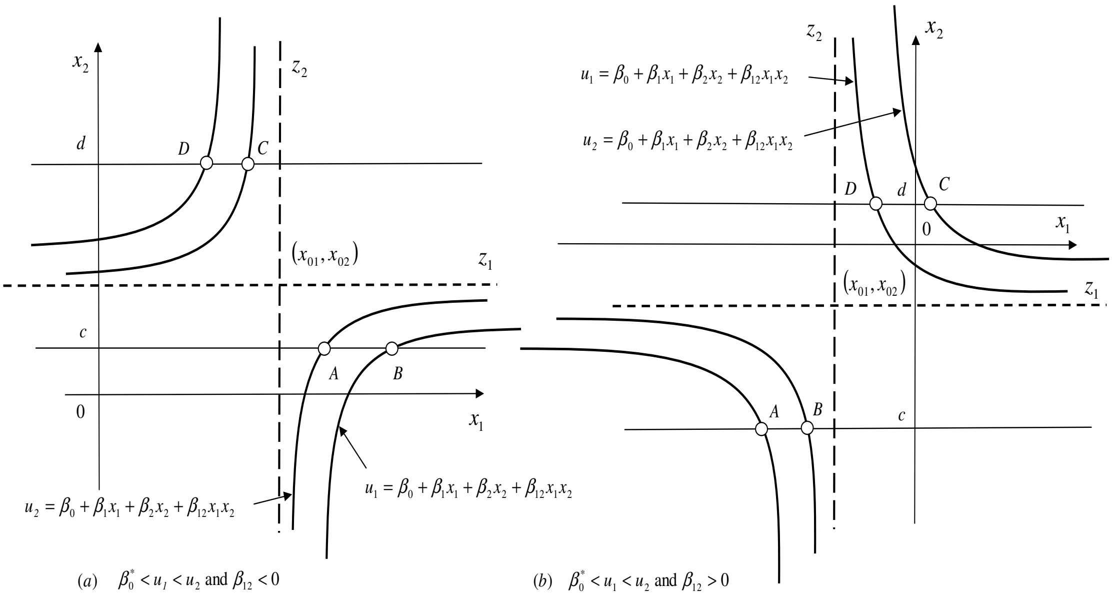


Figure 7.2: Design patterns of the candidate D-optimal design for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the (x_1, x_2) design space $\mathbb{R} \times [c, d]$ and for the case of $\beta_0^* < u_1 < u_2$: (a) $\beta_{12} < 0$; and (b) $\beta_{12} > 0$. Circles are support points.

3. $u_1 < u_2 < \beta_0^*$. Figures 7.3 (a) and (b) display the respective for this case when $\beta_{12} < 0$ and $\beta_{12} > 0$. The support points of the 4-point candidate D-optimal design are A, B, C and D which are in the following two quadrants of the (z_1, z_2) design space: (a) first and third quadrants when $\beta_{12} < 0$; and (b) second and fourth quadrants when $\beta_{12} > 0$.

Proposition 7.2 formalizes the fact that each of the 4-point designs with design patterns displayed in Figures 7.1 to 7.3 is D-optimal.

Proposition 7.2. The design

$$\xi^* = \left\{ \begin{array}{cccc} \left(\frac{u^*-\beta_0-\beta_2 c}{\beta_1+\beta_{12} c}, c\right) & \left(\frac{-u^*-\beta_0-\beta_2 c}{\beta_1+\beta_{12} c}, c\right) & \left(\frac{u^*-\beta_0-\beta_2 d}{\beta_1+\beta_{12} d}, d\right) & \left(\frac{-u^*-\beta_0-\beta_2 d}{\beta_1+\beta_{12} d}, d\right) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array} \right\}, \tag{7.6}$$

where $u^* = 1.5434$, for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $\mathbb{R} \times [c, d]$ where c and d are constants, is D-optimal.

Proof

Consider Figures 7.1 to 7.3. In each of the figures, the support points of a candidate 4-point D-optimal design ξ for model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design

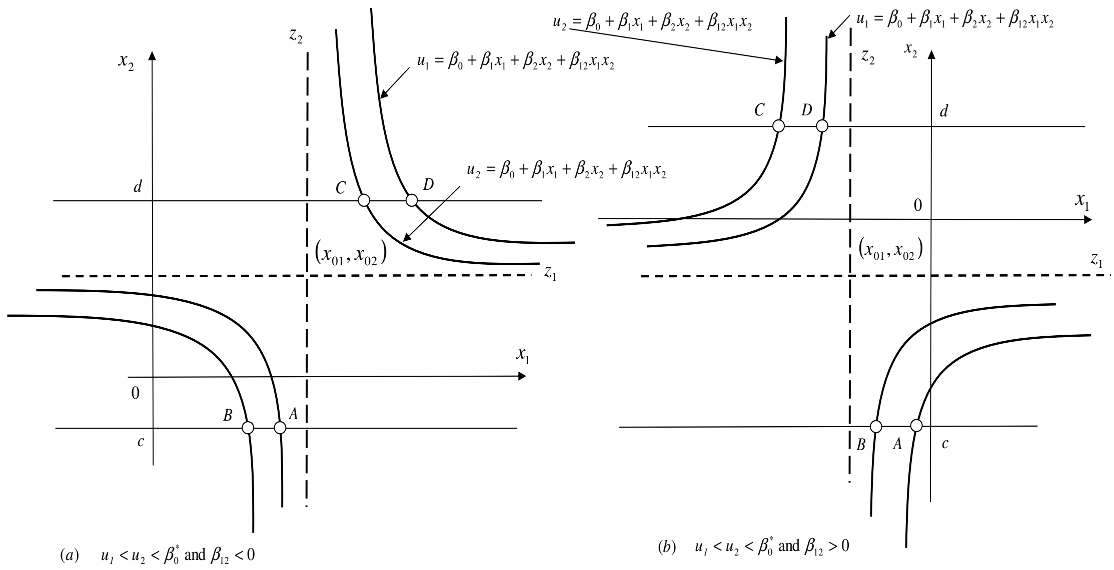


Figure 7.3: Design patterns of the candidate D-optimal design for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the (x_1, x_2) design space $\mathbb{R} \times [c, d]$ and for the case of $u_1 < u_2 < \beta_0^*$: (a) $\beta_{12} < 0$; and (b) $\beta_{12} > 0$. Circles are support points.

space $\mathbb{R} \times [c, d]$ are $A\left(\dfrac{u_2 - \beta_0 - \beta_2 c}{\beta_1 + \beta_{12} c}, c\right)$, $B\left(\dfrac{u_1 - \beta_0 - \beta_2 c}{\beta_1 + \beta_{12} c}, c\right)$, $C\left(\dfrac{u_2 - \beta_0 - \beta_2 d}{\beta_1 + \beta_{12} d}, d\right)$ and $D\left(\dfrac{u_1 - \beta_0 - \beta_2 d}{\beta_1 + \beta_{12} d}, d\right)$. As model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ has 4 parameters, then the weights associated with the 4 support points of ξ are equal. Then the candidate 4-point D-optimal design for model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the (x_1, x_2) design space $\mathbb{R} \times [c, d]$ is of the form

$$\xi = \left\{ \begin{array}{cccc} \left(\frac{u_2 - \beta_0 - \beta_2 c}{\beta_1 + \beta_{12} c}, c\right) & \left(\frac{u_1 - \beta_0 - \beta_2 c}{\beta_1 + \beta_{12} c}, c\right) & \left(\frac{u_2 - \beta_0 - \beta_2 d}{\beta_1 + \beta_{12} d}, d\right) & \left(\frac{u_1 - \beta_0 - \beta_2 d}{\beta_1 + \beta_{12} d}, d\right) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array} \right\}. \tag{7.7}$$

Let $M(\xi; \boldsymbol{\beta})$ be the information matrix for the parameters of model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ evaluated at the design ξ given by (7.7). The determinant of $M(\xi; \boldsymbol{\beta})$ is

$$D_1 = \frac{(d-c)^4 e^{2(u_1+u_2)} (u_1 - u_2)^4}{256(\beta_1 + \beta_{12} c)^2 (\beta_1 + \beta_{12} d)^2 (1 + e^{u_1})^4 (1 + e^{u_2})^4}$$

which is proportional to

$$D = \frac{e^{2(u_1+u_2)} (u_1 - u_2)^4}{(1 + e^{u_1})^4 (1 + e^{u_2})^4} \tag{7.8}$$

which, in turn, is proportional to the determinant (4.47) in Section 4.6.2. Therefore, D and thus D_1 is maximum at $u_2 = -u_1 = u^* = 1.5434$. Consequently, the candidate D-optimal design of the form (7.7) is

$$\xi^* = \left\{ \begin{array}{cccc} \left(\frac{u^*-\beta_0-\beta_2 c}{\beta_1+\beta_{12}c}, c\right) & \left(\frac{-u^*-\beta_0-\beta_2 c}{\beta_1+\beta_{12}c}, c\right) & \left(\frac{u^*-\beta_0-\beta_2 d}{\beta_1+\beta_{12}d}, d\right) & \left(\frac{-u^*-\beta_0-\beta_2 d}{\beta_1+\beta_{12}d}, d\right) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array} \right\}, \tag{7.9}$$

where $u^* = 1.5434$. Clearly, the candidate D-optimal design ξ^* given in (7.9) is the same as the design (7.6). It remains to prove D-optimality of the design (7.9) by showing that the standardized variance function is less than or equal to the number of parameters 4 with equality holding at the support points of ξ^*. The standardized variance function associated with the design (7.9) is

$$d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) = \frac{2e^{u-u^*}(1+e^{u^*})^2 \left\{\left[u^{*2} + (s_0 + s_1 x_1)^2\right](x_2 - c)^2 + \left[u^{*2} + (q_0 + q_1 x_1)^2\right](x_2 - d)^2\right\}}{(d-c)^2 u^{*2}(1+e^u)^2} \tag{7.10}$$

where $s_0 = \beta_0 + \beta_2 c$, $s_1 = \beta_1 + \beta_{12} c$, $q_0 = \beta_0 + \beta_2 d$, $q_1 = \beta_1 + \beta_{12} d$ and $\boldsymbol{x} = (x_1, x_2)^T \in \mathbb{R} \times [c, d]$. Note that as $x_1 \in \mathbb{R}$, then $s_0 + s_1 x_1 = \beta_0 + \beta_2 c + \beta_1 x_1 + \beta_{12} x_1 c$ is the branch intersecting the line $x_2 = c$ of an arbitrary hyperbola $u = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$. Similarly, $q_0 + q_1 x_1 = \beta_0 + \beta_2 d + \beta_1 x_1 + \beta_{12} d x_1$ is the branch intersecting the line $x_2 = d$ of the arbitrary hyperbola $u = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$. Thus, the standardized variance function (7.10) can be simplified to

$$d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) = \frac{2e^{u-u^*}(1+e^{u^*})^2 (u^{*2} + u^2)\left[(x_2 - c)^2 + (x_2 - d)^2\right]}{(d-c)^2 u^{*2}(1+e^u)^2} \tag{7.11}$$

Clearly, $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) = 4$ at all the 4 support points of the design (7.9) at which $u = \pm u^*$ and $x_2 = c$ or $x_2 = d$. It remains to show that $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) \leq 4$ for all $(x_1, x_2)^T \in \mathbb{R} \times [c, d]$. In (7.11) the quadratic polynomial $f(x_2) = (x_2 - c)^2 + (x_2 - d)^2 = 2x_2^2 - 2(c+d)x_2 + c^2 + d^2$ has the global minimum $\dfrac{(d-c)^2}{2}$ attained at $x_2 = \dfrac{c+d}{2}$. Then, since $c \leq x_2 \leq d$, then $f(x_2)$ has the global maximum $(d-c)^2$ attained at $x_2 = c$ or $x_2 = d$. Hence,

$$d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) \leq g(u) = \frac{2e^{u-u^*}(1+e^{u^*})^2(u^2 + u^{*2})}{u^{*2}(1+e^u)^2} \leq 4$$

since $\dfrac{e^{u-u^*}(1+e^{u^*})^2(u^2 + u^{*2})}{u^{*2}(1+e^u)^2} \leq 2$ (see Theorem 4.1 in Section 4.3). Thus, the design (7.6) is D-optimal. ■

7.3.2 D-optimal designs on the design space $[0, b] \times [0, d]$

In practical situations the design space is bounded. In the case of two drugs, say A and B, the doses or log-doses x_1 and x_2 can be from a rectangular design space of the form $[a, b] \times [c, d]$ where a, b, c and d are real numbers. Because of the restrictions $a \leq x_1 \leq b$ and $c \leq x_2 \leq d$ on both x_1 and x_2, the number of points of intersection of the curves of the hyperbolae u_1 and u_2 in each of the Figures 7.1 to 7.3 and the design space boundaries is no longer necessarily 4, but can vary from 4 to 8. Hence, various design patterns with the number of support points varying from 4 to 8 can result depending on the values of the model parameters and the design space $[a, b] \times [c, d]$. In such cases, it can be conjectured that the optimal positive logit is not necessary $u^* = 1.5434$.

As in Section 6.2 of Chapter 6, suppose that the original design space $[a, b] \times [c, d]$ is transformed to a new design space $[a_1, b_1] \times [c_1, d_1]$ for the sake of ease of constructing the designs on the new space. Then as in (6.3), the support points on $[a_1, b_1] \times [c_1, d_1]$ are (x_{1new}, x_{2new}) given by

$$\begin{cases} x_{1new} = \frac{(b_1-a_1)x_1-(ab_1-a_1b)}{b-a} \\ x_{2new} = \frac{(d_1-c_1)x_2-(cd_1-c_1d)}{d-c}. \end{cases} \tag{7.12}$$

However, because of the additional interaction term $\beta_{12}x_1x_2$ in model (7.1), the new parameters are no longer given by (6.4), but are given by

$$\begin{cases} \beta_{1new} = \frac{(b-a)\beta_1}{b_1-a_1} + \frac{(b-a)(cd_1-c_1d)\beta_{12}}{(b_1-a_1)(d_1-c_1)} \\ \beta_{2new} = \frac{(d-c)\beta_2}{d_1-c_1} + \frac{(d-c)(ab_1-a_1b)\beta_{12}}{(d_1-c_1)(b_1-a_1)} \\ \beta_{12new} = \frac{(b-a)(d-c)\beta_{12}}{(b_1-a_1)(d_1-c_1)} \\ \beta_{0new} = \beta_0 + \frac{(ab_1-a_1b)\beta_1}{b_1-a_1} + \frac{(cd_1-c_1d)\beta_2}{d_1-c_1} + \frac{(ab_1-a_1b)(cd_1-c_1d)\beta_{12}}{(b_1-a_1)(d_1-c_1)}. \end{cases} \tag{7.13}$$

In order to reduce complications in the notations and calculations, the design space that will be used throughout this chapter is the square $[0, b] \times [0, b]$ obtained by setting $a = c = 0$ and $d = b$ in the design space $[a, b] \times [c, d]$. If required, the transformations (7.12) and (7.13) will be used to transform the designs from the (x_1, x_2) design space $[0, b] \times [0, b]$ to the (x_{1new}, x_{2new}) design space $[a_1, b_1] \times [c_1, d_1]$.

As discussed in Chapter 2, if $\beta_0 < 0$, $\beta_1 > 0$ and $\beta_2 > 0$, then there is synergy between two drugs if $\beta_{12} > 0$, no interaction between the drugs if $\beta_{12} = 0$, and antagonism between the drugs if $\beta_{12} < 0$. The construction of D-optimal designs on the square $[0, b] \times [0, b]$ will first be discussed assuming that the two drugs interact synergically, and then the results will be extended to some cases of antagonistic interaction. Other cases of antagonism will be discussed independently when deduction from synergy is difficult.

7.4. D-optimal designs for synergy

Consider the two-variable binary logistic model with interaction (7.1) on the design space $[0,b] \times [0,b]$. It has to be noted that under the assumptions that $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_{12} > 0$, $\beta_0^* = \beta_0 - \frac{\beta_1\beta_2}{\beta_{12}} < \beta_0 < 0$. Furthermore, in the case of synergy ($\beta_{12} > 0$) with logit u fixed, $\beta_1 > 0$ and $\beta_2 > 0$, the centre $\left(-\dfrac{\beta_2}{\beta_{12}}, -\dfrac{\beta_1}{\beta_{12}}\right)$ of the hyperbola given by (7.1) is outside the design space $[0,b] \times [0,b]$ since the coordinates of the centre are all negative. Because the branches of any hyperbola are symmetric with respect to its centre, the pattern of the candidate D-optimal design will depend on the location of the 4 branches of the hyperbolae $\pm u = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ relative to the design space $[0,b] \times [0,b]$. As discussed in Section 4.6.2 and Section 7.3, these branches may be spread in the four quadrants of $\mathbb{R}^2$ if $|\beta_0^*| < 1.5434$ or in just two quadrants if $|\beta_0^*| > 1.5434$, where $\beta_0^* = \beta_0 - \dfrac{\beta_1\beta_2}{\beta_{12}}$.

Section 7.4.1 discusses D-optimal designs for the case when $(x_1, x_2) \in [0,\infty)\times[0,\infty)$ and $|\beta_0^*| < 1.5434$ while Section 7.4.2 considers the designs for the cases when $(x_1, x_2) \in [0,b] \times [0,b]$ and $|\beta_0^*| < 1.5434$. Section 7.4.3 discusses the designs for the cases when $(x_1, x_2) \in [0,\infty) \times [0,\infty)$ and $\beta_0^* < -1.5434$, and when $(x_1, x_2) \in [0,b] \times [0,b]$ and $\beta_0^* < -1.5434$.

7.4.1 D-optimal designs on $[0,\infty) \times [0,\infty)$ when $|\beta_0^*| < 1.5434$

Consider the two-variable binary logistic model (7.1) where $(x_1, x_2) \in [0,\infty) \times [0,\infty)$. Recall from Section 4.6.2 and Section 7.3 that if $|\beta_0^*| < u = u^* = 1.5434$ and $-\infty < x_1 < \infty$ and $c \le x_2 \le d$, then the 4 hyperbolic branches corresponding to logits $\pm u = 1.5434$ are spread in the 4 quadrants of $\mathbb{R}^2$ generated by the asymptote lines $x_1 = -\dfrac{\beta_2}{\beta_{12}}$ and $x_2 = -\dfrac{\beta_1}{\beta_{12}}$. If $(x_1, x_2) \in [0,\infty) \times [0,\infty)$, then only points located in the first quadrant of $\mathbb{R}^2$ can be the support points of the candidate D-optimal design. Figure 7.4, without the upper limits $x_1 = b$ and $x_2 = b$, displays the pattern of one such candidate D-optimal design. Also, since $\beta_0^* = \beta_0 - \dfrac{\beta_1\beta_2}{\beta_{12}} < \beta_0 \le 0$, then hyperbolic curves in Figure 7.4 satisfy the condition $-u = -u^* < \beta_0^* < \beta_0 \le 0$ where $u^* = 1.5434$, or, equivalently

$$0 \le -\beta_0 < -\beta_0^* \le 1.5434. \tag{7.14}$$

The construction of the D-optimal design with the design pattern displayed in Figure 7.4 can be done as follows. As in Section 5.3.3 on the 3-point D-optimal design, point $E(0,0)$ is one of the support points of the candidate D-optimal design. Therefore one of the optimal logits

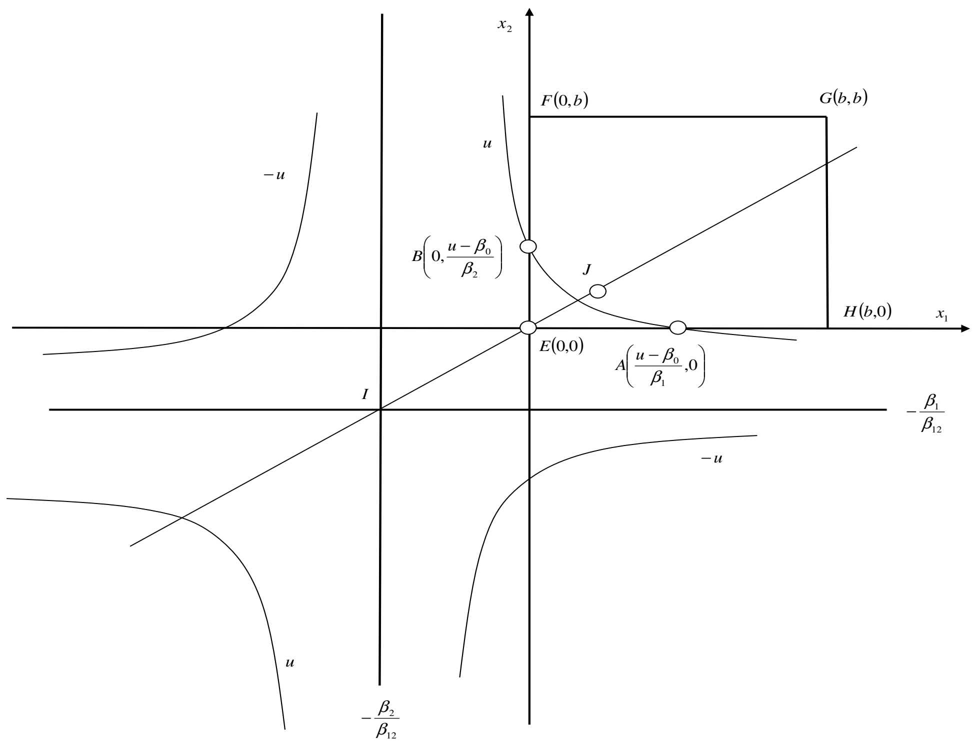


Figure 7.4: Design pattern of the candidate D-optimal design for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, \infty) \times [0, \infty)$ where $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$, $\beta_{12} > 0$, $\beta_0^* = \beta_0 - \frac{\beta_1 \beta_2}{\beta_{12}}$ and $0 < -\beta_0 < -\beta_0^* < 1.5434 < \infty$. Circles are support points.

is $u = \beta_0$. As with the case of the two-variable binary logistic model without interaction, the other two support points are $A(\dfrac{u-\beta_0}{\beta_1}, 0)$ and $B(0, \dfrac{u-\beta_0}{\beta_2})$, the points of intersection of the logit line $u = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ and the boundary of the design space $[0, \infty] \times [0, \infty]$. However, a design with support points A, B, and E is singular for estimating the interaction parameter β_{12} since the interaction term in model (7.1) vanishes at the three support points. Hence, there is a need of an interior support point, say J in Figure 7.4 at which the interaction term in the model does not vanish. In order to account for the potency of the second drug with respect to the first drug, the coordinates of the interior support point J should satisfy the condition $x_2 = \dfrac{1}{\rho} x_1$ where $\rho = \dfrac{\beta_2}{\beta_1}$ is the potency of the second drug relative to the first drug (see Section 2.4.4). Hence, $x_2 = \dfrac{\beta_1}{\beta_2} x_1$, i.e. J is the point of intersection of logit line, say u_1, and the interior ray joining $E(0, 0)$ and the centre $(-\frac{\beta_2}{\beta_{12}}, -\frac{\beta_1}{\beta_{12}})$ of the hyperbolae logit $\pm u$. Ray designs for similar models are also discussed in O'Brien (2004), and in Kupchak (2000, pp. 126-138). The coordinates of the interior point J are positive solutions for x_1 and x_2 to the simultaneous equations

$$\begin{cases} x_2 &= \frac{\beta_1 x_1}{\beta_2} \\ u_1 &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2. \end{cases} \tag{7.15}$$

The solution for (x_1, x_2) to the equations (7.15) is

$$\begin{cases} x_{1J} &= -\frac{\beta_2}{\beta_{12}} + \frac{\sqrt{\beta_1\beta_2[\beta_1\beta_2+\beta_{12}(u_1-\beta_0)]}}{\beta_1\beta_{12}} \\ x_{2J} &= -\frac{\beta_1}{\beta_{12}} + \frac{\sqrt{\beta_1\beta_2[\beta_1\beta_2+\beta_{12}(u_1-\beta_0)]}}{\beta_2\beta_{12}}. \end{cases} \tag{7.16}$$

Hence, the candidate D-optimal design ξ with support points represented by circles in Figure 7.4 is

$$\xi = \left\{ \begin{matrix} (x_1, x_2) \\ \lambda \\ u \end{matrix} \right\} = \left\{ \begin{matrix} (0,0) & \left(0, \frac{u-\beta_0}{\beta_2}\right) & \left(\frac{u-\beta_0}{\beta_1}, 0\right) & (x_{1J}, x_{2J}) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \beta_0 & u & u & u_1 \end{matrix} \right\} \tag{7.17}$$

where x_{1J} and x_{2J} are given by (7.16). The information matrix for the parameter vector $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$ evaluated at design (7.17) is

$$M(\xi; \boldsymbol{\beta}) = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & m_{44} \end{bmatrix}$$

where

$$\begin{aligned}
m_{11} &= \frac{1}{4}\left(\frac{2e^u}{(1+e^u)^2} + \frac{e^{u_1}}{(1+e^{u_1})^2} + \frac{e^{\beta_0}}{\left(1+e^{\beta_0}\right)^2}\right) \\
m_{12} &= m_{21} = \frac{e^u(u-\beta_0)}{4\beta_1(1+e^u)^2} + \frac{e^{u_1}\left(\sqrt{\beta_1\beta_2(-\beta_0\beta_{12}+u_1\beta_{12}+\beta_1\beta_2)}-\beta_1\beta_2\right)}{4\beta_1\beta_{12}(1+e^{u_1})^2} \\
m_{13} &= m_{31} = \frac{e^u(u-\beta_0)4\beta_2(1+e^u)^2}{+}\frac{e^{u_1}\left(\sqrt{\beta_1\beta_2(-\beta_0\beta_{12}+u_1\beta_{12}+\beta_1\beta_2)}-\beta_1\beta_2\right)}{4\beta_2\beta_{12}(1+e^{u_1})^2} \\
m_{14} &= m_{41} = \frac{e^{u_1}\left(\sqrt{\beta_1\beta_2(-\beta_0\beta_{12}+u_1\beta_{12}+\beta_1\beta_2)}-\beta_1\beta_2\right)^2}{4\beta_1\beta_{12}^2\beta_2(1+e^{u_1})^2} \\
m_{22} &= \frac{e^u(\beta_0-u)^2}{4\beta_1^2(1+e^u)^2} + \frac{e^{u_1}\left(\sqrt{\beta_1\beta_2(-\beta_0\beta_{12}+u_1\beta_{12}+\beta_1\beta_2)}-\beta_1\beta_2\right)^2}{4\beta_1^2\beta_{12}^2(1+e^{u_1})^2} \\
m_{23} &= m_{32} = \frac{e^{u_1}\left(\sqrt{\beta_1\beta_2(-\beta_0\beta_{12}+u_1\beta_{12}+\beta_1\beta_2)}-\beta_1\beta_2\right)^2}{4\beta_1\beta_{12}^2\beta_2(1+e^{u_1})^2} \\
m_{24} &= m_{42} = \frac{e^{u_1}\left(\sqrt{\beta_1\beta_2(-\beta_0\beta_{12}+u_1\beta_{12}+\beta_1\beta_2)}-\beta_1\beta_2\right)^3}{4\beta_1^2\beta_{12}^3\beta_2(1+e^{u_1})^2} \\
m_{33} &= \frac{e^u(\beta_0-u)^2}{4\beta_2^2(1+e^u)^2} + \frac{e^{u_1}\left(\sqrt{\beta_1\beta_2(-\beta_0\beta_{12}+u_1\beta_{12}+\beta_1\beta_2)}-\beta_1\beta_2\right)^2}{4\beta_2^2\beta_{12}^2(1+e^{u_1})^2} \\
m_{34} &= m_{43} = \frac{e^{u_1}\left(\sqrt{\beta_1\beta_2(-\beta_0\beta_{12}+u_1\beta_{12}+\beta_1\beta_2)}-\beta_1\beta_2\right)^3}{4\beta_1\beta_{12}^3\beta_2^2(1+e^{u_1})^2} \\
m_{44} &= \frac{e^{u_1}\left(\sqrt{\beta_1\beta_2(-\beta_0\beta_{12}+u_1\beta_{12}+\beta_1\beta_2)}-\beta_1\beta_2\right)^4}{4\beta_1^2\beta_{12}^4\beta_2^2(1+e^{u_1})^2}.
\end{aligned} \tag{7.18}$$

The determinant of $M(\xi;\boldsymbol{\beta})$ is

$$D = \frac{(u-\beta_0)^4 e^{\beta_0+2u+u_1}\left\{\beta_{12}^2(u_1-\beta_0)^2 + 8\beta_1\beta_2[\beta_1\beta_2+\beta_{12}(u_1-\beta_0)] - 4[2\beta_1\beta_2+\beta_{12}(u_1-\beta_0)]\sqrt{\beta_1\beta_2[\beta_1\beta_2+\beta_{12}(u_1-\beta_0)]}\right\}}{256\beta_1^2\beta_2^2\beta_{12}^4(1+e^{\beta_0})^2(1+e^{u})^4(1+e^{u_1})^2}$$

which is proportional to the product of

$$D_u = \frac{(u-\beta_0)^4 e^{2u}}{(1+e^u)^4} \tag{7.19}$$

and

$$D_{u_1} = \frac{\left\{\beta_{12}^2(u_1-\beta_0)^2 + 8\beta_1\beta_2[\beta_1\beta_2+\beta_{12}(u_1-\beta_0)] - 4[2\beta_1\beta_2+\beta_{12}(u_1-\beta_0)]\sqrt{\beta_1\beta_2[\beta_1\beta_2+\beta_{12}(u_1-\beta_0)]}\right\} e^{u_1}}{(1+e^{u_1})^2}. \tag{7.20}$$

Expression (7.19) is the same as the determinant (5.36) which was shown to have equal maximum values at $u = \beta_0$ and $u = u^*$ where u^* is the unique solution, on $[0,\infty)$, of the equation

$$2 - \beta_0 + u + 2e^u + \beta_0 e^u - ue^u = 0 \tag{7.21}$$

for $0 \le -\beta_0 \le 1.5434$ (see equation (5.38)). The second row of Table 5.3 contains values of u^* for selected values of β_0.

For a given value of the parameter vector $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$, the value of u_1 which maximizes (7.20) can be found numerically. The D-optimality of the numerically constructed design (7.17) can be checked using the graph of the standardized variance function $d(\boldsymbol{x}, \xi; \boldsymbol{\beta})$ versus $\boldsymbol{x}^T = (x_1, x_2) \in [0,\infty] \times [0,\infty]$ as illustrated in the following numerical example.

Example 7.1. Consider the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ where $(x_1, x_2) \in [0,\infty] \times [0,\infty]$ and with parameter values $\beta_0 = -0.5$, $\beta_1 = 2$, $\beta_2 = 1.5$ and $\beta_{12} = 3$. In this example, $\beta_{12} = 3 > 0$ and $\beta_0^* = \beta_0 - \frac{\beta_1\beta_2}{\beta_{12}} = -1.5$ so that $0 < -\beta_0 = 0.5 < -\beta_0^* = 1.5 < 1.5434$. Hence, condition (7.14) is satisfied. Therefore, the expected D-optimal design is of the form (7.17). Since, $\beta_0 = -0.5 > -1.5434$, the second row of Table 5.3 gives $u^* = 2.075$ as solution for u to the equation (7.21). Furthermore, the function (7.20) is

$$D_{u_1} = \frac{\left\{110.25 + 81u_1 + 9u_1^2 - (51.96 + 20.79u_1)\sqrt{4.5 + 3u_1}\right\} e^{u_1}}{(1+e^{u_1})^2}.$$

The graph of D_{u_1} versus $u_1 \ge \beta_0 = -0.5$ displayed in Figure 7.5 indicates that the D_{u_1} has only one maximum on $[-0.5, \infty)$. Differentiating D_{u_1} with respect to u_1 and numerically solving $\dfrac{\partial D_{u_1}}{\partial u_1} = 0$ for u_1 gives the solution $u_1^* = 2.832$. Setting $\beta_0 = -0.5$, $\beta_1 = 2$, $\beta_2 = 1.5$, $\beta_{12} = 3$, $u = 2.075$ and $u_1 = 2.832$ in design (7.17) gives

$$\xi^* = \left\{\begin{array}{cccc} (0,0) & (0,1.717) & (1.287,0) & (0.541,0.721) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ -0.5 & 2.075 & 2.075 & 2.832 \end{array}\right\}. \tag{7.22}$$

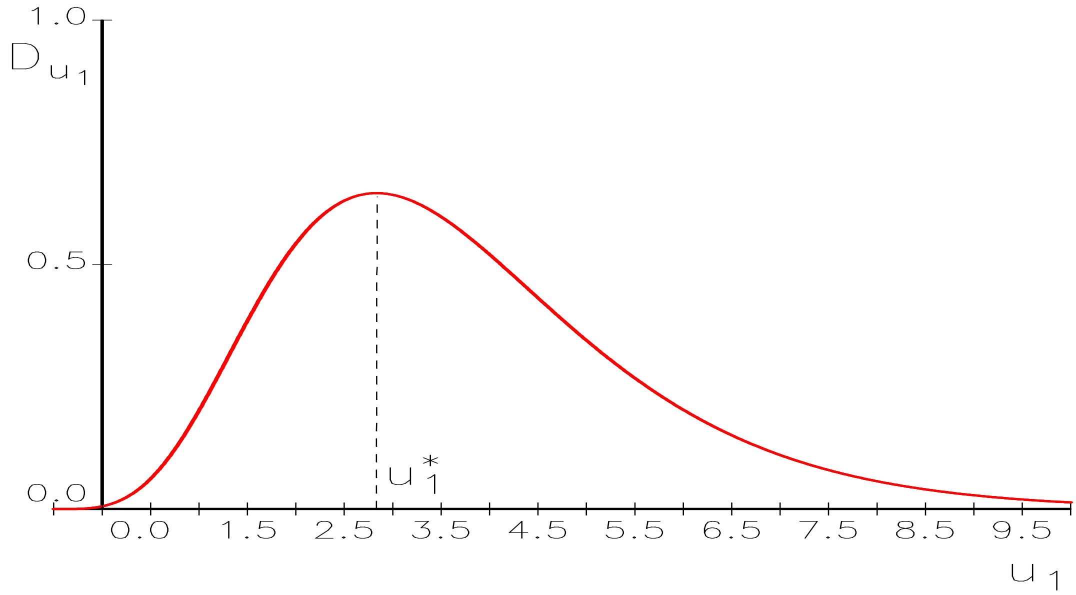


Figure 7.5: Graph of the function $D_{u_1} = \frac{\left\{110.25+81u_1+9u_1^2-(51.96+20.79u_1)\sqrt{4.5+3u_1}\right\}e^{u_1}}{(1+e^{u_1})^2}$ versus $u_1 \geq \beta_0 = -0.5$.

The support points of the design (7.22) are represented by triangles and a diamond in Figure 7.6 (a). The standardized variance function associated with design (7.22) is

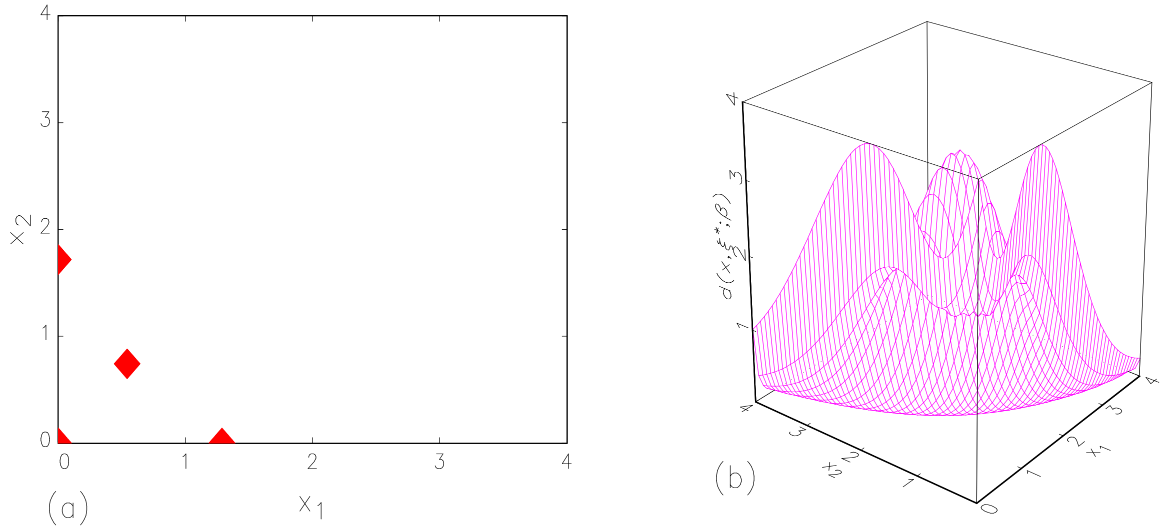


Figure 7.6: (a) Support points of design (7.22) and (b) the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ on the design space $[0, \infty] \times [0, \infty]$ where $\boldsymbol{\beta} = (-0.5, 2, 1.5, 3)^T$. The triangles and a diamond are support points of design (7.22).

$$d(\boldsymbol{x},\xi^*;\boldsymbol{\beta}) = \tfrac{e^u}{(1+e^u)^2}\Big\{17.02 - 26.44x_1 + 34.62x_1^2 - 19.83x_2 + 1.42x_1x_2 - 56.69x_1^2x_2 \\ +19.47x_2^2 - 42.52x_1x_2^2 + 597.84x_1^2x_2^2\Big\} \tag{7.23}$$

where $u = -0.5 + 2x_1 + 1.5x_2 + 3x_1x_2$ and $(x_1, x_2) \in [0,\infty) \times [0,\infty)$. The function (7.23) plotted in Figure 7.6 (b) has maximum value of 4 at the 4 support point of the candidate D-optimal design (7.22). Hence, the design (7.22) is D-optimal.

7.4.2 D-optimal designs on $[0, b] \times [0, b]$ when $|\beta_0^*| < 1.5434$

Reconsider Figure 7.4 but with the design space restricted to the square $[0, b] \times [0, b]$. Restricting the design space to $[0, b] \times [0, b]$ is practical as was argued in Section 7.3.2, and gives rise to the following candidate D-optimal design cases.

Case when $1.5434 \leq \min(\beta_0 + \beta_1 b, \beta_0 + \beta_2 b)$

In this case, the points $(0, b)$ and $(b, 0)$ lie above the logit line $u = 1.5434$. This condition combined with condition (7.14) gives the design condition

$$0 < -\beta_0 < -\beta_0^* < 1.5434 \leq \min(\beta_0 + \beta_1 b, \beta_0 + \beta_2 b) \tag{7.24}$$

where $\beta_0^* = \beta_0 - \frac{\beta_1\beta_2}{\beta_{12}}$. The candidate D-optimal design which satisfies condition (7.24) is design (7.17). Furthermore, the parameter values in Example 7.1 satisfy condition (7.24) since

$$0 < -\beta_0 = 0.5 < -\beta_0^* = 1.5 < 1.5434 \leq \ \min(\beta_0 + \beta_1 b, \beta_0 + \beta_2 b) = 2.5.$$

Case when $\beta_0 + \beta_1 b < 1.5434 < \beta_0 + \beta_2 b$

In this case, the points $(b, 0)$ and $(0, b)$ lie below and above the logit line $u = 1.5434$. This condition together with condition (7.14) gives the design condition

$$-1.5434 < \beta_0^* < \beta_0 < \beta_0 + \beta_1 b \leq 1.5434 \leq \beta_0 + \beta_2 b \tag{7.25}$$

where $\beta_0^* = \beta_0 - \dfrac{\beta_1\beta_2}{\beta_{12}}$. The candidate D-optimal design which satisfies condition (7.25) is design (7.17) with support point $(\dfrac{u-\beta_0}{\beta_1}, 0)$ replaced by $(b, 0)$. That is, design

$$\xi = \left\{ \begin{array}{cccc} (0,0) & (0, \frac{u-\beta_0}{\beta_2}) & (b,0) & (x_{1J}, x_{2J}) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \beta_0 & u & \beta_0 + \beta_1 b & u_1 \end{array} \right\}. \tag{7.26}$$

The determinant of the information matrix for $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$ evaluated at design (7.26) is the product of (7.20) and

$$D = \frac{(u-\beta_0)^2(\beta_0+\beta_1 b-\beta_0)^2 e^{u+\beta_0+\beta_1 b}}{(1+e^u)^2(1+e^{\beta_0+\beta_1 b})^2} = \frac{\beta_1^2 b^2 (u-\beta_0)^2 e^{u+\beta_0+\beta_1 b}}{(1+e^u)^2(1+e^{\beta_0+\beta_1 b})^2}$$

which is proportional to $D_u = \frac{(u-\beta_0)^2 e^u}{(1+e^u)^2}$, and the latter is the square root of (7.19). Hence, the optimal value of u is still the solution for u to the equation (7.21).

Example 7.2. Consider the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0,2] \times [0,2]$ where $\beta_0 = -0.5$, $\beta_1 = 1$, $\beta_2 = 2$ and $\beta_{12} = 3$. In this example, $\beta_0^* = \beta_0 - \frac{\beta_1 \beta_2}{\beta_{12}} \simeq -1.17$, $\beta_0 + \beta_1 b = 1.5$ and $\beta_0 + \beta_2 b = 3.5$ so that condition (7.25) is satisfied. Thus, the expected D-optimal design should be (7.26) with optimal logits $\beta_0 = -0.5$ and $\beta_0 + \beta_1 b = 1.5$, and the respective optimal logits u and u_1 which maximize (7.19) and (7.20). For $\beta_0 = -0.5$, Table 5.3 gives $u = 2.075$ as the optimal value of logit u. Setting $\beta_0 = -0.5$, $\beta_1 = 1$, $\beta_2 = 2$ and $\beta_{12} = 3$ in D_{u_1} given by (7.20), and then solving $\frac{\partial D_{u_1}}{\partial u_1} = 0$ for u_1 gives $u_1 = 2.725$ as the optimal value of logit u_1. Hence, the candidate D-optimal design (7.26) is

$$\xi^* = \left\{ \begin{array}{cccc} (0,0) & (0,1.287) & (2,0) & (0.944, 0.472) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ -0.5 & 2.075 & 1.5 & 2.725 \end{array} \right\}. \tag{7.27}$$

The support points of design (7.27) are represented by triangles and a diamond in Figure 7.7 (a). The standardized variance function corresponding to design (7.27) is

$$\begin{aligned} d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) \; = \; & \tfrac{e^u}{(1+e^u)^2}\Big\{17.02 - 17.02x_1 + 10.96x_1^2 - 26.44x_2 + 0.89x_1x_2 - 22.25x_1^2x_2 \\ & +34.62x_2^2 - 42.01x_1x_2^2 + 408.56x_1^2x_2^2\Big\} \end{aligned} \tag{7.28}$$

where $u = -0.5 + x_1 + 2x_2 + 3x_1x_2$ and $(x_1, x_2) \in [0,2] \times [0,2]$. The function (7.28) plotted in Figure 7.7 (b) clearly indicates that design (7.27) is D-optimal.

Case when $\beta_0 + \beta_2 b < 1.5434 \le \beta_0 + \beta_1 b$

This case is similar to the case when $\beta_0 + \beta_1 b < 1.5434 \le \beta_0 + \beta_2 b$ (discussed above) except that now the respective points $(0, b)$ and $(b, 0)$ lie below and above the logit line $u = 1.5434$. Thus, the design condition (7.25) becomes

$$-1.5434 < \beta_0^* < \beta_0 < \beta_0 + \beta_2 b < 1.5434 \le \beta_0 + \beta_1 b \tag{7.29}$$

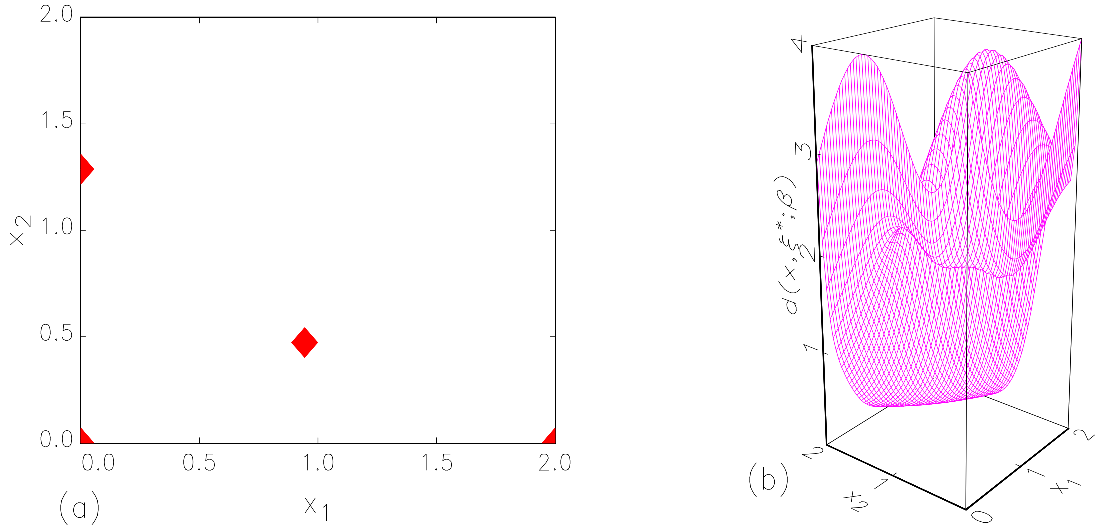


Figure 7.7: (a) Support points of design (7.27) and (b) the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ on the design space $[0, 2] \times [0, 2]$ where $\boldsymbol{\beta} = (-0.5, 1, 2, 3)^T$. The triangles and a diamond are support points of design (7.27).

where $\beta_0^* = \beta_0 - \frac{\beta_1\beta_2}{\beta_{12}}$. The candidate D-optimal design which satisfies condition (7.29) is design (7.17) with support point $B(0, \frac{u-\beta_0}{\beta_2})$ replaced by $C(0, b)$. Furthermore, the construction of the D-optimal design which satisfies condition (7.29) is as for the case when $\beta_0 + \beta_1 b < 1.5434 < \beta_0 + \beta_2 b$ above.

Case when $[\beta_0 + \beta_1 b, \beta_0 + \beta_2 b] \leq 1.5434$

In this case, the points $(b, 0)$ and $(0, b)$ lie on or below the logit line $u = 1.5434$. This condition together with condition (7.14) gives the design condition

$$-1.5434 < \beta_0^* < \beta_0 < [\beta_0 + \beta_1 b, \beta_0 + \beta_2 b] \leq 1.5434. \tag{7.30}$$

The candidate D-optimal design which satisfies condition (7.30) is design (7.17) with the respective support points $\left(\frac{u-\beta_0}{\beta_1}, 0\right)$ and $\left(0, \frac{u-\beta_0}{\beta_2}\right)$ replaced by $(b, 0)$ and $(0, b)$. That is, design

$$\xi = \left\{ \begin{array}{cccc} (0,0) & (0,b) & (b,0) & (x_{1J}, x_{2J}) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \beta_0 & \beta_0 + \beta_2 b & \beta_0 + \beta_1 b & u_1 \end{array} \right\}. \tag{7.31}$$

In this case, the determinant of the information matrix for $\beta = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$ evaluated at design (7.31) is proportional to (7.20).

Example 7.3. Consider the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0,2] \times [0,2]$ where $\beta_0 = -1$, $\beta_1 = 0.9$, $\beta_2 = 1.2$ and $\beta_{12} = 3$. The parameter values in this example satisfy condition (7.30) since

$$-1.5434 < \beta_0^* = -1.36 < \beta_0 = -1 < \beta_0 + \beta_1 b = 0.8 < \beta_0 + \beta_2 b = 1.4 < 1.5434.$$

Hence, the candidate D-optimal is (7.31) where the logit u_1 maximizes (7.20). Setting $\beta_0 = -1$, $\beta_1 = 0.9$, $\beta_2 = 1.2$ and $\beta_{12} = 3$ in (7.20), and then solving $\dfrac{\partial D_{u_1}}{\partial u_1} = 0$ for u_1 gives $u_1 = 2.151$ so that the candidate D-optimal design is

$$\xi^* = \left\{ \begin{array}{cccc} (0,0) & (0,2) & (2,0) & (0.867, 0.650) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ -1 & 1.4 & 0.8 & 2.251 \end{array} \right\}. \tag{7.32}$$

The support points of design (7.32) are represented by triangles and a diamond in Figure 7.8 (a). The standardized variance function associated with design (7.32) is

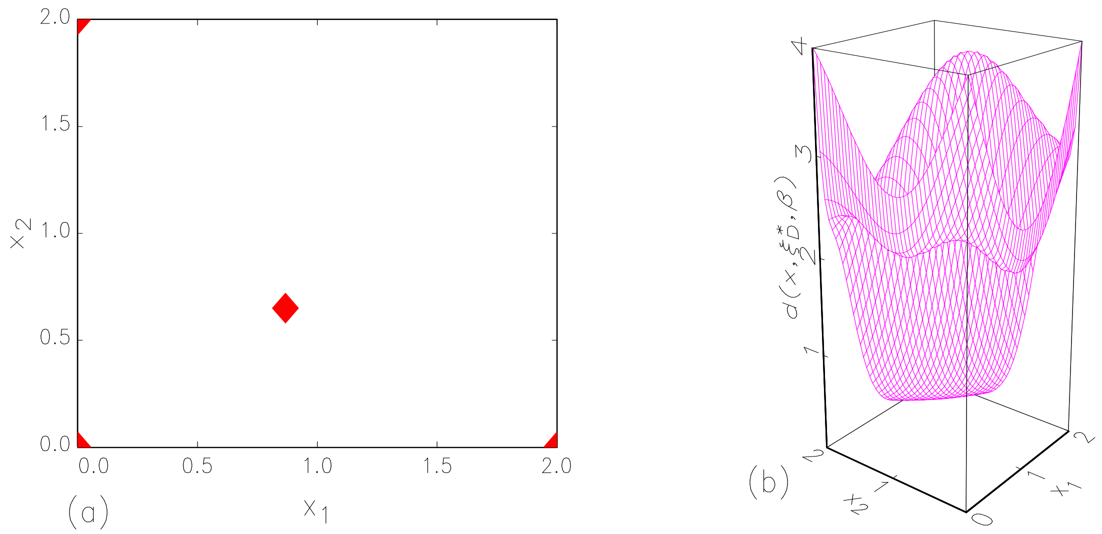


Figure 7.8: (a) Support points of design (7.32) and (b) the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ on the design space $[0,2] \times [0,2]$ where $\boldsymbol{\beta} = (-1, 0.9, 1.2, 3)^T$. The triangles and a diamond are support points of design (7.32).

$$\begin{aligned} d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) \;\; = \;\; & \tfrac{e^u}{(1+e^u)^2} \Big\{ 20.34 - 20.34 x_1 + 9.76 x_1^2 - 20.34 x_2 + 7.27 x_1 x_2 - 5.66 x_1^2 x_2 \\ & + 11.39 x_2^2 - 5.82 x_1 x_2^2 + 169.31 x_1^2 x_2^2 \Big\} \end{aligned} \tag{7.33}$$

where $u = -1 + 0.9x_1 + 1.2x_2 + 3x_1 x_2$ and $(x_1, x_2) \in [0,2] \times [0,2]$. The graph of function (7.33) displayed in Figure 7.8 (b) attains the maximum of 4 at the 4 support points of design (7.32), and hence design (7.32) is D-optimal.

7.4.3 D-optimal designs on $[0,\infty)\times[0,\infty)$ or $[0,b]\times[0,b]$ when $\beta_0^* < -1.5434$

Now, consider the case when $\beta_0^* < -1.5434$. As discussed in Section 4.6.2, with the help of Figure 4.8 (b), the hyperbolic curves corresponding to this case are in the first and third quadrants generated by the asymptotes $x_1 = -\dfrac{\beta_2}{\beta_{12}}$ and $x_2 = -\dfrac{\beta_1}{\beta_{12}}$. However, in the case of synergy on the design space $[0,\infty)\times[0,\infty)$, only hyperbolic branches in the first quadrant intersect the design space. Furthermore, as $\beta_0^* < \beta_0$, condition $\beta_0^* < -1.5434$ can be split into two cases: (a) $\beta_0^* < -1.5434 \le \beta_0 < 1.5434$; and (b) $\beta_0^* < \beta_0 < -1.5434$. The construction of D-optimal designs for these two cases will be discussed separately.

<u>D-optimal designs on $[0,\infty)\times[0,\infty)$ or $[0,b]\times[0,b]$ when $\beta_0^* < -1.5434 \le \beta_0 < 1.5434$</u>

Firstly, consider the design pattern of the candidate D-optimal design, on the design space $[0,\infty)\times[0,\infty)$ and for the case when $\beta_0^* < -1.5434 \le \beta_0 < 1.5434$, displayed in Figure 7.9 (a). Since the design space and the condition $-1.5434 \le \beta_0$ are as in Section 7.4.1, all the results of Section 7.4.1 apply to this case. Hence, the candidate D-optimal design in this case is of the form (7.17).

Secondly, consider the design pattern of the candidate D-optimal design, on the design space $[0,b]\times[0,b]$ and for the case when $\beta_0^* < -1.5434 \le \beta_0 < 1.5434$, displayed in Figure 7.9 (a). The D-optimal designs are constructed on this design space in the same way as discussed in Section 7.4.2. The following numerical example illustrates the construction of the designs.

Example 7.4. Consider the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0,2]\times[0,2]$ where $\beta_0 = -1.5$, $\beta_1 = 3$, $\beta_2 = 2$ and $\beta_{12} = 4$. The parameter values in this example satisfy the condition $\beta_0^* < -1.5434 \le \beta_0 < 1.5434$ since $\beta_0^* = -3 < -1.5434 < \beta_0 = -1.5 < 1.5434$. Hence, a 4-point D-optimal design with $\beta_0 = -1.5$ as one of the optimal logits is expected. The other two optimal logits u and u_1 are obtained as follows. Table 5.3 gives $u = 1.562$ as the optimal value of logit u if $\beta_0 = -1.5$. Setting $\beta_0 = -1.5$, $\beta_1 = 3$, $\beta_2 = 2$ and $\beta_{12} = 4$ in D_{u_1} given by (7.20), and then solving $\dfrac{\partial D_{u_1}}{\partial u_1} = 0$ for u_1 gives $u_1 = 2.272$ as the optimal value of logit u_1. Hence, the candidate D-optimal design (7.17) is

$$\xi^* = \left\{ \begin{array}{cccc} (0,0) & (1.021,0) & (0,1.531) & (0.437,0.656) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ -1.5 & 1.562 & 1.562 & 2.272 \end{array} \right\}. \tag{7.34}$$

The support points of design (7.34) are represented by triangles and a diamond in Figure 7.10

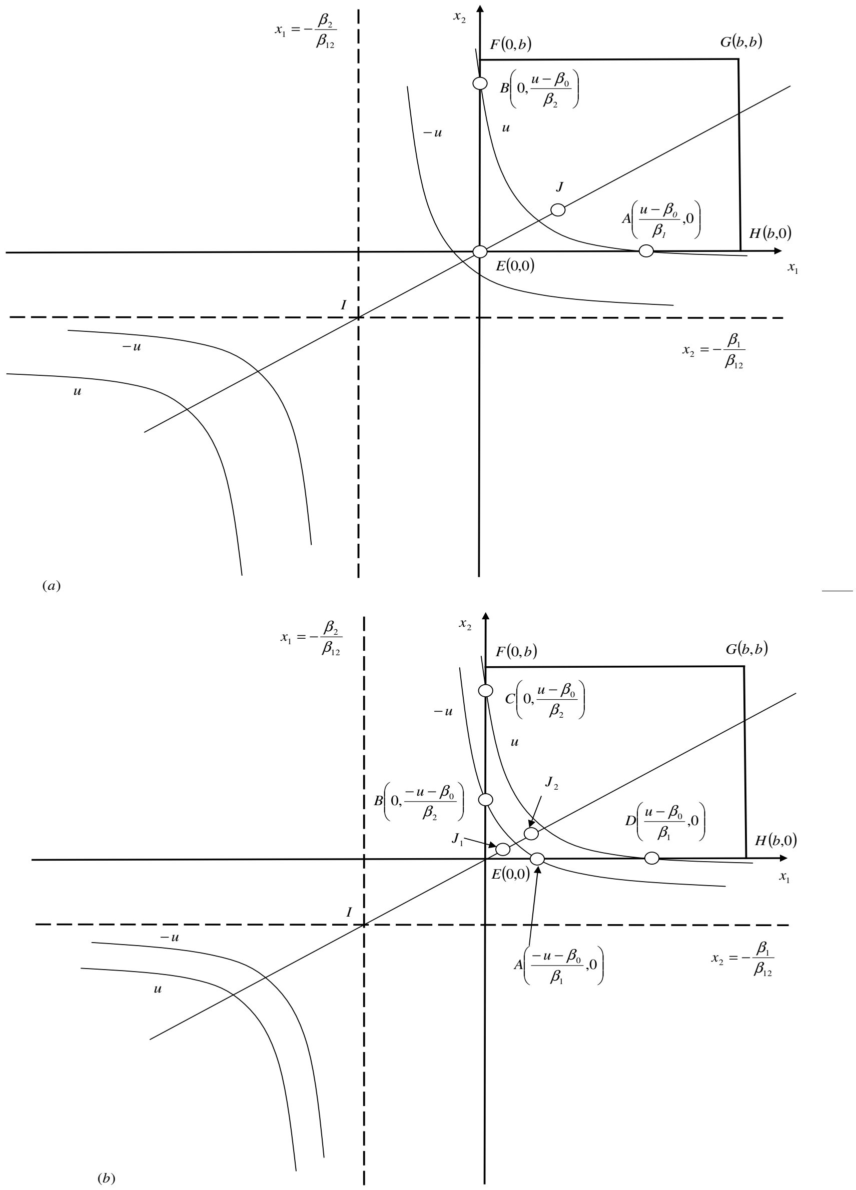


Figure 7.9: Design patterns of the candidate D-optimal design for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ where $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_{12} > 0$, and (a) $\beta_0^* < -1.5434 \leq \beta_0 < 1.5434$ and (b) $\beta_0^* < \beta_0 < -1.5434$. Circles are support points.

(a). The standardized variance function associated with design (7.34) is

$$\begin{aligned} d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) \quad = \quad & \tfrac{e^u}{(1+e^u)^2} \Big\{ 26.82 - 52.54x_1 + 52.52x_1^2 - 35.04x_2 + 7.46x_1x_2 - 55.39x_1^2x_2 \\ & +23.35x_2^2 - 36.962x_1x_2^2 + 704.70x_1^2x_2^2 \Big\} \end{aligned} \tag{7.35}$$

where $u = -1.5 + 3x_1 + 2x_2 + 4x_1x_2$ and $(x_1, x_2) \in [0, 2] \times [0, 2]$. The graph of function (7.35)

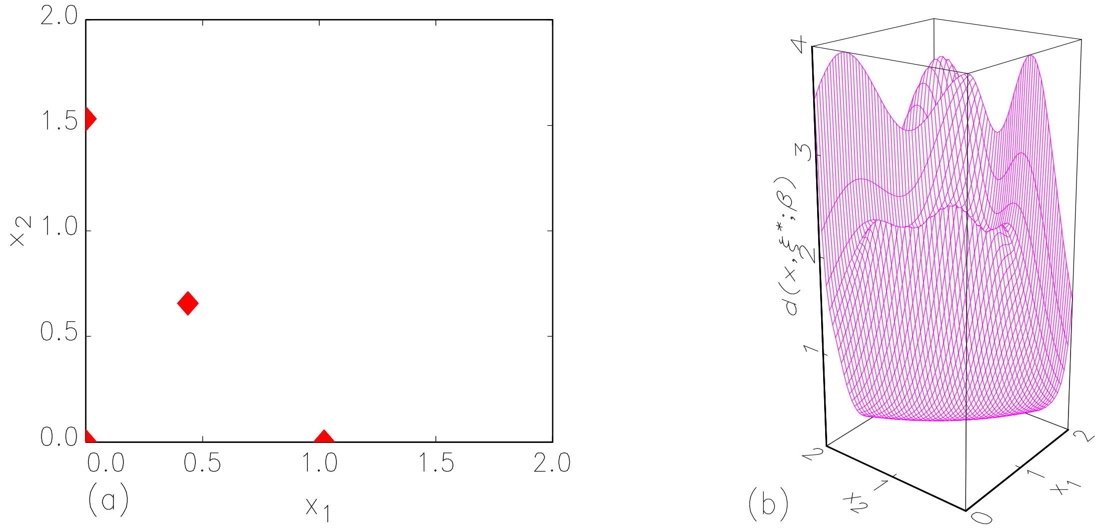


Figure 7.10: (a) Support points of design (7.34) and (b) the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ given in (7.35) on the design space $[0, 2] \times [0, 2]$ and with parameter vector $\boldsymbol{\beta} = (-1.5, 3, 2)^T$. The triangles and a diamond are support points of design (7.34).

displayed in Figure 7.10 (b) attains the maximum of 4 at the 4 support points of design (7.34), and hence the design (7.34) is D-optimal.

D-optimal designs on $[0, \infty) \times [0, \infty)$ or $[0, b] \times [0, b]$ when $\beta_0^* < \beta_0 < -1.5434$

Consider the design pattern of the candidate D-optimal design, on the design space $[0, \infty) \times [0, \infty)$ and for the case when $\beta_0^* < \beta_0 < -1.5434$, displayed in Figure 7.9 (b). In this case both logit lines $\pm u = 1.5434$ pass through the design space $[0, \infty) \times [0, \infty)$. As in Section 7.4.1, the support points of the candidate D-optimal design are: the points of intersection A, B, C and D of the logit lines of the form $\pm u$ and the boundary of the design space $[0, \infty) \times [0, \infty)$; and the respective points of intersection J_1 and J_2 of the ray $x_2 = \dfrac{\beta_1}{\beta_2} x_1$ and the logit lines u_1 and u_2 where $u_1 < u_2$ (see Figure 7.9 (b)). Thus, the candidate D-optimal design has at most 6 support points. The analytical construction of this design is cumbersome because of the large number of support points and the fact that the support points are not all equally weighted. However, the design can be numerically constructed. Under the assumption that the support points on the same logit line are equally weighted, the candidate D-optimal design on the design space $[0,\infty)\times[0,\infty)$ is given by

$$\xi=\left\{\begin{array}{cccccc}\left(\frac{-u-\beta_0}{\beta_1},0\right) & \left(0,\frac{-u-\beta_0}{\beta_2}\right) & \left(0,\frac{u-\beta_0}{\beta_2}\right) & \left(\frac{u-\beta_0}{\beta_1},0\right) & (x_{1J_1},x_{2J_1}) & (x_{1J_2},x_{2J_2})\\ \frac{\lambda_1}{2} & \frac{\lambda_1}{2} & \frac{\lambda_2}{2} & \frac{\lambda_2}{2} & \lambda_3 & 1-\lambda_1-\lambda_2-\lambda_3\\ -u & -u & u & u & u_1 & u_2\end{array}\right\} \tag{7.36}$$

where

$$\left\{\begin{array}{rcl} x_{1J_1} & = & -\frac{\beta_2}{\beta_{12}}+\frac{\sqrt{\beta_1\beta_2[\beta_1\beta_2+\beta_{12}(u_1-\beta_0)]}}{\beta_1\beta_{12}}\\ x_{2J_1} & = & -\frac{\beta_1}{\beta_{12}}+\frac{\sqrt{\beta_1\beta_2[\beta_1\beta_2+\beta_{12}(u_1-\beta_0)]}}{\beta_2\beta_{12}}\end{array}\right. \tag{7.37}$$

and

$$\left\{\begin{array}{rcl} x_{1J_2} & = & -\frac{\beta_2}{\beta_{12}}+\frac{\sqrt{\beta_1\beta_2[\beta_1\beta_2+\beta_{12}(u_2-\beta_0)]}}{\beta_1\beta_{12}}\\ x_{2J_2} & = & -\frac{\beta_1}{\beta_{12}}+\frac{\sqrt{\beta_1\beta_2[\beta_1\beta_2+\beta_{12}(u_2-\beta_0)]}}{\beta_2\beta_{12}}.\end{array}\right. \tag{7.38}$$

Now, consider constructing candidate D-optimal designs on the design space $[0,b]\times[0,b]$ and for the case when $\beta_0^*<\beta_0<-1.5434$. The restriction of the design space to $[0,b]\times[0,b]$ implies that the design of the form (7.36) exists only when the logit line $u^*=1.5434$ lies below the vertices $(0,b)$ and $(b,0)$ of the design space $[0,b]\times[0,b]$. If the logit line $u^*=1.5434$ lies above at least one of the vertices $(0,b)$ and $(b,0)$, it can be conjectured that a candidate D-optimal design does not necessarily follow the pattern of design (7.36). The conditions of existence of possible D-optimal designs on the design space $[0,b]\times[0,b]$ are given below, and selected numerical examples are used to demonstrate that the number of support points range from 4 to 6, and associated weights are not necessarily equal.

<u>Case when $\beta_0<-1.5434<1.5434<\min(\beta_0+\beta_1 b,\beta_0+\beta_2 b)$</u>

In this case, the point $(0,0)$ lies below the logit line $u=-1.5434$ and the points $(0,b)$ and $(b,0)$ lie above the logit line $u=1.5434$ as indicated in Figure 7.9 (b). This condition together with condition $\beta_0^*<\beta_0<-1.5434$ gives the design condition

$$\beta_0^*<\beta_0<-1.5434<1.5434<\min(\beta_0+\beta_1 b,\beta_0+\beta_2 b). \tag{7.39}$$

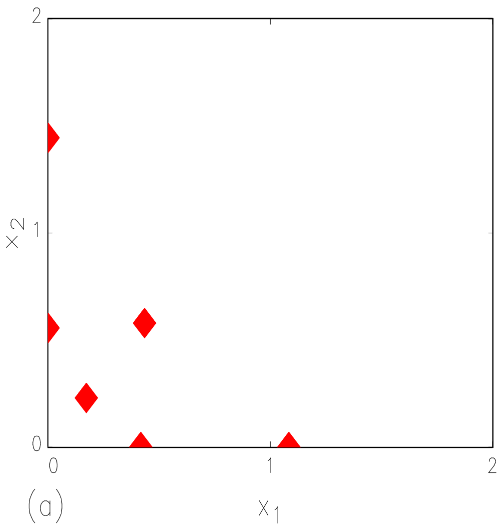
The candidate D-optimal design which satisfies condition (7.39) is still design (7.36).

Example 7.5. Consider the two-variable binary logistic model with interaction $u=\text{logit}(p)=\beta_0+\beta_1x_1+\beta_2x_2+\beta_{12}x_1x_2$ on the design space $[0,2]\times[0,2]$ where $\beta_0=-3$, $\beta_1=4$, $\beta_2=3$ and $\beta_{12}=5$. In this case,
$\beta_0^*=-5.4<\beta_0=-3<-1.5434<1.5434<\min(\beta_0+\beta_1 b,\beta_0+\beta_2 b)=3$. Therefore, condition (7.39) is satisfied, and the D-optimal design is expected to be of the form (7.36). The design

was calculated using the Gauss program in Appendix C and the result was the 6-point design

$$\xi^* = \left\{ \begin{array}{cccccc} (0.417, 0) & (1.083, 0) & (0, 0.556) & (0, 1.444) & (0.172, 0.229) & (0.434, 0.579) \\ 0.0895 & 0.2429 & 0.0894 & 0.2429 & 0.0870 & 0.2483 \\ -1.333 & 1.333 & -1.333 & 1.333 & -1.431 & 1.734 \end{array} \right\}. \quad (7.40)$$

Figure 7.11 (a) displays the location of the support points of design (7.40) on the design space $[0, 2] \times [0, 2]$, and Figure 7.11 (b) displays the graph of the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$. The graph in Figure 7.11 (b) suggests that design (7.40) is D-optimal.



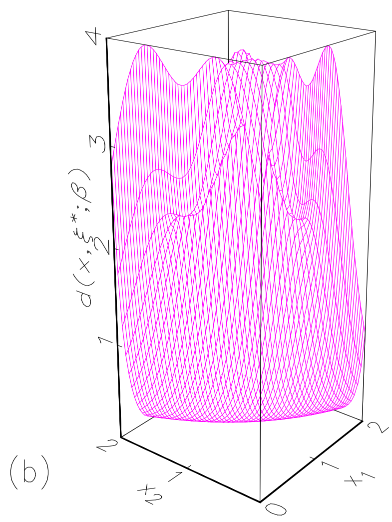


Figure 7.11: (a) Support points and (b) the standardized variance function of the D-optimal design for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, 2] \times [0, 2]$ and with parameters $\beta_0 = -3$, $\beta_1 = 4$, $\beta_2 = 3$ and $\beta_{12} = 5$. The triangles and diamonds are support points of design (7.40).

<u>Case when $\beta_0 < -1.5434 < \beta_0 + \beta_1 b \leq 1.5434 < \beta_0 + \beta_2 b$</u>

In this case, the point $(0, 0)$ lies below the logit line $u = -1.5434$, the respective points $(b, 0)$ and $(0, b)$ lie below and above the logit line $u = 1.5434$. This condition together with the condition $\beta_0^* < \beta_0 < -1.5434$ gives the design condition

$$\beta_0^* < \beta_0 < -1.5434 < \beta_0 + \beta_1 b \leq 1.5434 < \beta_0 + \beta_2 b. \quad (7.41)$$

The candidate D-optimal design which satisfies condition (7.41) is design (7.36) with support point $\left(\frac{u-\beta_0}{\beta_1}, 0\right)$ replaced by $(b, 0)$. In the candidate D-optimal design, the support points $\left(\frac{u-\beta_0}{\beta_1}, 0\right)$ and $(b, 0)$ no longer lie on the same logit line. The effect of this is that the weights of these points as well as those of the support points $\left(\frac{-u-\beta_0}{\beta_1}, 0\right)$ and $\left(0, \frac{-u-\beta_0}{\beta_2}\right)$ may no longer conjectured to be equal. This is confirmed by the following examples.

Example 7.6. Consider the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, 2] \times [0, 2]$ where $\beta_0 = -3$, $\beta_1 = 2$, $\beta_2 = 3$ and $\beta_{12} = 5$. In this example, $\beta_0^* = -4.2 < \beta_0 = -3 < -1.5434 < \beta_0 + \beta_1 b = 1 < 1.5434 < \beta_0 + \beta_2 b = 3$ so that condition (7.41) is satisfied. The candidate D-optimal design calculated using the Gauss program in Appendix C is the 6-point design

$$\xi^* = \left\{ \begin{array}{cccccc} (0.810, 0) & (2, 0) & (0, 0.537) & (0, 1.463) & (0.303, 0.205) & (0.727, 0.485) \\ 0.0050 & 0.2491 & 0.1114 & 0.2447 & 0.1420 & 0.2478 \\ -1.3801 & 1 & -1.390 & 1.390 & -1.469 & 1.670 \end{array} \right\}. \quad (7.42)$$

Figure 7.12 (a) indicates the location of the support points of design (7.42) on the design space $[0, 2] \times [0, 2]$, and Figure 7.12 (b) displays the graph of the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$. The graph in Figure 7.12 (b) suggests that design (7.42) is D-optimal.

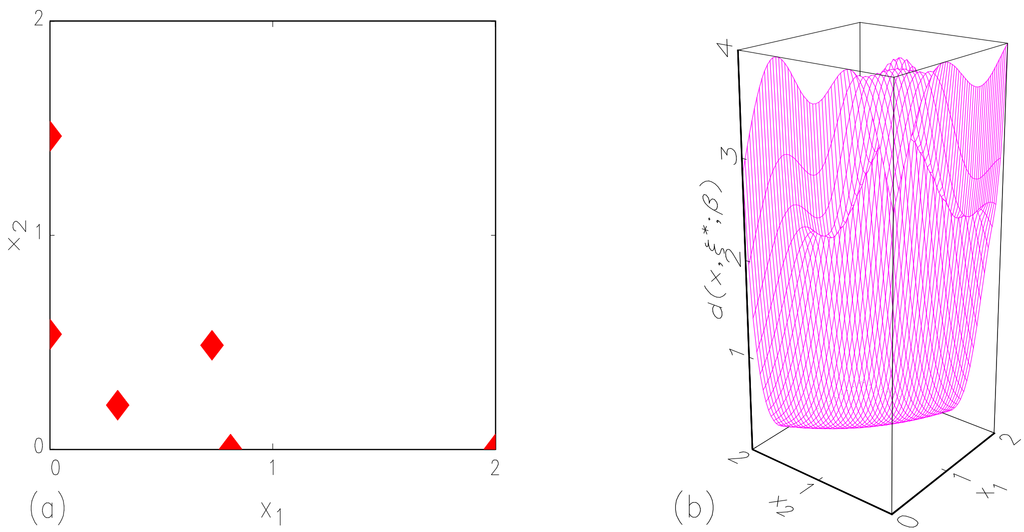


Figure 7.12: (a) Support points and (b) the standardized variance function of the D-optimal design for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, 2] \times [0, 2]$ and with parameters $\beta_0 = -3$, $\beta_1 = 2$, $\beta_2 = 3$ and $\beta_{12} = 5$. The triangles and diamonds are support points of design (7.42).

In Example 7.6, the pairs of support points which would otherwise lie on the same logit and hence be equally weighted in the absence of the restriction on the design space, lie on different logits and are unequally weighted. Depending of the values of the parameter vector

$\beta = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$ and the design space $[0, b] \times [0, b]$, the weights of some support points can be so small such that the 6-point D-optimal design is practically a less than 6-point D-optimal design or degenerates to a less than 6-point D-optimal design. Example 7.6 and the following example illustrates this.

Example 7.7. Consider the two-variable binary logistic model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, 2] \times [0, 2]$ where $\beta_0 = -3$, $\beta_1 = 2$, $\beta_2 = 3$ and $\beta_{12} = 1$. This example is the same as Example 7.6 except that here $\beta_{12} = 1$. In this example, condition (7.41) for expecting a 6-point D-optimal design is also satisfied. However, the candidate D-optimal design calculated using the Gauss program in Appendix C is the 5-point design

$$\xi^* = \left\{ \begin{array}{ccccc} (2,0) & (0,0.532) & (0,1.468) & (0.345,0.245) & (1.048,0.698) \\ 0.2494 & 0.1048 & 0.2460 & 0.1509 & 0.2489 \\ 1 & -1.404 & 1.404 & -1.490 & 1.921 \end{array} \right\}. \qquad (7.43)$$

Figure 7.13 (a) indicates the location of the support points of design (7.43) on the design space $[0, 2] \times [0, 2]$, and Figure 7.13 (b) displays the graph of the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$. The graph of Figure 7.13 (b) suggests that design (7.43) is D-optimal.

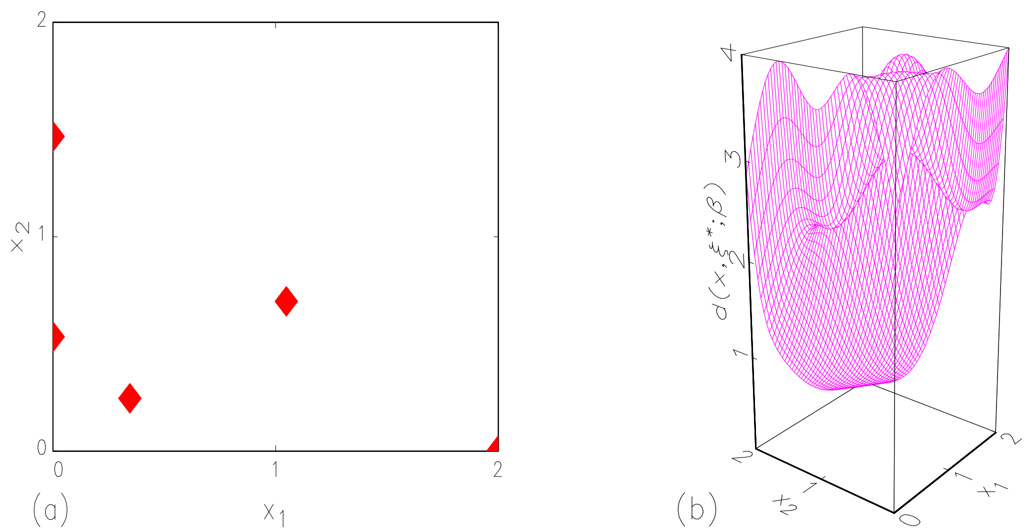


Figure 7.13: (a) Support points and (b) the standardized variance function of the D-optimal design for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, 2] \times [0, 2]$ where $\beta_0 = -3$, $\beta_1 = 2$, $\beta_2 = 3$ and $\beta_{12} = 1$. The triangles and diamonds are support points of design (7.43).

In Example 7.7, a 6-point D-optimal design with 2 support points on each of the x_1-axis,

x_2-axis and the ray $x_2 = \dfrac{2}{3}x_1$ was expected. However, the second support point on the x_1-axis was found to be (0.848,0) with weight 10^{-8} which meant that the 6-point D-optimal was practically the 5-point design (7.43). It should also be noted that the interior support point $(0.345, 0.245)$ of design (7.43) does not exactly lie on the ray $x_2 = \frac{2}{3}x_1$ as expected.

Case when $\beta_0 < -1.5434 < \beta_0 + \beta_2 b \leq 1.5434 < \beta_0 + \beta_1 b$

This case is similar to the case of $\beta_0 < -1.5434 < \beta_0 + \beta_1 b \leq 1.5434 < \beta_0 + \beta_2 b$ (above) except that in this case the respective points $(0, b)$ and (b,0) lie below and above the logit line $u = 1.5434$. Thus, the condition for the candidate 6-point D-optimal design is

$$\beta_0^* < \beta_0 < -u^* \leq \beta_0 + \beta_2 b \leq u^* \leq \beta_0 + \beta_1 b.$$

Furthermore, the construction of the D-optimal designs in this case is as for the case of

$$\beta_0 < -1.5434 < \beta_0 + \beta_1 b \leq 1.5434 < \beta_0 + \beta_2 b$$

above.

Case when $\beta_0 < -1.5434 < [\beta_0 + \beta_2 b, \beta_0 + \beta_1 b] \leq 1.5434$

In this case, the point $(0, 0)$ lies below the logit $u = -1.5434$, the respective points $(b, 0)$ and $(0, b)$ lie on or below the logit line $u = 1.5434$. This condition together with condition $\beta_0^* < \beta_0 < -1.5434$ gives the design condition

$$\beta_0^* < \beta_0 < -1.5434 < [\beta_0 + \beta_1 b, \beta_0 + \beta_2 b] \leq 1.5434. \tag{7.44}$$

The candidate D-optimal design which satisfies condition (7.44) is design (7.36) with respective support points $\left(\frac{u-\beta_0}{\beta_1}, 0\right)$ and $\left(0, \frac{u-\beta_0}{\beta_2}\right)$ replaced by $(b, 0)$ and $(0, b)$. The effect of this change is that the pattern of the other support points in design (7.36) may also be modified, and thus the support points of the resulting candidate D-optimal design may not have equal weights as in the following example.

Example 7.8. Consider the two-variable binary logistic model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, 2] \times [0, 2]$ where $\beta_0 = -3$, $\beta_1 = 2$, $\beta_2 = 2$ and $\beta_{12} = 5$. In this case $\beta_0^* = -3.8 < \beta_0 = -3 < -1.5434 < \beta_0 + \beta_1 b = \beta_0 + \beta_2 b = 1 < 1.5434$. Hence, condition (7.44) is satisfied. Therefore, the vertices $(2, 0)$ and $(0, 2)$ belong to the support points of the candidate D-optimal design. The Gauss program in Appendix C gives the following 6-point design.

$$\xi^* = \left\{ \begin{array}{cccccc} (0.770, 0) & (2, 0) & (0, 0.769) & (0, 2) & (0.273, 0.273) & (0.647, 0.647) \\ 0.0133 & 0.2490 & 0.0135 & 0.2490 & 0.2260 & 0.2492 \\ -1.461 & 1 & -1.462 & 1 & -1.534 & 1.684 \end{array} \right\}. \tag{7.45}$$

Figure 7.14 (a) indicates the location of the support points of design (7.43) on the design space $[0, 2] \times [0, 2]$, and Figure 7.14 (b) displays the graph of the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$. The graph of Figure 7.14 (b) suggests that design (7.45) is D-optimal.

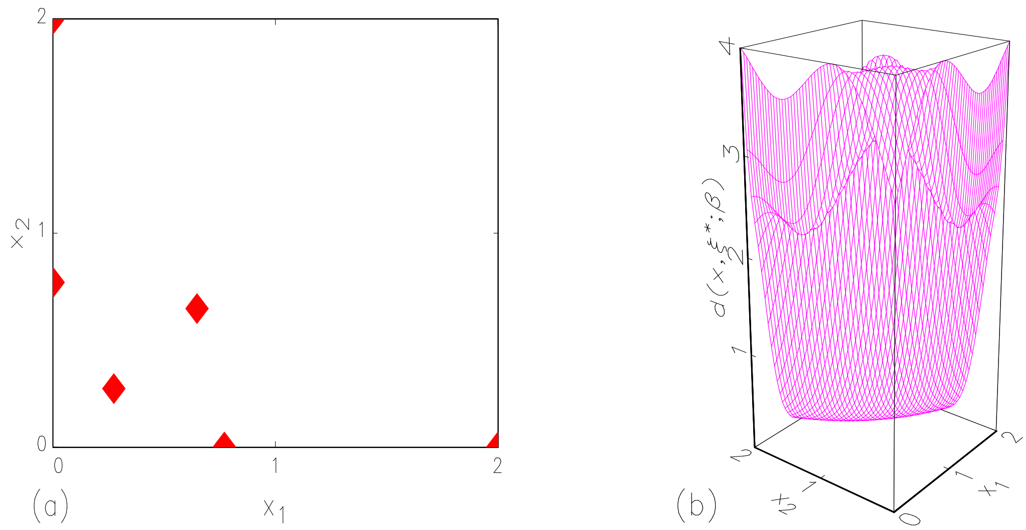


Figure 7.14: (a) Support points and (b) the standardized variance function of the D-optimal design for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, 2] \times [0, 2]$ and with parameters $\beta_0 = -3$, $\beta_1 = 2$, $\beta_2 = 2$ and $\beta_{12} = 5$. The triangles and diamonds are support points of design (7.45).

In Example 7.8, note that the coordinates of each interior support point are equal because $\beta_1 = \beta_2 = 2$ implies the interior ray is $x_2 = \frac{\beta_1}{\beta_2} x_1 = x_1$. That the support points $(0.770, 0)$ and $(0, 0.769)$, the corresponding logits -1.461 and 1.462, and the corresponding weights 0.0133 and 0.0135 are slightly different can be attributed to accumulated small numerical errors during the optimization process. Note that derivatives or gradients used during optimization were calculated numerically by the optimization routines.

The following example conjectures that a 4-point D-optimal design can be obtained when a 6-point D-optimal design is expected as a result of a large difference between the values of the parameters β_1 and β_2.

Example 7.9. Consider the two-variable binary logistic model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, 2] \times [0, 2]$ and with parameter values $\beta_0 = -3$, $\beta_1 = 2$, $\beta_2 = 1.5$ and $\beta_{12} = 5$. This example is the same as Example 7.8 except that here $\beta_2 = 1.5$. In this example, condition (7.44) for expecting a 6-point D-optimal design with $(0, 2)$ and $(2, 0)$

as two of its support points is satisfied. Furthermore, the interior support points are expected to lie on the ray $x_2 = \dfrac{4}{3}x_1$ which is not the axis of symmetry of the design space $[0,2] \times [0,2]$ as was the ray $x_2 = x_1$ in Example 7.8. The D-optimal design in this case was found to be the 4-point design

$$\xi^* = \left\{ \begin{array}{cccc} (2,0) & (0,2) & (0.427, 0.154) & (0.564, 0.805) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 1 & -1.587 & 1.604 \end{array} \right\} \tag{7.46}$$

Figure 7.15 (a) indicates the location of the support points of design (7.46) on the design space $[0,2] \times [0,2]$, and Figure 7.15 (b) displays the graph of the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$. The graph of Figure 7.15 (b) suggests that the design (7.46) is D-optimal.

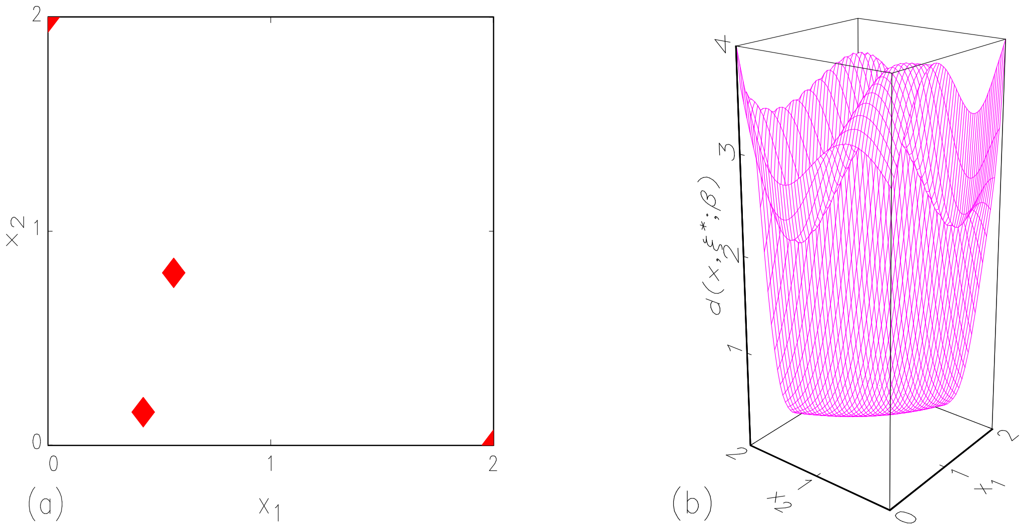


Figure 7.15: (a) Support points of the D-optimal design and (b) the standardized variance function for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0,2] \times [0,2]$ and with parameters $\beta_0 = -3$, $\beta_1 = 2$, $\beta_2 = 1.5$ and $\beta_{12} = 5$. The triangles and diamonds are support points of design (7.43).

In Example 7.9, note that in addition to the reduction of number of support points of the D-optimal design from 6 to 4, none of the interior support points is located on the ray $x_2 = \dfrac{4}{3}x_1$. Hence, it can be conjectured that a complete loss of "balance" was generated by the presence of the support points $(0,2)$ and $(2,0)$ and the asymmetric ray $x_2 = \dfrac{4}{3}x_1$ with respect to design space $[0,2] \times [0,2]$.

7.5. D-optimal designs for antagonism

This section discusses the construction of D-optimal designs for detecting drug interaction using the two variable binary logistic model with interaction

$$u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 \tag{7.47}$$

on the design space $[0, b] \times [0, b]$ where $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_{12} < 0$. As discussed in Chapter 2 this case represents antagonism. It was also found that antagonism can also be observed with the same model when $\beta_0 > 0$, $\beta_1 < 0$, $\beta_2 < 0$ and $\beta_{12} > 0$. However, as it was earlier stated the construction of D-optimal design is to be discussed only for the first case. The construction of D-optimal designs in the second case will be easily deduced from the first case. A brief discussion on the construction of D-optimal designs for antagonism using model (7.47) in the second case will be presented in Section 7.6.

For model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ with $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_{12} < 0$, $\pm u$ are hyperbolae whose centre is $\left(-\frac{\beta_2}{\beta_{12}}, -\frac{\beta_1}{\beta_{12}}\right) \in \mathbb{R}^{2+}$. The logit value at the centre of the hyperbolae $\pm u$ is $\beta_0^* = \beta_0 - \dfrac{\beta_1 \beta_2}{\beta_{12}}$. The location of the centre of the hyperbolae $\pm u$ and the logit value β_0^* relative to the design space $[0, b] \times [0, b]$ determine the patterns of the candidate D-optimal design. This section is subdivided as follows. Sections 7.5.1 discusses the construction of D-optimal designs on the design space $[0, b] \times [0, b]$ when $|\beta_0^*| < 1.5434$ and $-\frac{\beta_2}{\beta_{12}} > b$. Section 7.5.2 discusses the construction of D-optimal designs on the design space $[0, b] \times [0, b]$ when $\beta_0^* > 1.5434$, and $-\frac{\beta_2}{\beta_{12}} > b$ and $-\frac{\beta_1}{\beta_{12}} > b$. Sections 7.5.3 discusses the construction of D-optimal designs of the case when $|\beta_0^*| < 1.5434$, and $-\frac{\beta_2}{\beta_{12}} < b$ and $-\frac{\beta_1}{\beta_{12}} < b$. Section 7.5.4 discusses the construction of D-optimal designs of the case when $|\beta_0^*| < 1.5434$, and either $-\frac{\beta_2}{\beta_{12}} < b$ and $-\frac{\beta_1}{\beta_{12}} > b$, or $-\frac{\beta_2}{\beta_{12}} > b$ and $-\frac{\beta_1}{\beta_{12}} < b$. Section 7.5.5 discusses the construction of D-optimal designs of the case when $\beta_0^* > 1.5434$, and $-\frac{\beta_1}{\beta_{12}} < b$ and $-\frac{\beta_1}{\beta_{12}} < b$. Section 7.5.6 discusses the construction of D-optimal designs of the case when $\beta_0^* < -1.5434$, and $-\frac{\beta_1}{\beta_{12}} < b$ and $-\frac{\beta_1}{\beta_{12}} < b$. Section 7.5.7 discusses the construction of D-optimal designs of the case when $\beta_0^* > 1.5434$, and either $-\frac{\beta_2}{\beta_{12}} < b$ and $-\frac{\beta_1}{\beta_{12}} > b$, or $-\frac{\beta_2}{\beta_{12}} > b$ and $-\frac{\beta_1}{\beta_{12}} < b$. Finally, Section 7.5.8 discusses the construction of D-optimal designs of the case when $\beta_0^* < -1.5434$, and either $-\frac{\beta_2}{\beta_{12}} < b$ and $-\frac{\beta_1}{\beta_{12}} > b$, or $-\frac{\beta_2}{\beta_{12}} > b$ and $-\frac{\beta_1}{\beta_{12}} < b$.

7.5.1 D-optimal designs when $|\beta_0^*| < 1.5434$, and $-\frac{\beta_2}{\beta_{12}} > b$ and $-\frac{\beta_1}{\beta_{12}} > b$

Consider model (7.47) on the design space $[0,b] \times [0,b]$ where $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_{12} < 0$, and where $-1.5434 < \beta_0^* = \beta_0 - \frac{\beta_1\beta_2}{\beta_{12}} < 1.5434$, $-\dfrac{\beta_1}{\beta_{12}} > b$ and $-\dfrac{\beta_2}{\beta_{12}} > b$. For this model, Figure 7.16 (a) indicates that the centre $\left(-\frac{\beta_2}{\beta_{12}}, -\frac{\beta_1}{\beta_{12}}\right)$ of the hyperbolae $\pm u$ is outside the design space $[0,b] \times [0,b]$ and that the branches of hyperbolae $\pm u$ are in 4 different quadrants generated by the asymptotes $x_1 = -\frac{\beta_2}{\beta_{12}}$ an $x_2 = -\frac{\beta_1}{\beta_{12}}$. The support points of the candidate D-optimal design for this case are points A, B, G and J marked by circles in Figure 7.16 (a).

The relationship between the support points (x_1, x_2) in Figure 7.16 and the support points (y_1, y_2) in Figure 7.4 is given by

$$(x_1, x_2) = (-y_2 + b, -y_1 + b). \tag{7.48}$$

This transformation is the composition of the translation from point $(0,0)$ to point (b,b) followed by the reflection with respect to the line passing through point G and perpendicular to line GI. By inspection of Figure 7.16 (a), it is clear that each curve u in Figure 7.16 (a) is the image of curve $\widetilde{u} = -u$ in Figure 7.4 by the transformation (7.48). The relationship between the parameter vector $\boldsymbol{\beta}' = (\beta_0', \beta_1', \beta_2', \beta_{12}')^T$ of the model in terms of (y_1, y_2) and parameter vector $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$ of the model in terms of (x_1, x_2) is derived from the transformation (7.48) as follows.

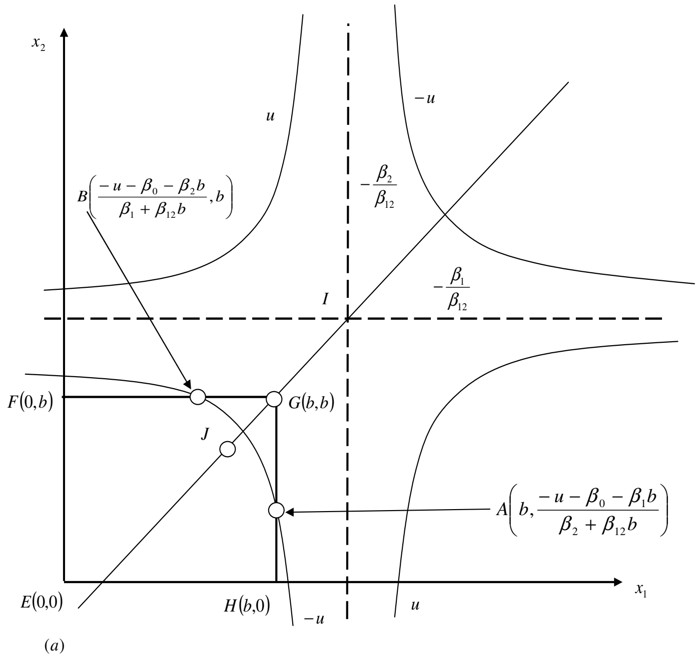
$$\begin{aligned}
\beta_0' + \beta_1' y_1 + \beta_2' y_2 + \beta_{12}' y_1 y_2 &= \widetilde{u} = -u \\
&= -\beta_0 - \beta_1 x_1 - \beta_2 x_2 - \beta_{12} x_1 x_2 \\
&= -\beta_0 - \beta_1(-y_2 + b) - \beta_2(-y_1 + b) - \beta_{12}(-y_2 + b)(-y_1 + b) \\
&= -\beta_0 - \beta_1 b - \beta_2 b - \beta_{12} b^2 + (\beta_2 + \beta_{12} b) y_1 + (\beta_1 + \beta_{12} b) y_2 - \beta_{12} y_1 y_2
\end{aligned}$$

and, hence

$$\left\{\begin{array}{lcl}
\beta_0' &=& -\beta_0 - \beta_1 b - \beta_2 b - \beta_{12} b^2 \\
\beta_1' &=& \beta_2 + \beta_{12} b \\
\beta_2' &=& \beta_1 + \beta_{12} b \\
\beta_{12}' &=& -\beta_{12}
\end{array}\right.
\iff
\left\{\begin{array}{lcl}
\beta_0 &=& -\beta_0' - \beta_1' b - \beta_2' b - \beta_{12}' b^2 \\
\beta_1 &=& \beta_2' + \beta_{12}' b \\
\beta_2 &=& \beta_1' + \beta_{12}' b \\
\beta_{12} &=& -\beta_{12}'
\end{array}\right. \tag{7.49}$$

Then, using the transformations (7.48) and (7.49), the ray $y_2 = \dfrac{\beta_1'}{\beta_2'} y_1$ in the (y_1, y_2) space is transformed to the ray

$$x_2 = \frac{\beta_1 + \beta_{12} b}{\beta_2 + \beta_{12} b} x_1 + \frac{(\beta_2 - \beta_1) b}{\beta_2 + \beta_{12} b} \tag{7.50}$$



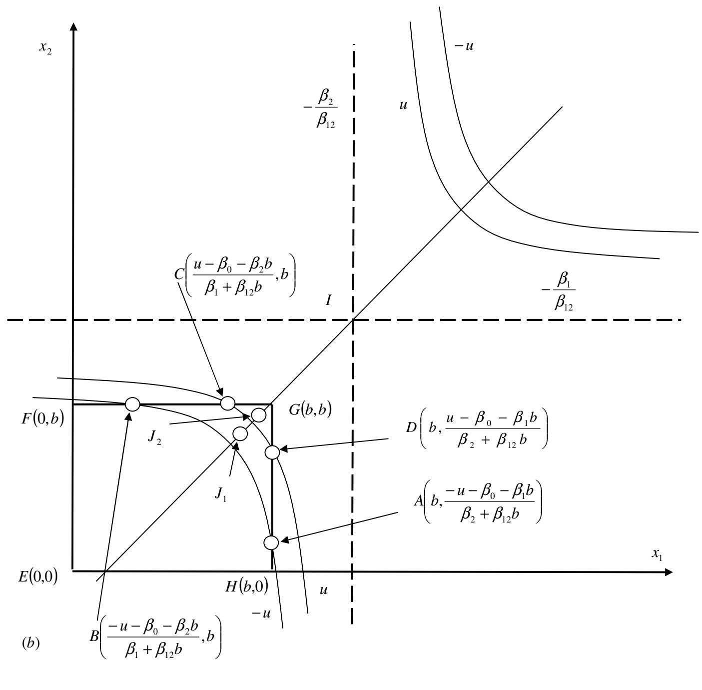


Figure 7.16: Design patterns for the candidate D-optimal design for the two-variable logistic model with interaction $u = \text{logit}(p) = \beta_0+\beta_1 x_1+\beta_2 x_2+\beta_{12}x_1x_2$ on the design space $[0, b]\times[0, b]$ with $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$, $\beta_{12} < 0$, $-\frac{\beta_2}{\beta_{12}} > b$ and $-\frac{\beta_1}{\beta_{12}} > b$ where (a) $|\beta_0^*| < 1.5434$ and (b) $\beta_0^* > 1.5434$. Circles are support points.

in the (x_1, x_2)-space. In summary, the candidate D-optimal design with the design pattern displayed in Figure 7.16 (a) is

$$\xi = \left\{ \begin{array}{c} (x_1, x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{cccc} (b, b) & \left(\frac{-u-\beta_0-\beta_2 b}{\beta_1+\beta_{12} b}, b\right) & \left(b, \frac{-u-\beta_0-\beta_1 b}{\beta_2+\beta_{12} b}\right) & (x_{1J}, x_{2J}) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \beta_0 & -u & -u & u_1 \end{array} \right\} \tag{7.51}$$

and its support points are related to those of design (7.17) by the transformations (7.48) and (7.49). The following numerical example demonstrates the application of transformations (7.48) and (7.49) to deriving design (7.51) from design (7.17).

Example 7.10. Consider the two-variable binary logistic model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, 2] \times [0, 2]$ with parameter values $\beta_0 = -18.5$, $\beta_1 = 7.5$, $\beta_2 = 8$ and $\beta_{12} = -3$. Transforming theses parameter values using the transformation (7.49) gives $\beta_0' = -0.5$, $\beta_1' = 2$, $\beta_2' = 1.5$ and $\beta_{12}' = 3$ which are the values of the parameters of the model in Example 7.1. Hence, it can be conjectured that the D-optimal design for the model in this example can be derived from the D-optimal design (7.22) in Example 7.1 using the transformation (7.48) with $b = 2$ to obtain

$$\xi^* = \left\{ \begin{array}{cccc} (2, 2) & (0.284, 2) & (2, 0.713) & (1.279, 1.459) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0.5 & -2.075 & -2.075 & -2.832 \end{array} \right\}.$$

Note that with the parameters of this example ray (7.50) is $x_2 = 0.75x_1 + 0.5$, and clearly the interior support point $(1.279, 1.459)$ belongs to this ray.

7.5.2 D-optimal designs when $\beta_0^* > 1.5434$, and $-\frac{\beta_2}{\beta_{12}} > b$ and $-\frac{\beta_1}{\beta_{12}} > b$

Consider the design pattern displayed in Figure 7.16 (b) where the support points of the candidate D-optimal design are marked by circles. The candidate D-optimal design ξ corresponding to the design displayed in Figure 7.16 (b) is

$$\xi = \left\{ \begin{array}{cccccc} \left(b, \frac{-u-\beta_0-\beta_2 b}{\beta_1+\beta_{12} b}\right) & \left(\frac{-u-\beta_0-\beta_2 b}{\beta_1+\beta_{12} b}, b\right) & \left(\frac{u-\beta_0-\beta_1 b}{\beta_2 b \beta_{12} b}, b\right) & \left(b, \frac{u-\beta_0-\beta_1 b}{\beta_2+\beta_{12} b}\right) & (x_{1J_1}, x_{2J_1}) & (x_{1J_2}, x_{2J_2}) \\ \frac{\lambda_1}{2} & \frac{\lambda_1}{2} & \frac{\lambda_2}{2} & \frac{\lambda_2}{2} & \lambda_3 & \lambda_4 \\ -u & -u & u & u & u_1 & u_2 \end{array} \right\} \tag{7.52}$$

where $\lambda_4 = 1 - \lambda_1 - \lambda_2 - \lambda_3$.

Clearly, the design pattern in Figure 7.16 (b) is the reflection of the design pattern in Figure 7.9 (b) with respect to line FH. Hence, the support points (x_1, x_2) in Figure 7.16 (b) can be derived from the support points (y_1, y_2) in Figure 7.9 (b) using the relationship

$$(x_1, x_2) = (-y_1 + b, -y_2 + b). \tag{7.53}$$

In addition, logit line u in the design of Figure 7.16 (b) corresponds to a logit line $\widetilde{u} = -u$ in the design of Figure 7.9 (b). That is, the candidate D-optimal design with the design pattern in Figure 7.16 (b) can be derived from the design (7.36) using (7.53). As in Section 7.5.1, the relationship between the parameter vectors $\boldsymbol{\beta}' = (\beta_0', \beta_1', \beta_2', \beta_{12}')^T$ in the design displayed in Figure 7.9 (b) and $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$ in the design displayed in Figure 7.16 (b) can be calculated as follows.

$$\begin{aligned}
\beta_0' + \beta_1' y_1 + \beta_2' y_2 + \beta_{12}' y_1 y_2 &= \widetilde{u} = -u \\
&= -\beta_0 - \beta_1 x_1 - \beta_2 x_2 - \beta_{12} x_1 x_2 \\
&= -\beta_0 - \beta_1(-y_1 + b) - \beta_2(-y_2 + b) - \beta_{12}(-y_1 + b)(-y_2 + b) \\
&= -\beta_0 - \beta_1 b - \beta_2 b - \beta_{12} b^2 + (\beta_1 + \beta_{12} b) y_1 + (\beta_2 + \beta_{12} b) y_2 - \beta_{12} y_1 y_2
\end{aligned}$$

so that

$$\left\{ \begin{array}{lcl}
\beta_0' & = & -\beta_0 - \beta_1 b - \beta_2 b - \beta_{12} b^2 \\
\beta_1' & = & \beta_1 + \beta_{12} b \\
\beta_2' & = & \beta_2 + \beta_{12} b \\
\beta_{12}' & = & -\beta_{12}
\end{array} \right. \iff \left\{ \begin{array}{lcl}
\beta_0 & = & -\beta_0' - \beta_1' b - \beta_2' b - \beta_{12}' b^2 \\
\beta_1 & = & \beta_1' + \beta_{12}' b \\
\beta_2 & = & \beta_2' + \beta_{12}' b \\
\beta_{12} & = & -\beta_{12}'
\end{array} \right. \tag{7.54}$$

Furthermore, the equation of the ray passing through points G and I in Figure 7.16 (b) is given by (7.50).

Example 7.11. Consider the two-variable binary logistic model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, 2] \times [0, 2]$ and with parameter values $\beta_0 = -31$, $\beta_1 = 14$, $\beta_2 = 13$ and $\beta_{12} = -5$. Transforming these parameter values using the transformation (7.54) gives $\beta_0' = -3$, $\beta_1' = 4$, $\beta_2' = 3$ and $\beta_{12}' = 5$ which are the values of the parameters of the model in Example 7.5. Hence, the D-optimal design for the model in this example can be derived from the D-optimal design (7.40) in Example 7.5 using the transformation (7.53) with $b = 2$ to obtain the 6-point D-optimal design

$$\xi^* = \left\{ \begin{array}{cccccc}
(1.583, 2) & (0.917, 2) & (2, 1.444) & (2, 0.556) & (1.828, 1.771) & (1.566, 1.421) \\
0.0895 & 0.2429 & 0.0895 & 0.2429 & 0.0870 & 0.2482 \\
1.333 & -1.333 & 1.333 & -1.333 & 1.431 & -1.734
\end{array} \right\}.$$

For the parameter values of this example, ray (7.50) is $x_2 = \dfrac{4}{3} x_1 - \dfrac{2}{3}$, and clearly the two interior support points $(1.566, 1.421)$ and $(1.828, 1.771)$ of the above design belong to this ray.

7.5.3 D-optimal designs when $|\beta_0^*| < 1.5434$, and $-\frac{\beta_2}{\beta_{12}} < b$ and $-\frac{\beta_1}{\beta_{12}} < b$

In Section 7.5.1, Figure 7.16 (a) was used to show that if $|\beta_0^*| < 1.5434$ then the branches of the hyperbolae $\pm u = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ are spread in the 4 quadrants generated by the asymptotes $x_1 = -\dfrac{\beta_2}{\beta_{12}}$ and $x_2 = -\dfrac{\beta_1}{\beta_{12}}$. If $0 < -\dfrac{\beta_2}{\beta_{12}} < b$ and $0 < -\dfrac{\beta_1}{\beta_{12}} < b$, then the centre of the hyperbolae $\pm u$ is inside the design space $[0, b] \times [0, b]$. Consequently, the branches of the hyperbolae $\pm u$ may intersect the boundaries of the design space at between 4 and 8 points. Thus, the number of support points of the candidate D-optimal design is expected to vary from 4 to 8. In this section, a candidate D-optimal design with a simple 8-point design pattern is used to demonstrate the analytical construction of the D-optimal designs in this case. Then, the existence conditions for more complicated design patterns are derived, and numerical examples are used to illustrate the construction of the corresponding D-optimal designs.

4 and 8 point D-optimal designs with a simple design patterns

Consider the design pattern displayed in Figure 7.17 where the design points are represented by circles. The conditions for the 8-point design pattern in the figure are that $0 < u \leq -\beta_0$, $\beta_0^* = \beta_0 - \frac{\beta_1\beta_2}{\beta_{12}} = 0$, and that the centre of the hyperbolae $\pm u$ is $\left(\dfrac{b}{2}, \dfrac{b}{2}\right)$ which is also the centre of the design space $[0, b] \times [0, b]$. These conditions imply that $\beta_1 = \beta_2 = -\frac{2\beta_0}{b}$ and $\beta_{12} = \frac{4\beta_0}{b^2}$ with $\beta_0 < 0$. Note that if $0 < u \leq -\beta_0$ and the centre of the hyperbolae $\pm u$ is $\left(\dfrac{b}{2}, \dfrac{b}{2}\right)$, then the 8-point design pattern becomes a 4-point design with support points $(0, 0)$, $(0, b)$, $(b, 0)$ and (b, b).

Next, consider the construction of the candidate 8-point D-optimal design, ξ_8, with the 8-point design pattern discussed above. The support points of the design are equidistant from the common centres of both hyperbolae $\pm u$ and the design space $[0, b] \times [0, b]$. Hence, the support points of the candidate 8-point D-optimal design ξ_8 can be assumed to be equally weighted.

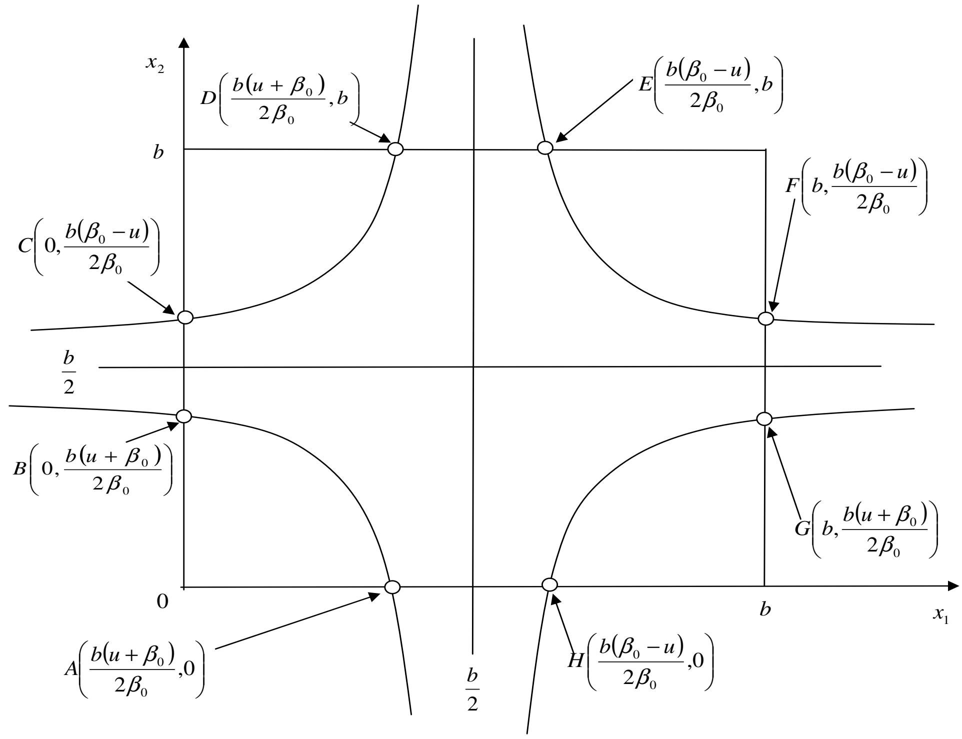


Figure 7.17: Support points of an 8-point D-optimal design for the two variable binary logistic model $u = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, b] \times [0, b]$ with $\beta_1 = \beta_2 = -\frac{2\beta_0}{b}$, $\beta_{12} = \frac{4\beta_0}{b^2}$, and $\beta_0 < 0$. Circles symbolize support points.

The design matrix associated with the equally weighted 8-point D-optimal design ξ_8 is

$$\boldsymbol{X} = \begin{bmatrix} 1 & \frac{(\beta_0+u)b}{2\beta_0} & 0 & 0 \\ 1 & 0 & \frac{(\beta_0+u)b}{2\beta_0} & 0 \\ 1 & 0 & \frac{(\beta_0-u)b}{2\beta_0} & 0 \\ 1 & \frac{(\beta_0+u)b}{2\beta_0} & b & \frac{(\beta_0+u)b^2}{2\beta_0} \\ 1 & \frac{(\beta_0-u)b}{2\beta_0} & b & \frac{(\beta_0-u)b^2}{2\beta_0} \\ 1 & b & \frac{(\beta_0-u)b}{2\beta_0} & \frac{(\beta_0-u)b^2}{2\beta_0} \\ 1 & b & \frac{(\beta_0+u)b}{2\beta_0} & \frac{(\beta_0+u)b^2}{2\beta_0} \\ 1 & \frac{(\beta_0-u)b}{2\beta_0} & 0 & 0 \end{bmatrix}, \tag{7.55}$$

where $0 < u \leq -\beta_0$, and the matrix of model weights is $\boldsymbol{W} = \frac{1}{8}\frac{e^u}{(1+e^u)^2} I_8$ where I_8 is the 8×8 identity matrix. Hence, the information matrix for the vector of parameters $\boldsymbol{\beta} =$

$(\beta_0, \beta_1, \beta_2, \beta_{12})^T$ is given by

$$M(\xi_8;\boldsymbol{\beta}) = \boldsymbol{X}^T\boldsymbol{W}\boldsymbol{X} = \frac{e^u}{(1+e^u)^2}\begin{bmatrix} 1 & \frac{b}{2} & \frac{b}{2} & \frac{b^2}{4} \\ \frac{b}{2} & \frac{b^2(3\beta_0^2+u^2)}{8\beta_0^2} & \frac{b^2}{4} & \frac{b^3(3\beta_0^2+u^2)}{16\beta_0^2} \\ \frac{b}{2} & \frac{b^2}{4} & \frac{b^2(3\beta_0^2+u^2)}{8\beta_0^2} & \frac{b^3(3\beta_0^2+u^2)}{16\beta_0^2} \\ \frac{b^2}{4} & \frac{b^3(3\beta_0^2+u^2)}{16\beta_0^2} & \frac{b^3(3\beta_0^2+u^2)}{16\beta_0^2} & \frac{b^4(\beta_0^2+u^2)}{8\beta_0^2} \end{bmatrix} \tag{7.56}$$

and its determinant is

$$D_1 = \frac{b^8u^2e^{4u}(u^2+\beta_0^2)^2}{1024\beta_0^6(1+e^u)^8}$$

which is proportional to

$$D = \frac{u^2e^{4u}(u^2+\beta_0^2)^2}{(1+e^u)^8}. \tag{7.57}$$

The derivative of $\ln D$, in (7.57), with respect to u is

$$\frac{d\ln D}{du} = \frac{2}{u} + 4 + \frac{4u}{u^2+\beta_0^2} - \frac{8e^u}{1+e^u} = \frac{2\left[\beta_0^2(1+e^u+2u-2ue^u)+u^2(3+3e^u+2u-2ue^u)\right]}{u(u^2+\beta_0^2)(1+e^u)} \tag{7.58}$$

and the second derivative of $\ln D$ is

$$\frac{d^2\ln D}{du^2} = \frac{-2(\beta_0^4+3u^4)}{u^2(u^2+\beta_0^2)^2} - \frac{8e^u}{(1+e^u)^2} < 0. \tag{7.59}$$

This means that equation $\frac{d\ln D}{du} = 0$ has a unique solution, u^*, for u which depends on β_0, and that D takes on its maximum at this unique solution for u. Note that solving for u in equation (7.58) is equivalent to solving for u in equation

$$\beta_0^2(1+e^u+2u-2ue^u)+u^2(3+3e^u+2u-2ue^u) = 0. \tag{7.60}$$

Furthermore, recall that the candidate 8-point D-optimal design discussed above becomes the 4-point D-optimal design if $-\beta_0 \geq u > 0$. This means that the branches of the hyperbolae $\pm u$ pass on or beyond the vertices of the design space $[0,b]\times[0,b]$. Thus, in the case of a 4-point design the optimal logit $u > 0$ is $u = -\beta_0$. The largest optimal value of u in the case of a candidate 4-point D-optimal design with support points $(0,0)$, $(0,b)$, $(b,0)$ and (b,b) is obtained by setting $u = -\beta_0$ in equation (7.60) and then numerically solving the equation for u to obtain $u = 1.5434$ as the unique solution. It follows that the candidate 8-point D-optimal design is possible if $-\infty < \beta_0 < -1.5434$, and that the candidate 4-point D-optimal design is possible if $-1.5434 \leq \beta_0 \leq 0$. Dividing the two sides of equation (7.60) by β_0^2, then setting β_0 equal to a very large number, in absolute value, in equation (7.60) and then numerically solving the equation for u, gives $u = 1.04363$. Table 7.1 provides the values of optimal $u = u^*$ for selected values of β_0 when $\beta_0 \leq -1.5434$. The results in Table 7.1 suggests that as β_0

Table 7.1: Relationship between β_0 and D-optimal values of u for the two-variable binary logistic model with interaction when $\beta_0 \leq -1.5434$.

β_0	u^*	β_0	u^*	β_0	u^*	β_0	u^*
-1000	1.0436	-7.5	1.0659	-5	1.0947	-3	1.19146
-500	1.0436	-7	1.0693	-4.75	1.1004	-2.75	1.2208
-50	1.0441	-6.5	1.0735	-4.5	1.1071	-2.5	1.25921
-10	1.0561	-6	1.0788	-4.25	1.1151	-2.25	1.3098
-9.5	1.0575	-5.5	1.0856	-4	1.1247	-2	1.37630
-9	1.0591	-5.75	1.0820	-3.75	1.1364	-1.75	1.4610
-8.5	1.0609	-5.5	1.0856	-3.5	1.1507	-1.6	1.51992
-8	1.0632	-5.25	1.0898	-3.25	1.1687	-1.5434	1.5434

increases from $-\infty$ to -1.5434, the optimal of u increase from 1.04363 to 1.5434. Hence, it can be conjectured that for the candidate 8-point D-optimal design discussed above, the optimal value of u is such that $1.04363 < u < 1.5434$. What remains to be proved is the D-optimality of the designs. The following theorem provides the proofs.

Theorem 7.1. Consider the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, b] \times [0, b]$ where $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 < 0$ and $\beta_{12} < 0$. Suppose that the $\beta_0^* = \beta_0 - \frac{\beta_1 \beta_2}{\beta_{12}} = 0$ and that the centre of the hyperbolae $\pm u$ is $\left(\dfrac{b}{2}, \dfrac{b}{2}\right)$, or, equivalently that $\beta_1 = \beta_2 = -\frac{2\beta_0}{b}$ and $\beta_{12} = \frac{4\beta_0}{b^2}$ with $\beta_0 < 0$. Then the following hold.

1. If $\beta_0 < -1.5434$, then the D-optimal design, ξ_8^*, for the model is an equally weighted 8-point design with support points $\left(\frac{(\beta_0+u^*)b}{2\beta_0}, 0\right)$, $\left(0, \frac{(\beta_0+u^*)b}{2\beta_0}\right)$, $\left(0, \frac{(\beta_0-u^*)b}{2\beta_0}\right)$, $\left(\frac{(\beta_0+u^*)b}{2\beta_0}, b\right)$, $\left(\frac{(\beta_0-u^*)b}{2\beta_0}, b\right)$, $\left(b, \frac{(\beta_0-u^*)b}{2\beta_0}\right)$, $\left(b, \frac{(\beta_0+u^*)b}{2\beta_0}\right)$ and $\left(\frac{(\beta_0-u^*)b}{2\beta_0}, 0\right)$ where u^* is the solution for u to the equation
$$\beta_0^2(1 + e^u + 2u - 2ue^u) + u^2(3 + 3e^u + 2u - 2ue^u) = 0.$$

2. If $-1.5434 \leq \beta_0 \leq 0$, then the D-optimal design for the model is the equally weighed

4-point design

$$\xi_4^* = \left\{ \begin{array}{c} (x_1,x_2) \\ \lambda \\ u \end{array} \right\} = \left\{ \begin{array}{cccc} (0,0) & (0,b) & (b,b) & (b,0) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \beta_0 & -\beta_0 & \beta_0 & -\beta_0 \end{array} \right\}.$$

Proof

1. Let $M(\boldsymbol{x};\boldsymbol{\beta})$ and $M(\xi_8^*;\boldsymbol{\beta})$ be the respective information matrices for $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$ evaluated at a single point $\boldsymbol{x} = (x_1, x_2)^T \in [0, b] \times [0, b]$ and at design ξ_8^*. It has to be shown that the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) = \mathrm{tr}[M(\boldsymbol{x};\boldsymbol{\beta})M^{-1}(\xi^*, \boldsymbol{\beta})] \leq 4$ with equality holding at the support points of ξ_8^*.

Let u be the logit at point $(x_1, x_2) \in [0, b] \times [0, b]$. Then, the conditions $\beta_1 = \beta_2 = -\frac{2\beta_0}{b}$ and $\beta_{12} = \frac{4\beta_0}{b^2}$ imply that $u = \mathrm{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 = \beta_0 - \frac{2\beta_0}{b} x_1 - \frac{2\beta_0}{b} x_2 + \frac{4\beta_0}{b^2} x_1 x_2 \in [\beta_0, -\beta_0]$. Using the simpler of the expression of u above, the standardized variance function $d(\boldsymbol{x}, \xi_8^*; \boldsymbol{\beta})$ can be written as

$$d(\boldsymbol{x}, \xi_8^*; \boldsymbol{\beta}) = \frac{e^{u-u^*}(1+e^{u^*})^2 \left\{ b^4 u^{*2}(3u^2 + u^{*2}) + b^4 \beta_0^2 (u^2 + 3u^{*2}) - 32\beta_0^2 u^{*2} x_1 x_2 (x_1 - b)(x_2 - b) \right\}}{b^4 (1+e^u)^2 u^{*2} (\beta_0^2 + u^{*2})} \tag{7.61}$$

$$\leq g(u) = \frac{e^{u-u^*}(1+e^{u^*})^2 \left[u^{*2}(3u^2 + u^{*2}) + \beta_0^2 (u^2 + 3u^{*2}) \right]}{u^{*2}(u^{*2} + \beta_0^2)(1+e^u)^2}. \tag{7.62}$$

The inequality (7.62) follows from the fact that $x_1 x_2 (x_1 - b)(x_2 - b) \geq 0$ on $[0, b] \times [0, b]$. Clearly, at the support points of ξ_8^*, $d(\boldsymbol{x}, \xi_8^*; \boldsymbol{\beta}) = g(u^*) = 4$ as required at the support points of a D-optimal design. It remains to be shown that $g(u) \leq 4$ for all $u \in [\beta_0, -\beta_0]$ where $\beta_0 < -1.5434$. The function $g(u)$ is even on $[\beta_0, -\beta_0]$ since $g(-u) = g(u)$. Hence, consider showing that $g(u) < 4$ for all $u \in [0, -\beta_0]$ since it will then follow that $g(u) < 4$ for all $u \in [\beta_0, 0]$.

The derivative of $g(u)$ with respect to u is $g'(u) = \dfrac{dg(u)}{du} = \dfrac{e^{u-u^*}(1+e^{u^*})^2 h(u)}{u^{*2}(u^{*2} + \beta_0^2)(1+e^u)^3}$ where

$$h(u) = \beta_0^2(2u + 2ue^u + u^2 - u^2 e^u + 3u^{*2} - 3u^{*2}e^u) + u^{*2}(6u + 6ue^u + 3u^2 - 3u^2 e^u + u^{*2} - u^{*2}e^u). \tag{7.63}$$

Thus, the solutions for u to the equation $g'(u) = 0$ are the same as the solutions for u to the equation $h(u) = 0$. Now, $h(0) = 0$ and

$$h(u^*) = 2u^*[\beta_0^2(1 + e^{u^*} + 2u^* - 2u^* e^{u^*}) + u^{*2}(3 + 3e^{u^*} + 2u^* - 2u^* e^{u^*})] = 0$$

since u^* is solution for u to the equation (7.60). This means that $u = 0$ and $u = u^*$ are at least two stationary points of $g(u)$. However, $u = 0$ corresponds to a minimum for the following reasons. The second derivative of $g(u)$ evaluated at $u = 0$ is

$$g''(0) = \frac{e^{-u^*}(1+e^{u^*})^2\left[u^{*2}(12-u^{*2})+\beta_0^2(4-3u^{*2})\right]}{8u^{*2}(\beta_0^2+u^{*2})} = \frac{e^{-u^*}(1+e^{u^*})^2 f(u^*)}{8u^{*2}(\beta_0^2+u^{*2})} \tag{7.64}$$

where

$$f(u^*) = u^{*2}(12-u^{*2}) + \beta_0^2(4-3u^{*2}) \tag{7.65}$$

has the same sign as $g''(0)$. The fact that u^* is solution for u to the equation (7.60) implies that

$$\beta_0^2 = \frac{u^{*2}(2u^*e^{u^*} - 3 - 2u^* - 3e^{u^*})}{1 + 2u^* + e^{u^*} - 2u^*e^{u^*}},$$

and hence $f(u^*)$ can be written as

$$f(u^*) = \frac{4u^{*3}\left[(u^{*2}-2u^*+4)e^{u^*} - (u^{*2}+2u^*+4)\right]}{(2u^*-1)e^{u^*} - (2u^*+1)} = \frac{f_1(u^*)}{f_2(u^*)} > 0 \tag{7.66}$$

where $f_1(u^*) = (u^{*2}-2u^*+4)e^{u^*} - (u^{*2}+2u^*+4) > 0$ and $f_2(u^*) = (2u^*-1)e^{u^*} - (2u^*+1) > 0$. The functions $f_1(u^*)$ and $f_2(u^*)$ are positive on $[1.04363, 1.5434)$ for the following reasons:
(1) The fact that u^* satisfies $\beta_0^2(1 + 2u^* + e^{u^*} - 2u^*e^{u^*}) + u^{*2}(3 + 3e^{u^*} + 2u^* - 2u^*e^{u^*}) = 0$ implies that $f_2(u^*) = (2u^*-1)e^{u^*} - (2u^*+1) = \dfrac{2u^{*2}(e^{u^*}+1)}{u^{*2}+\beta_0^2} > 0$.
(2) The derivative of $f_1(u^*)$ with respect to u^* is $f_1'(u^*) = 2(e^{u^*}-1) + u^*(u^*e^{u^*} - 2) > 0$ for all $u^* \in [1.04363, 1.5434]$, i.e. $f_1(u^*)$ is strictly increasing on $[1.04363, 1.5434)$ and hence its smallest value is $f_1(1.04363) = 1.348 > 0$. Thus, the inequality (7.66) holds true which implies that $g''(0) > 0$, and hence $u = 0$ corresponds to a minimum of $g(u)$. It remains to be shown that $u = u^*$ is the only stationary point at which $g(u)$ takes on its maximum value on $[0, -\beta_0]$. This is equivalent to showing that $h(u) = 0$ given by (7.63) posses roots 0 and u^* only on $[0, -\beta_0]$ or has only one stationary point on $[0, -\beta_0]$.

The continuity of $h(u)$ and the fact that $h(0) = h(u^*) = 0$ imply that $h(u)$ has at least one stationary point on $[0, u^*]$. Consider the behavior of $h(u)$ on $[0, \infty]$. The first derivative of $h(u)$ with respect to u is

$$h'(u) = s_1(u) - s_2(u) \tag{7.67}$$

where $s_1(u) = 2(1+u)(\beta_0^2 + 3u^{*2})$ is the equation of a straight line with intercept $2(\beta_0^2 + 3u^{*2})$ and slope $2(\beta_0^2 + 3u^{*2})$; and $s_2(u) = (u^{*4} + 3u^2u^{*2} + 3\beta_0^2u^{*2} + \beta_0^2u^2 - 2\beta_0^2 - 6u^{*2})e^u$ is a strictly convex function on $[0, \infty)$ since the second derivative of $s_2(u)$ with respect to u is $s_2''(u) = [u^{*2}(12u + 3u^2 + u^{*2}) + \beta_0^2(4u + u^2 + 3u^{*2})]e^u > 0$. At $u = 0$, $h'(0) = s_1(0) - s_2(0) = \beta_0^2(4 - 3u^{*2}) + u^{*2}(12 - u^{*2}) = f(u^*)$ given by (7.66) which was shown to be positive on $(1.04363, 1.5434)$. This means that $u = 0$ is not a stationary point of $h(u)$, and $s_1(u) > s_2(u)$

at $u = 0$. At $u = u^*$,

$$h'(u^*) = s_1(u^*) - s_2(u^*) = 2\left[\beta_0^2(1 + u^* + e^{u^*} - 2u^{*2}e^{u^*}) + u^*(3 + 3e^{u^*} + 3u^* - 2u^{*2}e^{u^*})\right] < 0$$

for the following reasons. The fact that u^* is the solution for u to the equation (7.60) implies that

$$\beta_0^2(1 + u^* + e^{u^*}) + u^{*2}(3 + 3e^{u^*}) = -\beta_0^2 u^{*2} + 2\beta_0^2 u^* e^{u^*} - 2u^{*3} + 2u^{*3}e^{u^*}.$$

Hence,

$$\begin{aligned} h'(u^*) = s_1(u^*) - s_2(u^*) &= 2\left[\beta_0^2(1 + u^* + e^{u^*}) + u^{*2}(3 + 3e^{u^*}) - 2\beta_0^2 u^{*2}e^{u^*} + 3u^{*3} - 2u^{*4}e^{u^*}\right] \\ &= 2\left[-\beta_0^2 u^{*2} + 2\beta_0^2 u^* e^{u^*} - 2u^{*3} + 2u^{*3}e^{u^*} - 2\beta_0^2 u^{*2}e^{u^*} + 3u^{*3} - 2u^{*4}e^{u^*}\right] \\ &= 2\left[2(1 - u^*)u^*(\beta_0^2 + u^{*2})e^{u^*} + u^*(u^{*2} - \beta_0^2)\right] < 0 \end{aligned}$$

because $u^* > 1.04363 > 1$ and $u^{*2} < \beta_0^2$. This means that $u = u^*$ is not a stationary point of $h(u)$, and $s_1(u) < s_2(u)$ at $u = u^*$. The facts that $s_1(u)$ is a straight line with $s_1(0) > s_2(0)$, $s_2(u)$ is convex on $[0, \infty)$ and $s_2(u^*) > s_1(u^*)$ imply that curves $s_1(u)$ and $s_2(u)$ intersect only once on $[0, \infty]$ and the point of intersection, which is clearly a maximum, is on $[0, u^*]$. Hence, $h'(u) = 0$ has a unique solution for u on $[0, \infty)$ which implies that $h(u)$ has a unique stationary point on $[0, \infty)$.

In conclusion, $g'(u) = 0$ has the unique solution $u = u^*$ on $[0, \infty)$ which implies that on $[0, -\beta_0]$, g(u) attains a global maximum of 4 at $u = u^*$ as was to be proved. The symmetry of g(u) about $u = 0$ implies that $[\beta_0, 0]$, g(u) attains a global maximum of 4 at $u^* = -u^*$ which completes the proof.

2. The proof for D-optimality of the candidate 4-point D-optimal design ξ_4^* given in Theorem 7.1 is deduced from the D-optimality of the 8-point design ξ_8^* as follows. Setting $u* = \beta_0$ implies that 8-point D-optimal design ξ_8^* degenerates to the 4-point design ξ_4^*. Thus, design ξ_4^* is the same as design ξ_8^* when pairs of support points on the same hyperbolic branch are taken at each of the 4 vertices of the design space $[0, b] \times [0, b]$. The fact that the 8-point design ξ_8^* has been shown to be D-optimal implies that the 4-point design ξ_4^* is also D-optimal. ■

Remark

Consider Figure 7.17. Let $\xi_4^{(1)}$ be the equally weighted 4-point design with support points A, C, E and G, and let $\xi_4^{(2)}$ be the equally weighted 4-point design with support points B, D, F and H. If $\beta_0 < -1.5434$, $\beta_1 = \beta_2 = -\frac{2\beta_0}{b}$ and $\beta_{12} = \frac{4\beta_0}{b^2}$, then the equally weighted 8-point D-optimal design with support points A , B, C, D, E, F, G, H is $\xi_8 = \frac{1}{2}\left(\xi_4^{(1)} + \xi_4^{(2)}\right)$.

Furthermore, the designs $\xi_4^{(1)}$, $\xi_4^{(2)}$ and ξ_8 are equally efficient as shown below. The respective design matrices associated with designs $\xi_4^{(1)}$ and $\xi_4^{(2)}$ are

$$\boldsymbol{X}_4^{(1)} = \begin{bmatrix} 1 & \frac{(\beta_0+u)b}{2\beta_0} & 0 & 0 \\ 1 & 0 & \frac{(\beta_0-u)b}{2\beta_0} & 0 \\ 1 & \frac{(\beta_0-u)b}{2\beta_0} & b & \frac{(\beta_0-u)b^2}{2\beta_0} \\ 1 & b & \frac{(\beta_0+u)b}{2\beta_0} & \frac{(\beta_0+u)b^2}{2\beta_0} \end{bmatrix} \tag{7.68}$$

and

$$\boldsymbol{X}_4^{(2)} = \begin{bmatrix} 1 & 0 & \frac{(\beta_0+u)b}{2\beta_0} & 0 \\ 1 & \frac{(\beta_0+u)b}{2\beta_0} & b & \frac{(\beta_0+u)b^2}{2\beta_0} \\ 1 & b & \frac{(\beta_0-u)b}{2\beta_0} & \frac{(\beta_0-u)b^2}{2\beta_0} \\ 1 & \frac{(\beta_0-u)b}{2\beta_0} & 0 & 0 \end{bmatrix}, \tag{7.69}$$

where $0 < u \leq -\beta_0$, and for both designs the matrix of model weights is $\boldsymbol{W} = \dfrac{1}{4}\dfrac{e^u}{(1+e^u)^2} I_4$ where I_4 is a 4×4 identity matrix. The respective information matrices for $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$ evaluated at designs $\xi_4^{(1)}$ and $\xi_4^{(2)}$ are $M(\xi_4^{(1)}; \boldsymbol{\beta})$ and $M(\xi_4^{(2)}; \boldsymbol{\beta})$ given by

$$M(\xi_4^{(1)}; \boldsymbol{\beta}) = M(\xi_4^{(2)}; \boldsymbol{\beta}) = \frac{e^u}{(1+e^u)^2} \begin{bmatrix} 1 & \frac{b}{2} & \frac{b}{2} & \frac{b^2}{4} \\ \frac{b}{2} & \frac{b^2(3\beta_0^2+u^2)}{8\beta_0^2} & \frac{b^2}{4} & \frac{b^3(3\beta_0^2+u^2)}{16\beta_0^2} \\ \frac{b}{2} & \frac{b^2}{4} & \frac{b^3(3\beta_0^2+u^2)}{8\beta_0^2} & \frac{b^3(3\beta_0^2+u^2)}{16\beta_0^2} \\ \frac{b^2}{4} & \frac{b^3(3\beta_0^2+u^2)}{16\beta_0^2} & \frac{b^3(3\beta_0^2+u^2)}{16\beta_0^2} & \frac{b^4(\beta_0^2+u^2)}{8\beta_0^2} \end{bmatrix}. \tag{7.70}$$

Clearly, the information matrix (7.70) is equal to the information matrix (7.56). Therefore

$$|M(\xi_4^{(1)}; \boldsymbol{\beta})| = |M(\xi_4^{(2)}; \boldsymbol{\beta})| = |M(\xi_8; \boldsymbol{\beta})|$$

which implies that designs $\xi_4^{(1)}$, $\xi_4^{(2)}$ and ξ_8 are equally efficient.

<u>Regression approximation of the 8-point D-optimal design</u>

Table 7.1 displays numerical solutions for u to the equation (7.60) for various values of $\beta_0 \leq -1.5434$. The data in Table 7.1 were used to determine a regression equation that can be used to determine optimal value, u^*, of u from a given β_0. For example, for $-5 \leq \beta_0 \leq -1.5434$, the equation was found to be

$$u^{**} = \widehat{u} = 1.0835 + 2.1598\exp(\beta_0) \tag{7.71}$$

with $R^2 = 0.999$. Figure 7.18 also indicates that equation (7.71) fits the data in Table 7.1 reasonably well. The following examples demonstrate numerically the application of equation (7.71) as well as verify Theorem 7.1.

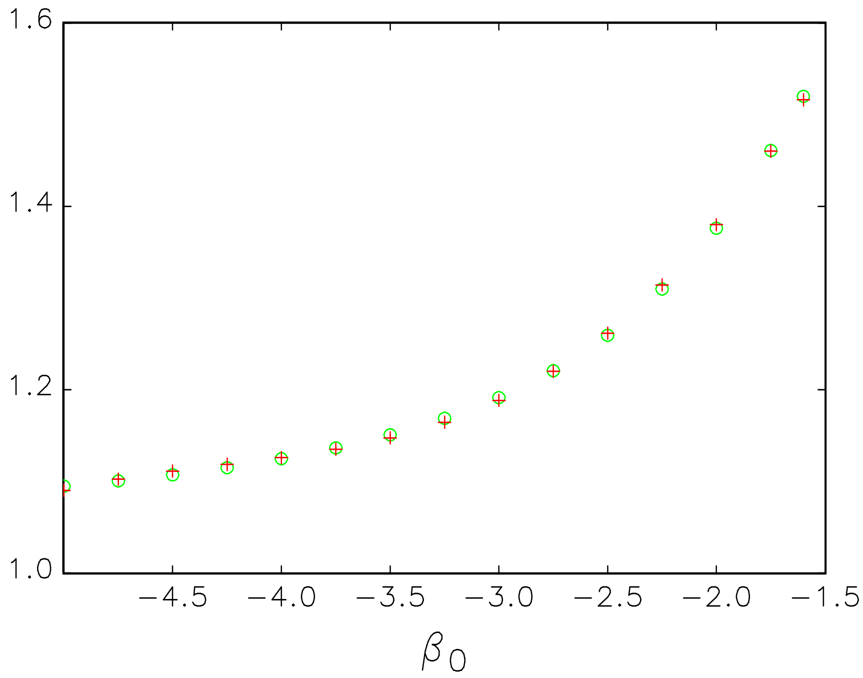


Figure 7.18: Scatter plot of the values of u for given values of β_0 for the two-variable binary logistic model with interaction with $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, b] \times [0, b]$ and with parameter values $\beta_1 = \beta_2 = -\frac{2\beta_0}{b}$, $\beta_{12} = \frac{4\beta_0}{b^2}$ and $-5 \leq \beta_0 \leq -1.5434$: The "O" represent the observed values of u, and the "+" represent the fitted values of u.

Example 7.12. Consider constructing the D-optimal design for parameters of model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, 2] \times [0, 2]$ and with parameters $\beta_0 = -3$, $\beta_1 = 3$, $\beta_2 = 3$ and $\beta_{12} = -3$. The hyperbola u is centred at $(1, 1)$ which is also the centre of the design space $[0, 2] \times [0, 2]$. Since, $\beta_0^* = 0$, $\beta_1 = \beta_2 = -\frac{2\beta_0}{b}$, $\beta_{12} = \frac{4\beta_0}{b^2}$ and $\beta_0 < -1.5434$, it follows from Theorem 7.1 that an equally weighted 8-point D-optimal design is expected. For $\beta_0 = -3$, Table 7.1 indicates that the optimal logit value is $u^* = 1.1915$. Therefore the candidate 8-point D-optimal design in this example is

$$\xi_8^* = \left\{ \begin{array}{cccccccc} (0.60, 0) & (1.40, 0) & (0, 0.60) & (0, 0.1.40) & (0.60, 2) & (1.40, 2) & (2, 0.60) & (2, 1.40) \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ -1.191 & 1.191 & -1.191 & 1.191 & 1.191 & -1.191 & 1.191 & -1.191 \end{array} \right\}. \tag{7.72}$$

The 8 support points of design (7.72) are represented by triangles in Figure 7.19 (a). The graph of the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ associated with design (7.72) is displayed in Figure 7.19 (b). The figure indicates that $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta}) \leq 4$ with equality holding at the 8 support points of design (7.72) which suggests that the design (7.72) is D-optimal.

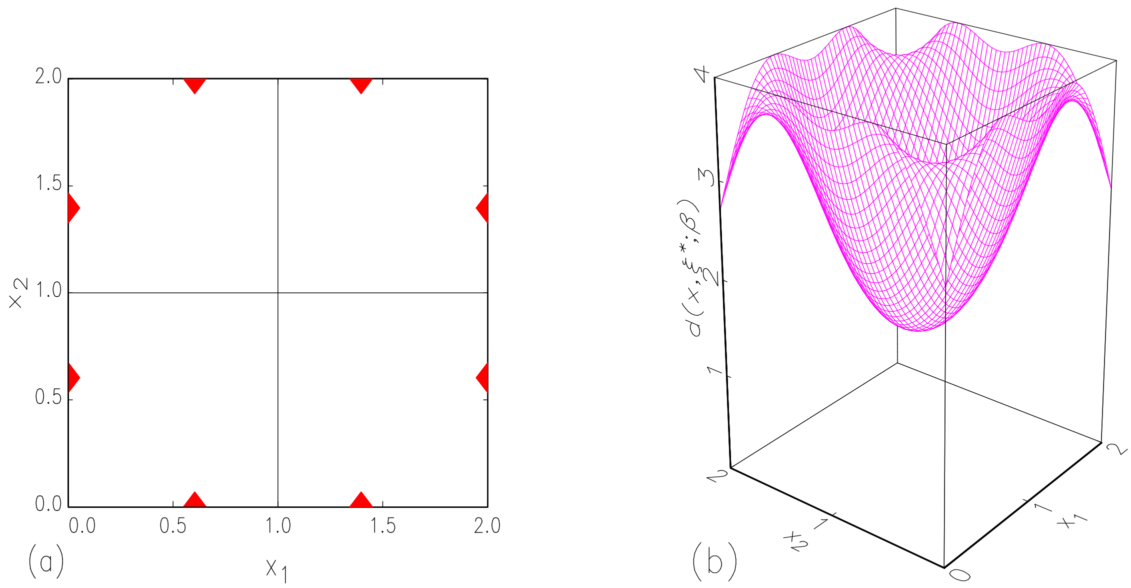


Figure 7.19: (a) Support points and (b) the standardized variance function of the D-optimal design for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, 2] \times [0, 2]$ and with parameter values $\beta_0 = \beta_{12} = -3$, $\beta_1 = \beta_2 = 3$. The triangles are support points of design (7.72).

Now, for $\beta_0 = -3$, equation (7.71) gives $u^{**} = \widehat{u} = 1.1910$ which is almost equal to the optimal value $u^* = 1.1915$ given in Table 7.1. The approximate value $u^{**} = 1.1910$ leads to an 8-point design, ξ_8, with the same values of support points and weights as the D-optimal design (7.72), and hence the efficiency of design ξ_8 relative to the D-optimal design ξ_8^* is $\text{eff}_D = \left(\frac{|M(\xi_8; \boldsymbol{\beta})|}{|M(\xi_8^*; \boldsymbol{\beta})|} \right)^{\frac{1}{4}} = 1$.

Example 7.13. Consider constructing the D-optimal design for parameters of model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, 2] \times [0, 2]$ and with parameters $\beta_0 = -1.4$, $\beta_1 = 1.4$, $\beta_2 = 1.4$ and $\beta_{12} = -1.4$. The parameter values and design space in this example are such that $\beta_1 = \beta_2 = -\frac{2\beta_0}{b}$, $\beta_{12} = \frac{4\beta_0}{b^2}$ with $b = 2$ and $-1.5434 < \beta_0 = -1.4 < 0$. Thus, it is expected that the D-optimal design for the model in this example is an equally weighted 4-point design with support points at the vertices of the design space $[0, 2] \times [0, 2]$. Therefore, the optimal logits are $u = -\beta_0 = 1.4$ and $-u = \beta_0 = -1.4$. Thus, the candidate 4-point D-optimal design is the 2^2 factorial design

$$\xi^* = \left\{ \begin{array}{cccc} (0,0) & (2,0) & (0,2) & (2,2) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ -1.4 & 1.4 & 1.4 & -1.4 \end{array} \right\}. \tag{7.73}$$

The support points of the design (7.73) are represented by symbols in Figure 7.20 (a). The standardized variance function of design (7.73) is given by Figure 7.20 (b). The figure indicates that design (7.73) is D-optimal.

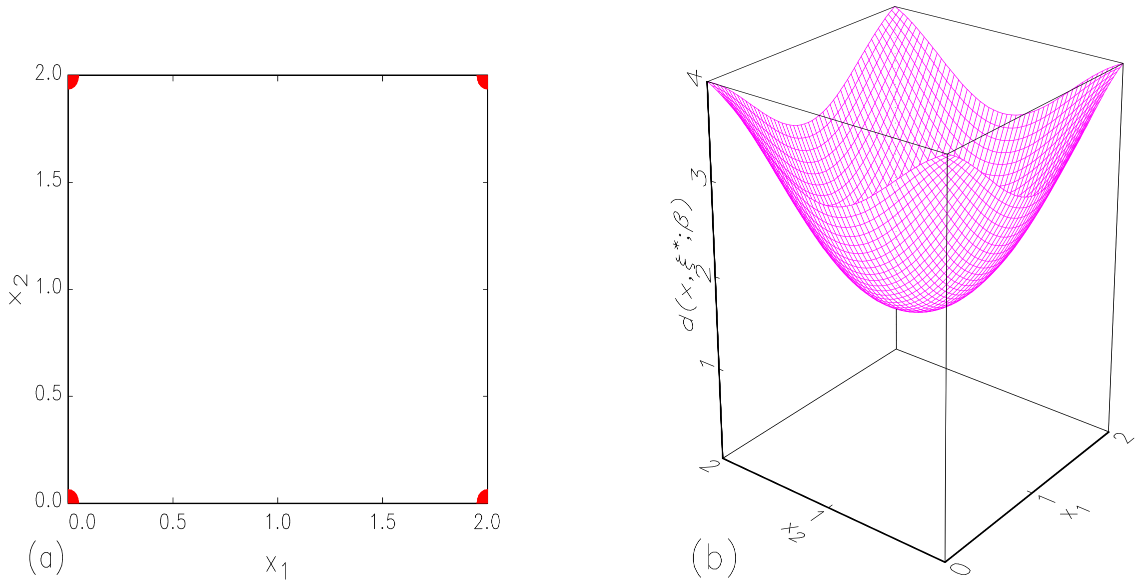


Figure 7.20: (a) Support points and (b) the standardized variance function of the D-optimal design for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0,2] \times [0,2]$ and with parameters $\beta_0 = \beta_{12} = -1.4$ and $\beta_1 = \beta_2 = 1.4$. The symbols are support points of design (7.73).

General design patterns

The 4-point and the 8-point D-optimal designs discussed above have simple design patterns as well as equally weighted support points because of the assumption that $\beta_0^* = \beta_0 - \frac{\beta_1\beta_2}{\beta_{12}} = 0$ and the centre $\left(\frac{b}{2}, \frac{b}{2}\right)$ of the design space $[0,b] \times [0,b]$ coincides with the centre $(-\frac{\beta_2}{\beta_{12}}, -\frac{\beta_1}{\beta_{12}})$ of the hyperbolae $\pm u$ where $u = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$. Without this assumption, the structure of the candidate D-optimal design can be complicated, and hence the analytical constructions of the designs is very difficult. However, even without the assumption, certain design patterns and structures of the candidate D-optimal designs can still be derived under certain symmetry conditions. This section discusses the construction of the D-optimal designs with such design patterns and structures of the candidate D-optimal designs. The construction of the D-optimal designs is done numerically due to the high complexity of analytically constructing the designs.

The design patterns and structures of the candidate D-optimal designs in this section are discussed with reference to Figure 7.21.

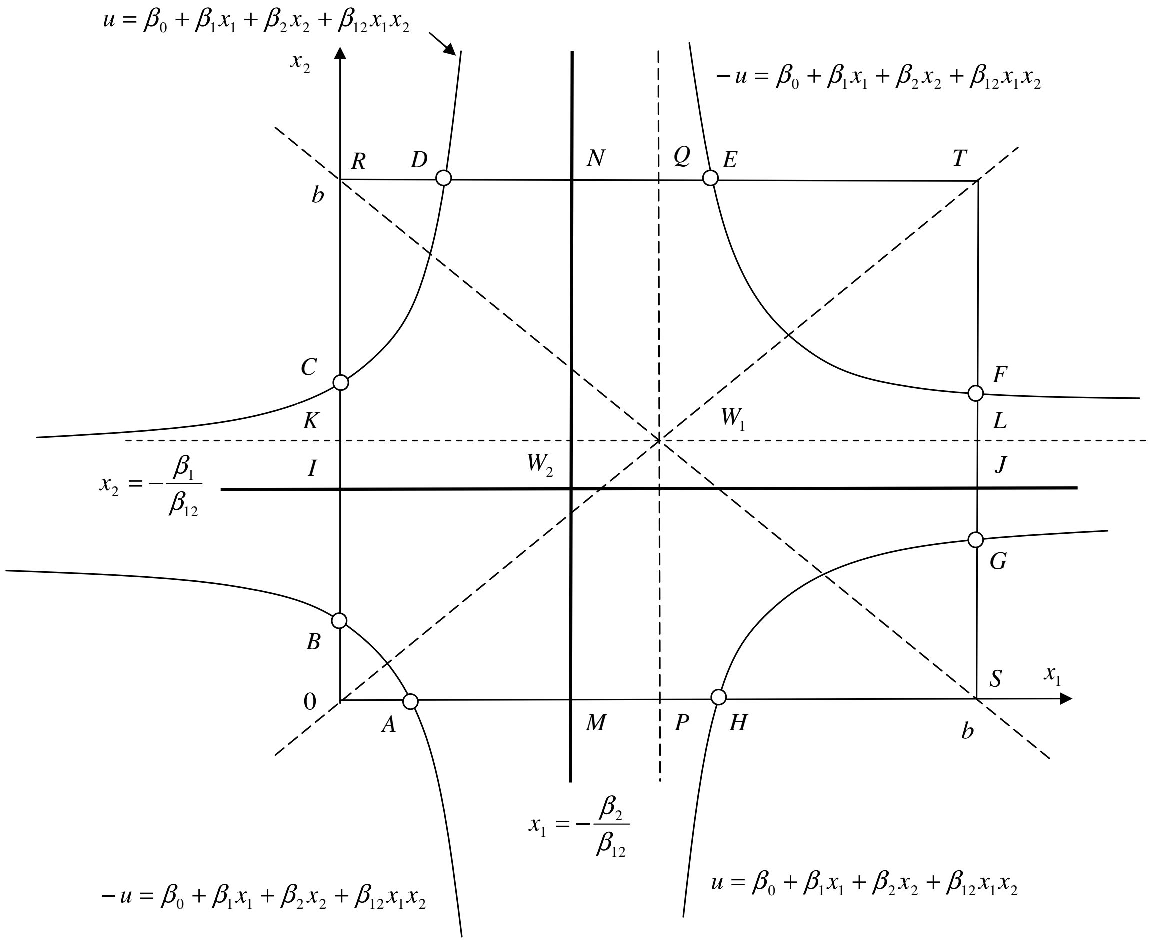


Figure 7.21: General design patterns for an 8-point D-optimal design for the two-variable binary logistic model $u = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, b] \times [0, b]$ where $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$, $\beta_{12} < 0$ and $|\beta_0^* = \beta_0 - \frac{\beta_1}{\beta_2}\beta_{12}| \leq u^* = 1.5434$. The candidate support points are represented by circles.

The following is a brief description of Figure 7.21. The point $W_1(\frac{b}{2}, \frac{b}{2})$ is the centre of the design space $[0, b] \times [0, b]$ while the point $W_2(-\frac{\beta_2}{\beta_{12}}, -\frac{\beta_1}{\beta_{12}})$ is the centre of the hyperbolae $\pm u = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$. The lines IJ and MN are the asymptotes of the hyperbolae $\pm u$, and the lines OT and RS are the diagonals of the design space $[0, b] \times [0, b]$ with respective equations $x_2 - x_1 = 0$ and $x_1 + x_2 = b$. The median lines of the design space $[0, b] \times [0, b]$ are the lines KL and MN with the respective equations $x_2 = \dfrac{b}{2}$ and $x_1 = \dfrac{b}{2}$. The support points of a candidate 8-point D-optimal design are A, B, C, D, E, F, G and H.

Symmetry conditions and their implications

Consider the 8-point design with support points A, B, C, D, E, F, G and H displayed in Figure 7.21.

Symmetry of the hyperbolae $\pm u$ about OT and RS

Symmetry about OT implies that OT passes through point W_2, or equivalently, that $\beta_1 = \beta_2$. In this case, the following pairs of points are expected to have equal weights: $\{A, B\}$, $\{C, H\}$, $\{D, G\}$, and $\{E, F\}$.

Symmetry about RS implies that RS passes through point W_2, or equivalently, that $\beta_1 + \beta_2 + b\beta_{12} = 0$. In this case, the following pairs of points are expected to have equal weights: $\{A, F\}$, $\{B, E\}$, $\{C, D\}$, and $\{H, G\}$.

Symmetry of the hyperbolae $\pm u$ about KL and PQ

Symmetry about KL implies that KL passes through point W_2 and C and B are at the same distance from KL, or equivalently, that $2\beta_0 + b\beta_2 = 0$ and $2\beta_1 + \beta_{12}b = 0$. In this case, the following points are expected to have equal weights: $\{A, D\}$, $\{B, C\}$, $\{E, H\}$, and $\{F, G\}$.

Symmetry about PQ implies that PQ passes through point W_2 and A and H are at the same distance from PQ, or equivalently, that $2\beta_2 + b\beta_1 = 0$ and $2\beta_2 + \beta_{12}b = 0$. In this case, the following points are expected to have equal weights: $\{A, H\}$, $\{B, G\}$, $\{C, F\}$, and $\{D, E\}$.

Symmetry of the hyperbolae $\pm u$ about the point W_1

Symmetry about $W_1 = W_2$ implies that the two lines MN and IJ pass through the centre of the design space $[0, b] \times [0, b]$, or equivalently that $\beta_1 = \beta_2$ and $\beta_1 + \beta_2 + b\beta_{12} = 0$. That is $\beta_1 = \beta_2$ and $2\beta_1 + b\beta_{12} = 0$. In this case, the following points are expected to have equal weights: $\{A, E\}$, $\{B, E\}$, $\{C, G\}$, and $\{D, H\}$.

Symmetry of the hyperbolae $\pm u$ about the point W_1, and about KL and PQ

In case, the following conditions are satisfied: $2\beta_0 + b\beta_1 = 0$, $2\beta_0 + b\beta_2 = 0$ and $2\beta_1 + b\beta_{12} = 0$, or equivalently, $\beta_1 = \beta_2 = -\frac{2\beta_0}{b}$ and $\beta_{12} = \frac{4\beta_0}{b^2}$ which is the case of the equally weighted 8-point design discussed above.

The implications of the above symmetry conditions on the weights of the support points of the candidate 8-point D-optimal design with the design pattern displayed in Figure 7.21 were derived assuming that the 8 support points are the expected ones. However, depending on the values of the model parameters and the design space, the number of support points of the candidate D-optimal design may not necessarily be equal to 8. For example, in Theorem 7.1, an 8-point D-optimal design exists if $\beta_0 < -1.5434$ and the design becomes a 4-point D-optimal design if $-1.5434 \leq \beta_0 \leq 0$. In general, the four hyperbolic branches of $\pm u$ may not be equidistant from the vertices of the design space $[0, b] \times [0, b]$ which results in a 4-point or an 8-point design. Thus, in theory the number of support points of a candidate D-optimal design may vary from 4 to 8 depending of the values of the model parameters and the design space $[0, b] \times [0, b]$.

D-optimal design patterns on $[0, b] \times [0, b]$ when $|\beta_0^*| < 1.5434$

Consider Figure 7.21. The condition that the curves of the hyperbolae $\pm u = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ be spread in the four quadrants generated by the asymptotes $x_1 = -\frac{\beta_2}{\beta_{12}}$ and $x_2 = -\frac{\beta_1}{\beta_{12}}$ was found to be $|\beta_0^*| < 1.5434$. Clearly, the hyperbolic curves of $\pm u$ will be within the design space $[0, b] \times [0, b]$ when the following conditions are satisfied:

$$0 \leq \frac{-u^* - \beta_0}{\beta_1} < -\frac{\beta_2}{\beta_{12}} < \frac{u^* - \beta_0}{\beta_1} \leq b$$

and

$$0 \leq \frac{u^* - \beta_0 - b\beta_2}{\beta_1 + b\beta_{12}} < -\frac{\beta_2}{\beta_{12}} < \frac{-u^* - \beta_0 - b\beta_2}{\beta_1 + b\beta_{12}} \leq b,$$

or equivalently,

$$\beta_0 \leq -u^* < \beta_0^* < u^* \leq \beta_0 + b\beta_1 \tag{7.74}$$

and

$$\beta_0 + b(\beta_1 + \beta_2) + b^2\beta_{12} \leq -u^* < \beta_0^* < u^* \leq \beta_0 + b\beta_2 \tag{7.75}$$

where $u^* = 1.5434$. Combining conditions (7.74) and (7.75) gives condition

$$\max\{\beta_0, \beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2\} \leq -1.5434 < \beta_0^* < 1.5434 \leq \min\{\beta_0 + \beta_1 b, \beta_0 + \beta_2 b\}. \tag{7.76}$$

Various violations of condition (7.76) results in candidate D-optimal designs with the number of support points varying from 4 to 8. The following are conditions of the candidate 4-point to 8-point D-optimal designs with corresponding numerical examples.

<u>Case of candidate 8-point D-optimal designs</u>

An 8-point design is expected if the inequalities in condition (7.76) are all strict, i.e.

$$\max\{\beta_0, \beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2\} < -1.5434 < \beta_0^* < 1.5434 < \min\{\beta_0 + \beta_1 b, \beta_0 + \beta_2 b\},$$

or, equivalently,

$$|\beta_0^*| < 1.5434 < \min\{-\beta_0, \beta_0 + \beta_1 b, \beta_0 + \beta_2 b, -\beta_0 - (\beta_1 + \beta_2)b - \beta_{12}b^2\}. \tag{7.77}$$

In such a case, all the branches of the hyperbolae $\pm u$ in Figure 7.21 intersect the design space $[0, b] \times [0, b]$, but no point of intersection is a vertex of the design space.

Table 7.2 displays examples of numerically constructed 8-point D-optimal designs on the design space $[0, 2] \times [0, 2]$ and for selected values of $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$ which satisfy conditions (7.77). The designs were constructed using the Gauss program in Appendix C and the D-optimality of the designs were checked using the plots of the standardized variance function.

Table 7.2: D-optimal designs with 8 support points for selected values of parameters for model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, b] \times [0, b]$ where $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$, $\beta_{12} < 0$, $\beta_1 + \beta_{12} b < 0$, $\beta_2 + \beta_{12} b < 0$, $|\beta_0^*| < 1.5434$ with $b = 2$.

$\boldsymbol{\beta}$	D-optimal designs									
-3		x_1	0.472	1.028	0	0	0.314	0.686	2	2
4	ξ_1^*	x_2	0	0	0.629	1.371	2	2	0.555	0.873
3		λ	0.1881	0.0960	0.0396	0.0387	0.1728	0.1405	0.1587	0.1657
-5		u	-1.114	1.114	-1.114	1.114	1.114	-1.114	1.114	-1.114
3		x_1	0.475	1.025	0	0	0.650	1.017	2	2
-4	ξ_2^*	x_2	0	0	0.475	1.025	2	2	0.650	1.017
-4		λ	0.1115	0.0870	0.1126	0.0862	0.1600	0.1414	0.1594	0.1419
5		u	-1.099	1.099	-1.099	1.099	1.099	-1.099	1.099	-1.099
-3		x_1	0.464	1.036	0	0	0.964	1.536	2	2
4	ξ_3^*	x_2	0	0	0.464	1.036	2	2	0.964	1.536
4		λ	0.1112	0.1380	0.1168	0.1340	0.1380	0.1112	0.1340	0.1168
-4		u	-1.142	1.142	-1.142	1.142	1.142	-1.142	1.142	-1.142

Consider the designs ξ_1^*, ξ_2^* and ξ_3^* in Table 7.2. The design ξ_1^* in Table 7.2 is an 8-point design because the model parameter values satisfy conditions (7.77), and the weights of the support points are all different because none of the symmetry conditions discussed above holds. For design ξ_2^*, the conditions are satisfied as well as the symmetry condition $\beta_1 = \beta_2$. Hence, the following pairs of support points are almost equally weighted.

$$\{(0.475, 0), (0, 0.475)\}, \{(1.025, 0), (0, 1.025)\}, \{(0.650, 2), (2, 0.650)\}, \{(1.017, 2), (2, 1.017)\}.$$

Finally, for design ξ_3^*, conditions (7.77) and conditions $\beta_1 = \beta_2$ and $2\beta_1 + b\beta_{12} = 0$ are satisfied, and hence there is symmetry relative to the center of the design space, and hence the following pairs of support points are equally weighted.

$$\{(0.464, 0), (1.536, 2)\}, \{(1.036, 0), (0.964, 2)\}, \{(0, 0.464), (2, 1.536)\}, \{(0, 1.036), (2, 0.964)\}.$$

Cases of 7-point design patterns

A 7-point D-optimal design is expected if one of the inequalities in condition (7.76) does not hold. The following are examples of when a 7-point D-optimal design on the design space $[0, b] \times [0, b]$ can be expected.

$$\beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2 < -1.5434 \leq \beta_0 < \beta_0^* < 1.5434 \leq \min\{\beta_0 + \beta_1 b, \beta_0 + \beta_2 b\}. \quad (7.78)$$

This is the condition for the support points A and B in Figure 7.21 to be both $(0, 0)$.

$$\beta_0 < -1.5434 \leq \beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2 < \beta_0^* < 1.5434 \leq \min\{\beta_0 + \beta_1 b, \beta_0 + \beta_2 b\}. \quad (7.79)$$

This is the condition for the support points E and F in Figure 7.21 to be both (b, b).

$$\max\{\beta_0, \beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2\} \leq -1.5434 < \beta_0^* < \beta_0 + \beta_1 b \leq 1.5434 < \beta_0 + \beta_2 b \quad (7.80)$$

This is the condition for the support points G and H in Figure 7.21 to be both $(b, 0)$.

$$\max\{\beta_0, \beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2\} \leq -1.5434 < \beta_0^* < \beta_0 + \beta_2 b \leq 1.5434 < \beta_0 + \beta_1 b \quad (7.81)$$

This is the condition for the support points C and D in Figure 7.21 to be both $(0, b)$.

In Table 7.3, the design ξ_1^* is an example of a numerically constructed 7-point design on the design space $[0, 2] \times [0, 2]$ and for the value of $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$ which satisfy condition (7.78). The design was constructed using the Gauss program in Appendix C and D-optimality of the design was checked using the plot of the standardized variance function. Note that none of the symmetry conditions discussed above is satisfied and hence the 7-point D-optimal design ξ_1^* is not symmetric.

Table 7.3: D-optimal designs with 4 to 7 support points for selected values of parameters for model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, b] \times [0, b]$ where $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$, $\beta_{12} < 0$, $\beta_1 + \beta_{12} b < 0$, $\beta_2 + \beta_{12} b < 0$ and $|\beta_0^*| < u^*$ with $b = 2$.

$\boldsymbol{\beta}$	D-optimal designs								
-1		x_1	0	0.714	0	0.493	0.821	2	2
3	ξ_1^*	x_2	0	0	0.766	2	2	0.535	0.854
2.8		λ	0.2252	0.0666	0.0207	0.1950	0.1504	0.1767	0.1654
-5		u	-1	1.143	1.143	1.146	-1.146	1.146	-1.146
-3		x_1	0.746	0	0	0.931	2	2	
2	ξ_2^*	x_2	0	0.672	2	2	0	0.850	
2.2		λ	0.0625	0.1963	0.2306	0.1786	0.2281	0.1039	
-2.5		u	-1.509	-1.522	1.4	-1.393	1	-1.379	
-3		x_1	0.755	0	0	2	2		
2	ξ_3^*	x_2	0	0.681	2	0	2		
2.2		λ	0.0696	0.1910	0.2450	0.2444	0.2500		
-1.7		u	-1.490	-1.503	1.4	1	-1.4		
-1.5		x_1	0	0	2	2			
1.3	ξ_4^*	x_2	0	2	0	2			
1.2		λ	0.25	0.25	0.25	0.25			
-1		u	-1.5	0.9	1.1	-0.5			

Case of candidate 6-point D-optimal designs

A 6-point D-optimal design is expected if 2 of the 4 vertices of the design space $[0, b] \times [0, b]$ in Figure 7.21 are support points of the candidate D-optimal design. This means that there are $\binom{4}{2} = 6$ conditions under which a 6-point D-optimal design can be expected. The following is the list of the 6 conditions

$$-1.5434 \leq [\beta_0, \beta_0 + (\beta_1 + \beta_2)b + \beta_{12} b^2] < \beta_0^* < 1.5434 < \min\{\beta_0 + \beta_1 b, \beta_0 + \beta_2 b\}. \quad (7.82)$$

This is the condition for the pairs $\{A, B\}$ and $\{E, F\}$ of support points in Figure 7.21 to be $(0, 0)$ and (b, b), respectively.

$$\max\{\beta_0, \beta_0 + (\beta_1 + \beta_2)b + \beta_{12} b^2\} < -1.5434 < \beta_0^* < [\beta_0 + \beta_1 b, \beta_0 + \beta_2 b] \leq 1.5434. \quad (7.83)$$

This is the condition for the pairs $\{C, D\}$ and $\{G, H\}$ of support points in Figure 7.21 to be $(0, b)$ and $(b, 0)$, respectively.

$$\beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2 < -1.5434 \leq \beta_0 < \beta_0^* < \beta_0 + \beta_2 b \leq 1.5434 < \beta_0 + \beta_1 b \tag{7.84}$$

This is the condition for the pairs $\{A, B\}$ and $\{C, D\}$ of support points in Figure 7.21 to be $(0, 0)$ and $(0, b)$, respectively.

$$\beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2 < -1.5434 \leq \beta_0 < \beta_0^* < \beta_0 + \beta_1 b \leq 1.5434 < \beta_0 + \beta_2 b. \tag{7.85}$$

This is the condition for the pairs $\{A, B\}$ and $\{G, H\}$ of support points in Figure 7.21 to be $(0, 0)$ and $(b, 0)$, respectively.

$$\beta_0 < -1.5434 \leq \beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2 < \beta_0^* < \beta_0 + \beta_2 b \leq 1.5434 < \beta_0 + \beta_1 b \tag{7.86}$$

This is the condition for the pairs $\{C, D\}$ and $\{E, F\}$ of support points in Figure 7.21 to be $(0, b)$ and (b, b), respectively.

$$\beta_0 < -1.5434 \leq \beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2 < \beta_0^* < \beta_0 + \beta_1 b \leq 1.5434 < \beta_0 + \beta_2 b \tag{7.87}$$

This is the condition for the pairs $\{G, H\}$ and $\{E, F\}$ of support points in Figure 7.21 to be $(b, 0)$ and (b, b), respectively.

Design ξ_2^* in Table 7.3 is an example of a numerically constructed 6-point D-optimal design on the design space $[0, 2] \times [0, 2]$ and for the value of $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$ which satisfy condition (7.83). The design was constructed numerically using the Gauss program in Appendix C and the D-optimality of the design checked using the plot of the standardized variance function. Note that none of the symmetry conditions discussed above is satisfied and hence the support points of the 6-point D-optimal design ξ_2^* are unequally weighted.

Cases of candidate 5-point D-optimal designs

A 5-point design is expected if 3 of the 4 vertices of the design space $[0, b] \times [0, b]$ in Figure 7.21 are support points of the candidate D-optimal design. This means that there are $\binom{4}{3} = 4$ conditions under which a 5-point D-optimal design can be expected. The following is the list of the 4 conditions.

$$-1.5434 \leq [\beta_0, \beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2] < \beta_0^* < \beta_0 + \beta_2 b \leq 1.5434 < \beta_0 + \beta_1 b. \tag{7.88}$$

This is the condition for the pairs $\{A, B\}$, $\{C, D\}$ and $\{E, F\}$ of support points in Figure 7.21 to be $(0, 0)$, $(0, b)$ and (b, b), respectively.

$$-1.5434 \leq [\beta_0, \beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2] < \beta_0^* < \beta_0 + \beta_1 b \leq 1.5434 < \beta_0 + \beta_2 b. \tag{7.89}$$

This is the condition for the pairs $\{A, B\}$, $\{G, H\}$ and $\{E, F\}$ of support points in Figure 7.21 to be $(0, 0)$, $(b, 0)$ and (b, b), respectively.

$$\beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2 < -1.5434 \leq \beta_0 < \beta_0^* < [\beta_0 + \beta_1 b, \beta_0 + \beta_2 b] \leq 1.5434. \tag{7.90}$$

This is the condition for the pairs $\{A, B\}$, $\{C, D\}$ and $\{G, H\}$ of support points in Figure 7.21 to be $(0, 0)$, $(0, b)$ and $(b, 0)$, respectively.

$$\beta_0 < -1.5434 \leq \beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2 < \beta_0^* < [\beta_0 + \beta_1 b, \beta_0 + \beta_2 b] \leq 1.5434. \tag{7.91}$$

This is the condition for the pairs $\{C, D\}$, $\{G, H\}$ and $\{E, F\}$ of support points in Figure 7.21 to be $(0, b)$, $(b, 0)$, and (b, b), respectively.

Design ξ_3^* in Table 7.3 is an example of a numerically constructed 5-point D-optimal design on the design space $[0, 2] \times [0, 2]$ and for the value of $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$ which satisfy condition (7.91). The design was constructed using the Gauss program in Appendix C and the D-optimality of the design checked using the plot of the standardized variance function. Note that none of the symmetry conditions discussed above is satisfied and hence the support points of the 5-point D-optimal design ξ_3^* are unequally weighted.

Case of candidate 4-point D-optimal design

A 4-point D-optimal design is expected if all the 4 vertices of the design space $[0, b] \times [0, b]$ in Figure 7.21 are support points of the candidate D-optimal design, or equivalently, if all the 4 inequalities in condition (7.76) do not hold. That is, if

$$|\beta_0^*| < \max\{-\beta_0, \beta_0 + \beta_1 b, \beta_0 + \beta_2 b, -\beta_0 - (\beta_1 + \beta_2)b - \beta_{12}b^2\} \leq 1.5434. \tag{7.92}$$

Design ξ_4^* in Table 7.3 is an example of a numerically constructed 4-point D-optimal design on the design space $[0, 2] \times [0, 2]$ and for the value of $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$ which satisfy condition (7.92). The design was constructed using the Gauss program in Appendix C and the D-optimality of the design checked using the plot of the standardized variance function.

7.5.4 D-optimal designs when $|\beta_0^*| < 1.5434$, and either $-\frac{\beta_2}{\beta_{12}} < b$ and $-\frac{\beta_1}{\beta_{12}} > b$, or $-\frac{\beta_2}{\beta_{12}} > b$ and $-\frac{\beta_1}{\beta_{12}} < b$

For $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_{12} < 0$, if $|\beta_0^*| = |\beta_0 - \dfrac{\beta_1\beta_2}{\beta_{12}}| < 1.5434$, then the hyperbolae $\pm u = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ are spread in the four quadrants generated by the asymptotes $x_1 = -\dfrac{\beta_2}{\beta_{12}}$ and $x_2 = -\dfrac{\beta_1}{\beta_{12}}$ as in Figures 7.22 (a) and (b). If in addition to the condition on

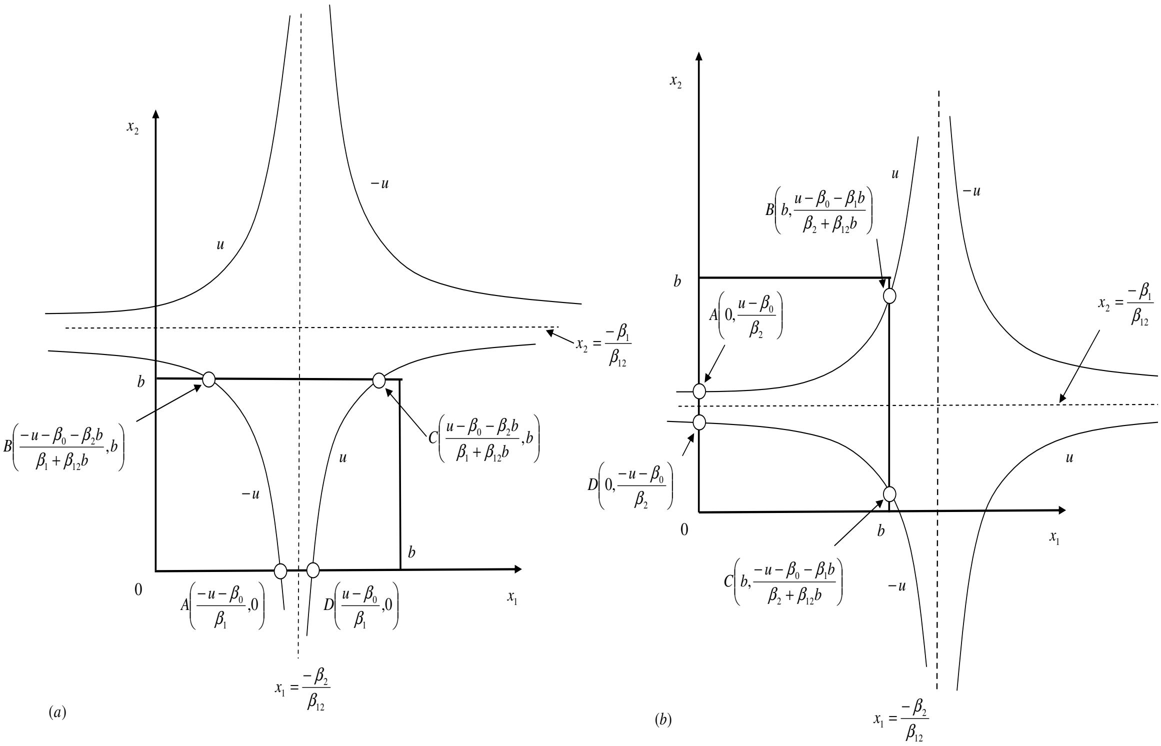


Figure 7.22: Design patterns of the candidate D-optimal designs for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, b] \times [0, b]$ where (a) $\beta_0 + \beta_2 b \leq -u < \beta_0^* < u \leq \beta_0 + (\beta_1 + \beta_2)b + \beta_{12} b^2$ and (b) $\beta_0 + \beta_1 b \leq -u < \beta_0^* < u \leq \beta_0 + (\beta_1 + \beta_2)b + \beta_{12} b^2$.

β_0^*, $0 < -\dfrac{\beta_2}{\beta_{12}} < b$ and $-\dfrac{\beta_1}{\beta_{12}} > b$, then the candidate D-optimal design has 4 support points indicated by circles in Figures 7.22 (a). The 4-point D-optimal design with the design pattern of Figure 7.22 (a) is possible if

$$0 \leq \frac{-u - \beta_0 - \beta_2 b}{\beta_1 + b\beta_{12}} < -\frac{\beta_2}{\beta_{12}} < \frac{u - \beta_0 - \beta_2 b}{\beta_1 + b\beta_{12}} \leq b,$$

or equivalently, if

$$\beta_0 + \beta_2 b \leq -u < \beta_0^* < u \leq \beta_0 + (\beta_1 + \beta_2)b + \beta_{12} b^2. \tag{7.93}$$

On the other hand, if in addition to the condition on β_0^*, $0 < -\dfrac{\beta_1}{\beta_{12}} < b$ and $-\dfrac{\beta_2}{\beta_{12}} > b$, then the candidate D-optimal design has 4 support points indicated by circles in Figures 7.22 (b). The 4-point D-optimal design with the design pattern of Figure 7.22 (b) is possible if

$$\beta_0 + \beta_1 b \leq -u < \beta_0^* < u \leq \beta_0 + (\beta_1 + \beta_2)b + \beta_{12} b^2. \tag{7.94}$$

Theorem 7.3 is about the D-optimality of the candidate 4-point D-optimal designs with the design patterns of the form displayed in Figures 7.22 (a) and (b).

Proposition 7.3. The design

$$\xi_1^* = \left\{ \begin{array}{cccc} \left(\frac{-1.5434-\beta_0}{\beta_1}, 0\right) & \left(\frac{-1.5434-\beta_0-\beta_2 b}{\beta_1+\beta_{12} b}, b\right) & \left(\frac{1.5434-\beta_0-\beta_2 b}{\beta_1+\beta_{12} b}, b\right) & \left(\frac{1.5434-\beta_0}{\beta_1}, 0\right) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array} \right\} \tag{7.95}$$

for the two-variable binary logistic model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, b] \times [0, b]$ with the model parameters satisfying the conditions $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_{12} < 0$, and $\beta_0 + \beta_2 b \leq -1.5434 < \beta_0^* < 1.5434 \leq \beta_0 + (\beta_1 + \beta_2)b + \beta_{12} b^2$, is D-optimal. Similarly, the design

$$\xi_2^* = \left\{ \begin{array}{cccc} \left(0, \frac{1.5434-\beta_0}{\beta_2}\right) & \left(b, \frac{1.5434-\beta_0-\beta_1 b}{\beta_2+\beta_{12} b}\right) & \left(b, \frac{-1.5434-\beta_0-\beta_1 b}{\beta_2+\beta_{12} b}\right) & \left(0, \frac{-1.5434-\beta_0}{\beta_2}\right) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array} \right\} \tag{7.96}$$

for the two-variable binary logistic model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, b] \times [0, b]$ with the model parameters satisfying the conditions $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_{12} < 0$, and $\beta_0 + \beta_1 b \leq -1.5434 < \beta_0^* < 1.5434 \leq \beta_0 + (\beta_1 + \beta_2)b + \beta_{12} b^2$, is D-optimal.

Proof

Consider design ξ_1^* given by (7.95) as the candidate D-optimal design for the 4-parameter two-variable binary logistic model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, b] \times [0, b]$ with design pattern displayed in Figure 7.22 (a) where the support points are represented by circles. The pattern of design points in Figure 7.22 (a) indicates that the candidate D-optimal design is of the form

$$\xi_1 = \left\{ \begin{array}{cccc} \left(\frac{-u-\beta_0}{\beta_1}, 0\right) & \left(\frac{-u-\beta_0-\beta_2 b}{\beta_1+\beta_{12} b}, b\right) & \left(\frac{u-\beta_0-\beta_2 b}{\beta_1+\beta_{12} b}, b\right) & \left(\frac{u-\beta_0}{\beta_1}, 0\right) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{array} \right\} \tag{7.97}$$

where u satisfying equation (7.93). Clearly, the design ξ_1 given (7.97) has the same form as the design (7.6), where $c = 0$ and $d = b$, which was shown to be D-optimal in Proposition 7.2. Hence, the optimal value of u in design (7.97) is $u = u^* = 1.5434$ so that the design (7.95) is D-optimal and condition (7.93) becomes

$$\beta_0 + \beta_2 b \leq -1.5434 < \beta_0^* < 1.5434 \leq \beta_0 + (\beta_1 + \beta_2)b + \beta_{12} b^2. \tag{7.98}$$

Now, consider the 4-point design ξ_2^* given by (7.96) corresponds to the case when $|\beta_0^*| < 1.5434$, and $-\frac{\beta_2}{\beta_{12}} > b$ and $-\frac{\beta_1}{\beta_{12}} < b$. Clearly, Figures 7.22 (a) and (b) suggest that the candidate

D-optimal design ξ_2^* with support points corresponding the design pattern in Figure 7.22 (b) can be obtained from design ξ_1^* given in (7.95) by exchanging the parameters β_1 and β_2. Hence, the design (7.96) is D-optimal provided that condition (7.94), where $u = u^* = 1.5434$, is satisfied. The transformation on parameter values and support points from design (7.95) to design (7.96) can be derived as follows.

A support point (x_1, x_2) of design (7.96) can be a transform of a support point (y_1, y_2) of design (7.95) using the transformation

$$(x_1, x_2) = (y_2, b - y_1). \tag{7.99}$$

Clearly, a logit line u in the design of Figure 7.22 (b) is image of a logit line $\widetilde{u} = -u$ in Figure 7.22 (a). The correspondence between the values of the parameters can be obtained as follows. Assume that $\widetilde{u} = \beta_0' + \beta_1' y_1 + \beta_2' y_2 + \beta_{12}' y_1 y_2$. Setting $\widetilde{u} = -u$ gives

$$\begin{aligned}\beta_0' + \beta_1' y_1 + \beta_2' y_2 + \beta_{12}' y_1 y_2 &= \widetilde{u} = -u \\ &- \beta_0 - \beta_1 y_1 - \beta_2(b - y_1) - \beta_{12} y_2 (b - y_1) \\ &= -\beta_0 - \beta_2 b + \beta_2 y_1 - (\beta_1 + \beta_{12} b) y_2 + \beta_{12} y_1 y_2,\end{aligned}$$

and hence

$$\begin{cases} \beta_0' &= -\beta_0 - \beta_2 b \\ \beta_1' &= \beta_2 \\ \beta_2' &= -\beta_1 - \beta_{12} b \\ \beta_{12}' &= \beta_{12} \end{cases} \iff \begin{cases} \beta_0 &= -\beta_0' - \beta_1' b \\ \beta_1 &= -\beta_2' - \beta_{12}' b \\ \beta_2 &= -\beta_1' \\ \beta_{12} &= \beta_{12}'. \end{cases} \tag{7.100}$$

For example, applying the transformations (7.99) and (7.100) to the second support point of design (7.95) gives the second support of design (7.96) since $x_1 = y_2 = b$ and

$$x_2 = b - y_1 = b - \frac{-1.5434 - \beta_0' - \beta_2' b}{\beta_1' + \beta_{12}' b} = \frac{1.5434 + (\beta_0' + b\beta_1') + b(\beta_2' + \beta_{12}' b)}{\beta_1' + \beta_{12}' b} = \frac{1.5434 - \beta_0 - \beta_1 b}{\beta_2 + \beta_{12} b}.$$

Other support points of design (7.96) can be obtained from the support points of design (7.95) in the same way. The weights of the corresponding support points of designs (7.95) and (7.95) remain equal. Hence, design ξ_2^* given in (7.96) is also D-optimal because it is the image of the D-optimal design (7.95) through the transformations (7.99) and (7.100). ■.

The following numerical example illustration the above theoretical results.

Example 7.14. Consider the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, 2] \times [0, 2]$ and with parameters $\beta_0 = -4$, $\beta_1 = 5$, $\beta_2 = 1$ and $\beta_{12} = -1.2$. In this example,

$$\beta_0 + \beta_2 b = -2 < -1.5434 < \beta_0^* \simeq 0.17 < 1.5434 < \beta_0 + (\beta_1 + \beta_2) b + \beta_{12} b^2 = 3.2.$$

Hence, the condition (7.98) is satisfied, and thus the candidate D-optimal design associated to this example given in (7.95) is

$$\xi_1^* = \left\{ \begin{array}{cccc} (0.491, 0) & (0.176, 2) & (1.363, 2) & (1.109, 0) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ -1.5434 & -1.5434 & 1.5434 & 1.5434 \end{array} \right\}.$$

The support points of the design ξ_1^* are represented by triangles in Figure 7.23 (a), and the standardized variance function is displayed in Figure 7.23 (b) which suggests that the above design is D-optimal.

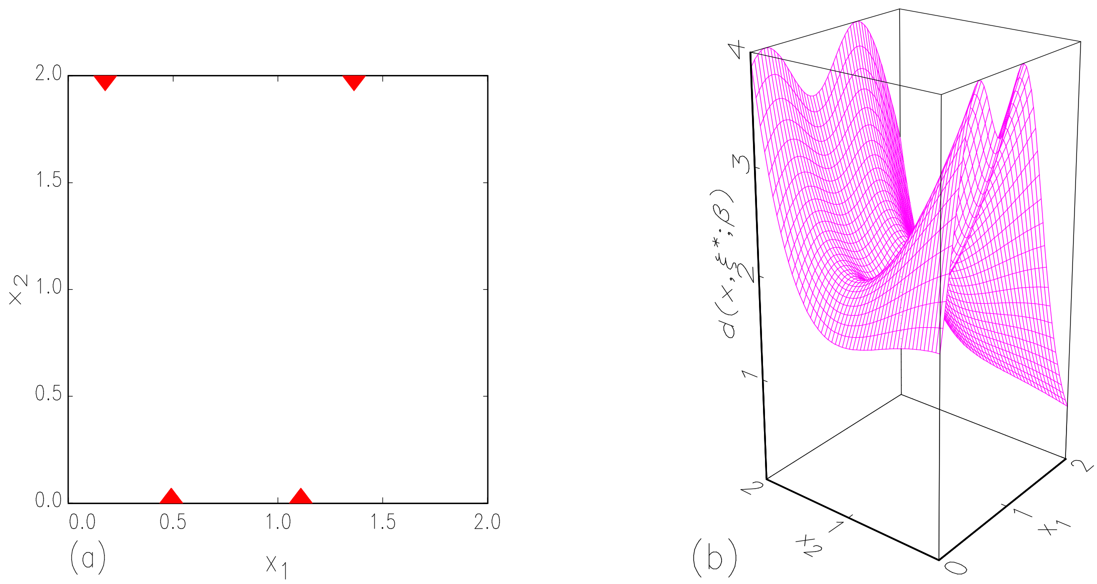


Figure 7.23: (a) Support points and (b) the standardized variance function of the candidate D-optimal design for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, 2] \times [0, 2]$ and with parameter vector $\boldsymbol{\beta} = (-4, 5, 1, -1.2)^T$.

Comments

If in Figure 7.22 (a) or Figure 7.22 (b) some branches of the optimal hyperbolae $\pm u = 1.5434$ pass through vertices of or outside the design space $[0, b] \times [0, b]$, then the vertices become support points of the D-optimal design instead of any other point outside the design space. Consequently, the logits at the support points, for example at B and C, are not necessarily equal to the optimal logits $\pm u = 1.5434$. The following numerical example is an illustration.

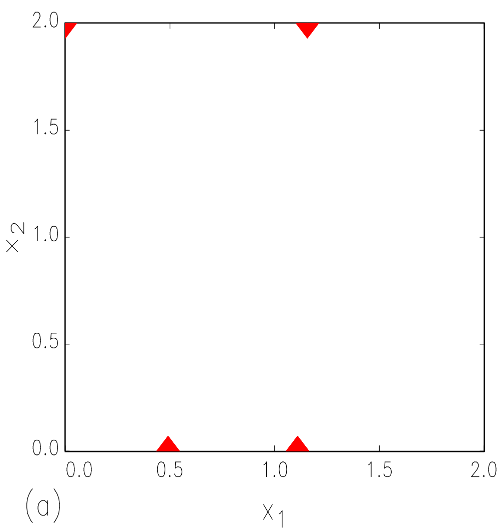
Example 7.15. Consider the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, 2] \times [0, 2]$ and with parameters $\beta_0 = -4$, $\beta_1 = 5$, $\beta_2 = 1.3$ and $\beta_{12} = -1.2$. The design space and the model parameter values in this example are the same as those in Example 7.14 with $\beta_2 = 1$ replaced by $\beta_2 = 1.3$. In this case

$$-1.5434 < \beta_0 + \beta_2 b = -1.4 < \beta_0^* = 1.42 < 1.5434 < \beta_0 + (\beta_1 + \beta_2)b + \beta_{12} b^2 = 3.8.$$

Hence, the D-optimal design has the pattern of design (7.95), but with the vertex $(0, 2)$ expected to be one of the support points. The presence of this vertex support point forces some of the support points to lie off the optimal logits ± 1.5434. Indeed, numerical calculations of the D-optimal design for the model in this example gave the 4-point design

$$\xi_2^* = \left\{ \begin{array}{cccc} (0.491, 0) & (0, 2) & (1.156, 2) & (1.109, 0) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ -1.5434 & -1.4 & 1.6053 & 1.5434 \end{array} \right\}.$$

Figures 7.24 (a) and (b) indicate the location of the support points on the designs space $[0, 2] \times [0, 2]$, and that the design is D-optimal.



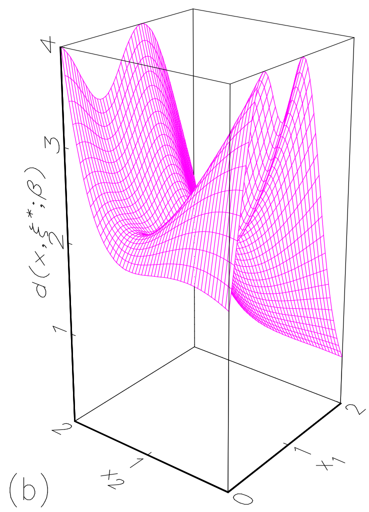


Figure 7.24: (a) Support points and (b) the standardized variance function of the D-optimal design for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, 2] \times [0, 2]$ and with parameters $\beta_0 = -4$, $\beta_1 = 5$, $\beta_2 = 1.3$ and $\beta_{12} = -1.2$. The triangles are support points of the design.

7.5.5 D-optimal designs when $\beta_0^* > 1.5434$, and $-\frac{\beta_1}{\beta_{12}} < b$ and $-\frac{\beta_1}{\beta_{12}} < b$

For $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_{12} < 0$, if $\beta_0^* = \beta_0 - \dfrac{\beta_1\beta_2}{\beta_{12}} > 1.5434$ then the hyperbolae $\pm u = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ are located in the first and third quadrants generated by the asymptotes $x_1 = -\dfrac{\beta_2}{\beta_{12}}$ and $x_2 = -\dfrac{\beta_1}{\beta_{12}}$ as illustrated in Figure 7.25. If in addition to the

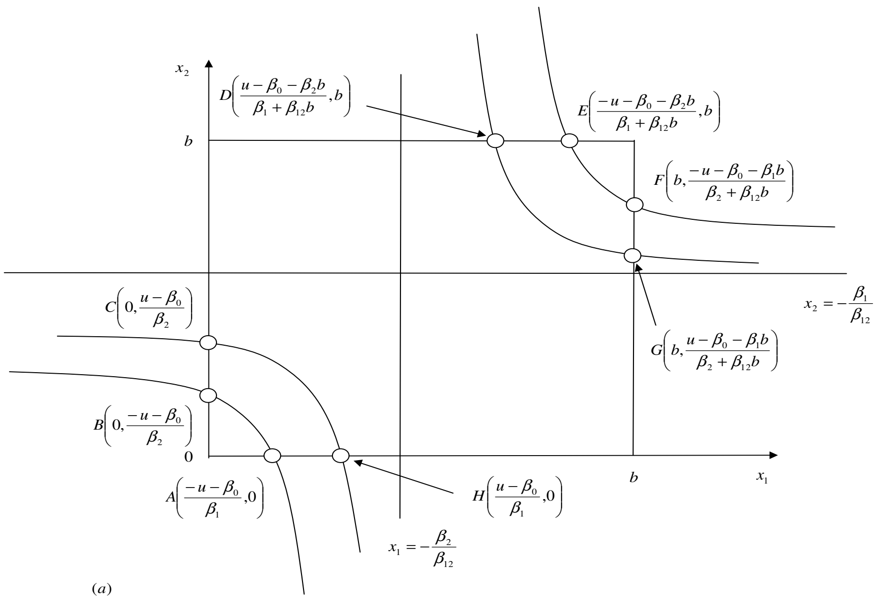


Figure 7.25: Design patterns of the candidate D-optimal designs for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, b] \times [0, b]$ where $\max(\beta_0, \beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2) < -u < u < \beta_0^* < \min(\beta_0 + \beta_1 b, \beta_0 + \beta_2 b)$. Circles represent the support points.

condition on β_0^*, $0 < -\dfrac{\beta_1}{\beta_{12}} < b$ and $0 < -\dfrac{\beta_2}{\beta_{12}} < b$, then the candidate D-optimal design has the 8 support points indicated by circles in Figure 7.25. Using the same reasoning as in Section 7.5.3, the 8-point D-optimal design with design pattern in Figure 7.25 is expected if

$$\max(\beta_0, \beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2) \leq -1.5434 < 1.5434 < \beta_0^* < \min(\beta_0 + \beta_1 b, \beta_0 + \beta_2 b). \qquad (7.101)$$

As in Section 7.5.3, violations of at least one of the inequalities in condition (7.101) leads to candidate D-optimal designs with fewer than 8 support points. The following are conditions of the candidate 4-point to 8-point D-optimal designs with corresponding numerical examples. The designs were constructed using the Gauss program in Appendix C and the D-optimality of the designs checked using the plot of the standardized variance function.

Case of candidate 8-point D-optimal designs

An 8-point D-optimal design is expected if condition (7.101) holds, but with strict inequality signs. For example, the designs ξ_1^* in Table 7.4 is an examples of numerically constructed 8-point D-optimal designs on the design space $[0,2] \times [0,2]$ and for the parameter values $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$ which satisfy condition (7.101). None of the symmetry conditions discussed in Section 7.5.3 are satisfied for design ξ_1^* and hence the support points of the 8-point D-optimal design ξ_1^* are unequally weighted.

Table 7.4: D-optimal designs for selected values of parameters for model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0,b] \times [0,b]$ where $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$, $\beta_{12} < 0$, $\beta_1 + \beta_{12} b < 0$, $\beta_2 + \beta_{12} b < 0$ and $\beta_0^* > 1.5434$ or $\beta_0^* < -1.5434$ with $b = 2$.

$\boldsymbol{\beta}$	D-optimal designs									
-3		x_1	0.311	0.689	0	0	0.966	1.534	2	2
6	ξ_1^*	x_2	0	0	0.466	1.034	2	2	1.311	1.689
4		λ	0.1327	0.1502	0.0831	0.1342	0.1309	0.0878	0.1526	0.1285
-5		u	-1.135	1.135	-1.135	1.135	1.135	-1.135	1.135	-1.135
9		x_1	1.689	1.311	2	2	1.034	0.466	0	0
-6	ξ_2^*	x_2	0	0	0.466	1.034	2	2	1.311	1.689
-6		λ	0.1327	0.1502	0.0831	0.1342	0.1309	0.0878	0.1526	0.1285
5		u	-1.135	1.135	-1.135	1.135	1.135	-1.135	1.135	-1.135

Case of candidate 7-point D-optimal designs

A 7-point D-optimal design is expected if one vertex of the design space is a support point of the candidate D-optimal design. The following are conditions of when a 7-point D-optimal design on the design space $[0,b] \times [0,b]$ can be expected.

$$\beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2 < -1.5434 \leq \beta_0 < 1.5434 < \beta_0^* < \min(\beta_0 + \beta_1 b, \beta_0 + \beta_2 b). \qquad (7.102)$$

This is the condition for the support points A and B in Figure 7.25 to be both $(0,0)$.

$$\beta_0 < -1.5434 \leq \beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2 < 1.5434 < \beta_0^* < \min(\beta_0 + \beta_1 b, \beta_0 + \beta_2 b). \qquad (7.103)$$

This is the condition for the support points E and F in Figure 7.25 to be both (b,b).

For example, in Table 7.5, the design ξ_1^* is an example of a numerically constructed 7-point D-optimal design on the design space $[0,2] \times [0,2]$ and for the value of $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$ which satisfy condition (7.103).

Table 7.5: D-optimal designs for selected values of parameters for model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, b] \times [0, b]$ where $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$, $\beta_{12} < 0$, $\beta_1 + \beta_{12} b < 0$, $\beta_2 + \beta_{12} b < 0$ and $\beta_0^* > 1.5434$ with $b = 2$.

$\boldsymbol{\beta}$	D-optimal designs								
-3	ξ_1^*	x_1	0.253	0.604	0	0	2	1.259	2
7		x_2	0	0	0.443	1.057	1.630	2	2
4		λ	0.1652	0.1863	0.0636	0.2388	0.1367	0.0048	0.2046
-5		u	-1.227	1.227	-1.227	1.227	1.219	1.222	-1
-1	ξ_2^*	x_1	0	1.032	0	2	1.288	2	
2.5		x_2	0	0	0.712	0.968	2	2	
3.5		λ	0.2203	0.1349	0.1448	0.1348	0.1449	0.2203	
-3		u	-1	1.580	1.490	1.580	1.490	-1	
-3	ξ_3^*	x_1	0.338	0.912	0	0	2		
4.8		x_2	0	0	0.406	1.094	2		
4		λ	0.1339	0.2411	0.1339	0.2411	0.250		
-2.5		u	-1.376	1.376	-1.376	1.376	3		

Case of 6-point design patterns

A 6-point D-optimal design is expected if 2 vertices of the design space $[0, b] \times [0, b]$ in Figure 7.25 are support points of the candidate D-optimal design. In such a case, the vertices $(0, 0)$ and (b, b) both lie on the logit line $u = -1.5434$ or are above and below the logit line $u = -1.5434$, respectively. Hence, a 6-point D-optimal design is expected if

$$-1.5434 \leq [\beta_0, \beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2] < 1.5434 < \min(\beta_0 + \beta_1 b, \beta_0 + \beta_2 b). \tag{7.104}$$

This is the condition for the pairs $\{A, B\}$ and $\{B, F\}$ of support points in Figure 7.25 to be $(0, 0)$ and (b, b), respectively. In Table 7.5, the design ξ_2^* is an example of a numerically constructed 6-point D-optimal design on the design space $[0, 2] \times [0, 2]$ and for the value of $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$ which satisfies condition (7.104).

Case of 5-point design patterns

Consider Figure 7.25. A 5-point candidate D-optimal design is expected in either of the following cases: (1) two branches, one of $-u$ and one of u lie at or below point $(0, 0)$, and the other two branches one of $-u$ and one of u lie above point $(0, 0)$ but below point (b, b); or (2) two branches, one of $-u$ and one of u lie at or above point (b, b) and the other two branches, one of $-u$ and one of u lie below point (b, b) but above point $(0, 0)$. Hence, a 5-point

D-optimal design may exist if one of the following two conditions is satisfied.

$$\beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2 < -1.5434 < 1.5434 \leq \beta_0 < \beta_0^* < \min(\beta_0 + \beta_1 b, \beta_0 + \beta_2 b). \qquad (7.105)$$

This the condition for the support points A and B in Figure 7.25 to be both $(0, 0)$.

$$\beta_0 < -1.5434 < 1.5434 \leq \beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2 < \beta_0^* < \min(\beta_0 + \beta_1 b, \beta_0 + \beta_2 b). \qquad (7.106)$$

This is the condition for the support points E and F in Figure 7.25 to be both (b, b). For example, in Table 7.5, the design ξ_3^* is an example of a numerically constructed 5-point D-optimal design on the design space $[0, 2] \times [0, 2]$ and for the value of $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$ which satisfy condition (7.106).

7.5.6 D-optimal designs when $\beta_0^* < -1.5434$, and $-\frac{\beta_2}{\beta_{12}} < b$ and $-\frac{\beta_1}{\beta_{12}} < b$

For $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_{12} < 0$, if $\beta_0^* = \beta_0 - \dfrac{\beta_1\beta_2}{\beta_{12}} < -1.5434$ then the branches of the hyperbolae $\pm u = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12}x_1x_2$ are located in the second and fourth quadrants generated by the asymptotes $x_1 = -\dfrac{\beta_2}{\beta_{12}}$ and $x_2 = -\dfrac{\beta_1}{\beta_{12}}$ as in Figure 7.26.

If in addition to the condition on β_0^*, $0 < -\dfrac{\beta_1}{\beta_{12}} < b$ and $0 < -\dfrac{\beta_2}{\beta_{12}} < b$, then the candidate D-optimal design has the 8 support points indicated by circles in Figure 7.26. Note that the relationship between support points (x_1, x_2) of this candidate 8-point D-optimal design and the support points (y_1, y_2) of the candidate D-optimal design with the design pattern displayed in Figure 7.25 is given by

$$(x_1, x_2) = (-y_1 + b, y_2). \qquad (7.107)$$

Clearly, if u is a logit line in the design of Figure 7.26, then the corresponding logit line in Figure 7.25 is $\widetilde{u} = -u$. The correspondence between the values of the parameters can be obtained as follows. Assume that $\widetilde{u} = \beta'_0 + \beta'_1 y_1 + \beta'_2 y_2 + \beta'_{12} y_1 y_2$. Setting $\widetilde{u} = -u$ gives

$$\begin{aligned}
\beta'_0 + \beta'_1 y_1 + \beta'_2 y_2 + \beta'_{12} y_1 y_2 &= \widetilde{u} = -u \\
&= -\beta_0 - \beta_1 x_1 - \beta_2 x_2 - \beta_{12} x_1 x_2 \\
&= -\beta_0 - \beta_1(-y_1 + b) - \beta_2 y_2 - \beta_{12}(-y_1 + b)y_2 \\
&= -\beta_0 - \beta_1 b + \beta_1 y_1 - (\beta_2 + \beta_{12} b)y_2 + \beta_{12} y_1 y_2
\end{aligned}$$

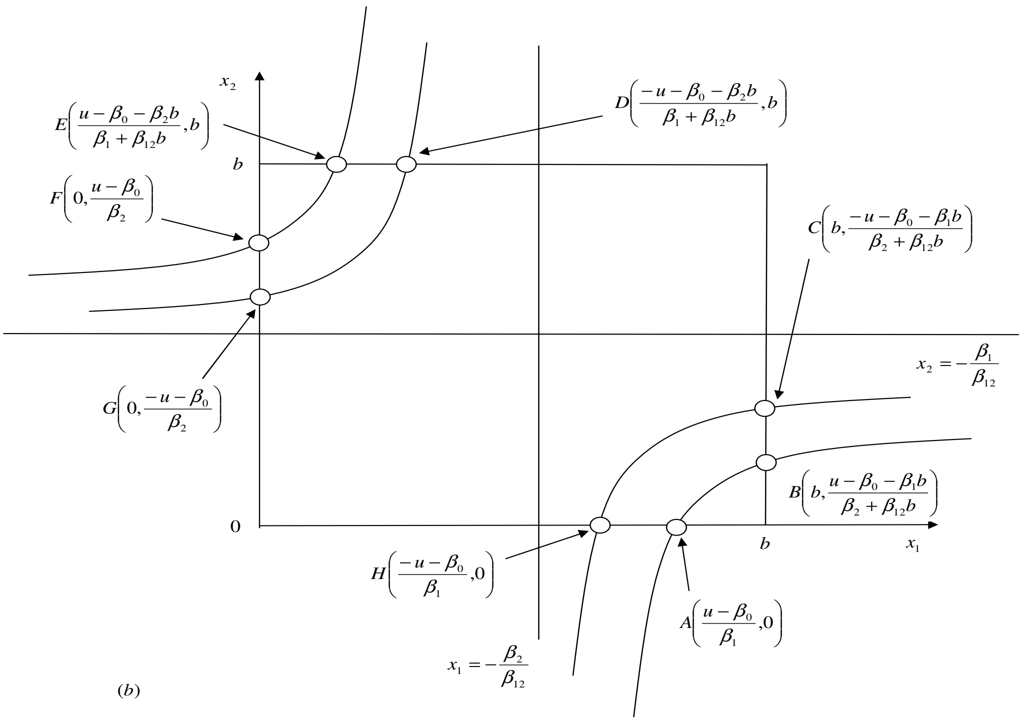


Figure 7.26: Design patterns of the candidate D-optimal designs for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, b] \times [0, b]$ where $\max(\beta_0, \beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2) < -u < u < \beta_0^* < \min(\beta_0 + \beta_1 b, \beta_0 + \beta_2 b)$.

and, hence

$$\left\{\begin{array}{lcl} \beta_0' & = & -\beta_0 - \beta_1 b \\ \beta_1' & = & \beta_1 \\ \beta_2' & = & -\beta_2 - \beta_{12} b \\ \beta_{12}' & = & \beta_{12} \end{array}\right. \iff \left\{\begin{array}{lcl} \beta_0 & = & -\beta_0' - \beta_1' b \\ \beta_1 & = & \beta_1' \\ \beta_2 & = & -\beta_2' - \beta_{12}' b \\ \beta_{12} & = & \beta_{12}'. \end{array}\right. \tag{7.108}$$

For example, consider point C in Figure 7.25. Using the notation above, assume that $C\left(0, \frac{\tilde{u}-\beta_0'}{\beta_2'}\right)$. The transformations (7.107) and (7.108) give $x_1 = -y_1 + b = b$ and

$$x_2 = \frac{u - \beta_0'}{\beta_2'} = \frac{u + \beta_0 + \beta_1 b}{-\beta_2 - \beta_{12} b} = \frac{-u - \beta_0 - \beta_1 b}{\beta_2 + \beta_{12} b}$$

and these are coordinates of point C in Figure 7.25. Also, note that if $\beta_0^{*'} = \beta_0' - \frac{\beta_1'\beta_2'}{\beta_{12}'}$ for design pattern in Figure 7.25 and $\beta_0^* = \beta_0 - \frac{\beta_1\beta_2}{\beta_{12}}$ for the design pattern in Figure 7.26, it follows from (7.108) that

$$\beta_0^{*'} = \beta_0' - \frac{\beta_1'\beta_2'}{\beta_{12}'} = -\beta_0 - \beta_1 b - \frac{\beta_1(-\beta_2 - \beta_{12} b)}{\beta_{12}} = -\left(\beta_0 - \frac{\beta_1\beta_2}{\beta_{12}}\right) = -\beta_0^*.$$

Consequently, for $u^* = -1.5434$, condition $\beta_0^{*'} > u^*$ is transformed to $-\beta_0^* > u^*$ or, equivalently, to $\beta_0^* < -u^*$.

For example, the support points of the D-optimal design ξ_2^* in Table 7.4 are obtained from those of the D-optimal design ξ_1^* in Table 7.4 using the transformation (7.107). The weights at support points of design ξ_2^* are the same as those of design ξ_1^*, but the optimal logit at each support point of design ξ_2^* has an opposite sign to that of the logit of the corresponding support point of design ξ_1^*.

7.5.7 D-optimal designs when $\beta_0^* > 1.5434$, and either $-\frac{\beta_2}{\beta_{12}} < b$ and $-\frac{\beta_1}{\beta_{12}} > b$, or $-\frac{\beta_2}{\beta_{12}} > b$ and $-\frac{\beta_1}{\beta_{12}} < b$

For $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$, $\beta_{12} < 0$ and $\beta_0^* = \beta_0 - \dfrac{\beta_1\beta_2}{\beta_{12}}$, if $\beta_0^* > 1.5434$, then the branches of the hyperbolae $\pm u = \beta_0 + b_1x_1 + b_2x_2 + \beta_{12}x_1x_2$ are located in the first and third quadrants generated by the asymptotes $x_1 = -\dfrac{\beta_2}{\beta_{12}}$ and $x_2 = -\dfrac{\beta_1}{\beta_{12}}$ as indicated in Figure 7.27 (a) and (b). The support points of the designs are represented by circles in the figures. If $-\dfrac{\beta_2}{\beta_{12}} < b$ and $-\dfrac{\beta_1}{\beta_{12}} > b$, then the candidate D-optimal design has the design pattern displayed in Figure 7.27 (a), and if $-\dfrac{\beta_2}{\beta_{12}} > b$ and $-\dfrac{\beta_1}{\beta_{12}} < b$, then the candidate D-optimal design has the design pattern displayed in Figure 7.27 (b).

Consider the design pattern displayed in Figure 7.27 (a). The support points at A, B, C and D are points of intersection of the boundary of the design space $[0, b] \times [0, b]$ and the two branches of the hyperbolae $\pm u$. However, the interaction parameter β_{12} may not be estimable from a design with support points A, B, C and D since the interaction term in model vanishes at the four support points. Hence, there is a need of an interior support point, say the vertex (b, b) at which the interaction term in the model does not vanish. The choice of vertex (b, b) as a support point is sensible because it is an obvious substitute for the imaginary points of intersection of the boundary of the design space and the two branches of the hyperbolae $\pm u$ which do not intersect the design space. Thus, the candidate D-optimal design with the design pattern displayed in Figure 7.27 (a) is

$$\xi = \left\{ \begin{array}{ccccc} \left(\frac{-u-\beta_0}{\beta_1}, 0\right) & \left(0, \frac{-u-\beta_0}{\beta_2}\right) & \left(0, \frac{u-\beta_0}{\beta_2}\right) & \left(\frac{u-\beta_0}{\beta_1}, 0\right) & (b, b) \\ \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 & \lambda_5 \\ -u & -u & u & u & \beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2 \end{array} \right\} \quad (7.109)$$

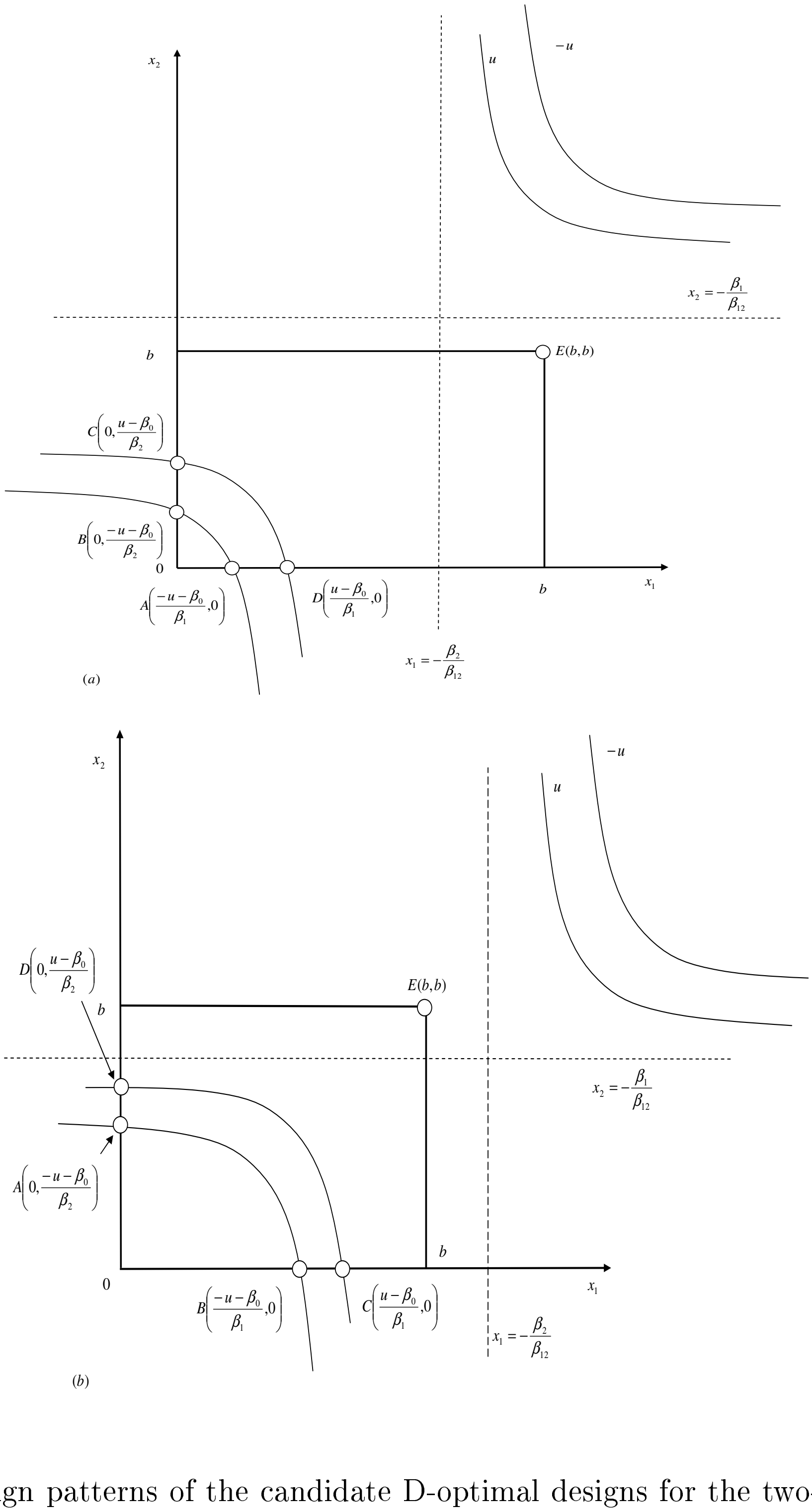


Figure 7.27: Design patterns of the candidate D-optimal designs for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, b] \times [0, b]$ and with parameter values with (a) $\beta_0 < -1.5434 < 1.5434 \leq \beta_0 + \beta_2 b < \beta_0^* < \beta_0 + \beta_1 b$ and (b) $\beta_0 < -1.5434 < 1.5434 \leq \beta_0 + \beta_1 b < \beta_0^* < \beta_0 + \beta_2 b$.

where $0 < \lambda_i < 1$ and $\sum_{i=1}^{5} \lambda_i = 1$. Clearly, the condition for the existence of design (7.109) is

$$\beta_0 < -1.5434 < 1.5434 \leq \beta_0 + \beta_2 b < \beta_0^* < \beta_0 + \beta_1 b \tag{7.110}$$

Table 7.5 displays two examples of the numerically constructed D-optimal designs of the form design (7.109) on the design space $[0, 2] \times [0, 2]$ and for selected values of the parameter vector $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$. Note that the values of the model parameters which obtained ξ_1^* in

Table 7.6: D-optimal for selected values of parameters for model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, b] \times [0, b]$ and with parameters satisfying conditions $-\frac{\beta_2}{\beta_{12}} < b$ and $-\frac{\beta_1}{\beta_{12}} > b$ or $-\frac{\beta_2}{\beta_{12}} > b$ and $-\frac{\beta_1}{\beta_{12}} < b$.

$\boldsymbol{\beta}$	D-optimal designs						
-3		x_1	0.325	0.875	0	0	2
5	ξ_1^*	x_2	0	0	0.541	1.459	2
3		λ^*	0.1339	0.2411	0.1339	0.2411	0.2500
-2		u^*	-1.376	1.376	-1.376	1.376	5
-3		x_1	0	0	0.541	1.459	2
5	ξ_2^*	x_2	0.325	0.875	0	0	2
2		λ^*	0.1339	0.2411	0.1339	0.2411	0.2500
-2		u^*	-1.376	1.376	-1.376	1.376	5

Table 7.6 satisfy condition (7.110). In fact,

$$\beta_0 = -3 < -1.5434 < 1.5434 \leq \beta_0 + \beta_2 b = 3 < \beta_0^* = 4.5 < \beta_0 + \beta_1 b = 7.$$

Now, consider the design pattern displayed in Figure 7.27 (b). The candidate D-optimal design with this pattern is

$$\xi = \left\{ \begin{array}{ccccc} \left(0, \frac{-u-\beta_0}{\beta_2}\right) & \left(\frac{-u-\beta_0}{\beta_1}, 0\right) & \left(0, \frac{u-\beta_0}{\beta_2}\right) & \left(\frac{u-\beta_0}{\beta_1}, 0\right) & (b, b) \\ \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 & \lambda_5 \\ -u & -u & u & u & \beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2 \end{array} \right\} \tag{7.111}$$

where $0 < \lambda_i < 1$ and $\sum_{i=1}^{5} \lambda_i = 1$. Note that a support point (x_1, x_2) of the design (7.111) is related to a support point (y_1, y_2) of design (7.109) by the transformation

$$(x_1, x_2) = (y_2, y_1) \tag{7.112}$$

Clearly, if u is a logit line in the design of Figure 7.27 (b), then the corresponding logit line in Figure 7.27 (a) is $u' = u$. The correspondence between the values of the parameters can be obtained as follows. Assume that $u' = \beta_0' + \beta_1' y_1 + \beta_2' y_2 + \beta_{12}' y_1 y_2$. Setting $u' = u$ gives

$$\begin{aligned}\beta_0' + \beta_1' y_1 + \beta_2' y_2 + \beta_{12}' y_1 y_2 = u' = u \\ &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 \\ &= \beta_0 + \beta_1 y_2 + \beta_2 y_1 + \beta_{12} y_2 y_1\end{aligned}$$

Hence, the respective parameter vectors $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_{12})^T$ and $\boldsymbol{\beta}' = (\beta_0', \beta_1', \beta_2', \beta_{12}')^T$ are linked by

$$\begin{cases} \beta_0' &= \beta_0 \\ \beta_1' &= \beta_2 \\ \beta_2' &= \beta_1 \\ \beta_{12}' &= \beta_{12}. \end{cases} \tag{7.113}$$

For example, the vector of parameters $\boldsymbol{\beta} = (-3, 5, 3, -2)^T$ for the candidate D-optimal design ξ_2^* is linked to the vector of parameters $\boldsymbol{\beta} = (-3, 3, 5, -5)$ for the design ξ_1^* in Table 7.6 by the relations (7.113). Thus the support points of the design ξ_2^* can be derived from those of design ξ_1 by the transformation (7.112). The weights and the optimal logits remain unchanged.

7.5.8 D-optimal designs when $\beta_0^* < -1.5434$, and either $-\frac{\beta_2}{\beta_{12}} < b$ and $-\frac{\beta_1}{\beta_{12}} > b$, or $-\frac{\beta_2}{\beta_{12}} > b$ and $-\frac{\beta_1}{\beta_{12}} < b$

For $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$, $\beta_{12} < 0$ and $\beta_0^* = \beta_0 - \dfrac{\beta_1\beta_2}{\beta_{12}}$, if $\beta_0^* < -1.5434$, then the branches of the hyperbolae $\pm u = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ are located in the second and fourth quadrants generated by the asymptotes $x_1 = -\dfrac{\beta_2}{\beta_{12}}$ and $x_2 = -\dfrac{\beta_1}{\beta_{12}}$ as indicated in Figures 7.28 (a) and (b). The figures display the design patterns of the candidate D-optimal designs, and the support points of the designs are represented by circles in the figures. If $-\frac{\beta_2}{\beta_{12}} < b$ and $-\frac{\beta_1}{\beta_{12}} > b$, then the candidate D-optimal design has the design pattern displayed in Figure 7.28 (a), and if $-\frac{\beta_2}{\beta_{12}} > b$ and $-\frac{\beta_1}{\beta_{12}} < b$, then the candidate D-optimal design has the design pattern displayed in Figure 7.28 (b).

Consider the design pattern displayed in Figure 7.28 (a). The candidate D-optimal design with the displayed design pattern is

$$\xi = \left\{ \begin{array}{ccccc} \left(\frac{u-\beta_0}{\beta_1}, 0\right) & \left(b, \frac{u-\beta_0-\beta_1 b}{\beta_2+\beta_{12} b}\right) & \left(b, \frac{-u-\beta_0-\beta_1 b}{\beta_2+\beta_{12} b}\right) & \left(0, \frac{u-\beta_0}{\beta_2}\right) & (0, b) \\ \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 & \lambda_5 \\ u & u & -u & -u & \beta_0 + \beta_2 b \end{array} \right\} \tag{7.114}$$

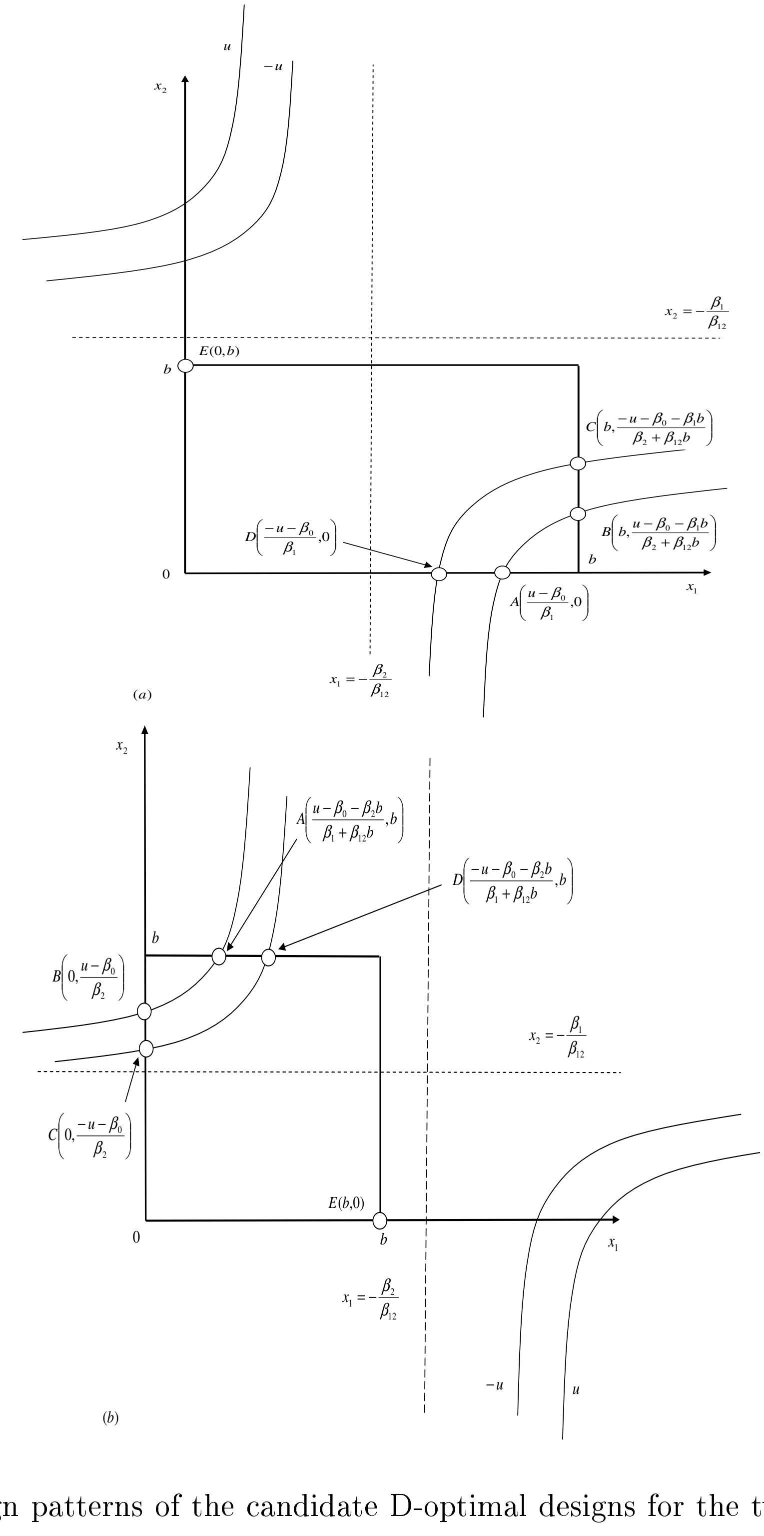


Figure 7.28: Design patterns of the candidate D-optimal designs for the two-variable binary logistic model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, b] \times [0, b]$ with (a) $\beta_0 < \beta_0^* < \beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2 \leq -1.5434 < 1.5434 < \beta_0 + \beta_1 b$ and (b) $\beta_0 < \beta_0^* < \beta_0 + (\beta_1 + \beta_2)b + \beta_{12}b^2 \leq -1.5434 < 1.5434 < \beta_0 + \beta_2 b$.

where $0 < \lambda_i < 1$ and $\sum_{i=1}^{5} \lambda_i = 1$. Note that design (7.114) can be obtained from design (7.109) using the transformation

$$(x_1, x_2) = (b - y_1, y_2) \tag{7.115}$$

where (x_1, x_2) is a support point in design (7.114) and (y_1, y_2) is a support point in design (7.109). If u is a logit line for a design pattern in Figure 7.28 (a), then the corresponding logit line in Figure 7.27 (a) is $\widetilde{u} = -u$. Hence, the relationship between parameters can be derived as follows using the transformation (7.115).

$$\begin{aligned}
\beta'_0 + \beta'_1 y_1 + \beta'_2 y_2 + \beta'_{12} y_1 y_2 = \widetilde{u} &= -u \\
&= -\beta_0 - \beta_1 x_1 - \beta_2 x_2 - \beta_{12} x_1 x_2 \\
&= -\beta_0 - \beta_1(b - y_1) - \beta_2 y_2 - \beta_{12}(b - y_1) y_2 \\
&= -\beta_0 - \beta_1 b + \beta_1 y_1 - (\beta_2 + \beta_{12} b) y_2 + \beta_{12} y_1 y_2
\end{aligned}$$

and, thus

$$\left\{\begin{array}{lcl} \beta'_0 & = & -\beta_0 - \beta_1 b \\ \beta'_1 & = & \beta_1 \\ \beta'_2 & = & -\beta_2 - \beta_{12} b \\ \beta'_{12} & = & \beta_{12} \end{array}\right. \iff \left\{\begin{array}{lcl} \beta_0 & = & -\beta'_0 - \beta'_1 b \\ \beta_1 & = & \beta'_1 \\ \beta_2 & = & -\beta'_2 - \beta'_{12} b \\ \beta_{12} & = & \beta'_{12}. \end{array}\right. \tag{7.116}$$

Hence, the condition (7.110) for the existence of design (7.109) is transformed to condition

$$\beta_0 < \beta_0^* < \beta_0 + (\beta_1 + \beta_2) b + \beta_{12} b^2 \leq -1.5434 < 1.5434 < \beta_0 + \beta_1 b. \tag{7.117}$$

For example, the parameter values for design ξ_1^* in Table 7.7 satisfy condition (7.117), and hence a design of the form (7.114) can be conjectured. Furthermore, the parameters of design ξ_1^* in Table 7.7 are linked to those of design ξ_1^* in Table 7.6 by the transformation (7.116). Consequently, the support points and associated weights and logits of design ξ_1^*, in Table 7.7 can be calculated from those of design ξ_1^* in Table 7.6 using transformation (7.115).

Now, consider the design pattern displayed in Figure 7.28 (b). The candidate D-optimal design with the displayed design pattern is given by

$$\xi = \left\{\begin{array}{ccccc} \left(\frac{u-\beta_0-\beta_2 b}{\beta_1+\beta_{12} b}, b\right) & \left(0, \frac{u-\beta_0}{\beta_2}\right) & \left(0, \frac{-u-\beta_0}{\beta_2}\right) & \left(\frac{-u-\beta_0-\beta_2 b}{\beta_1+\beta_{12} b}, b\right) & (b, 0) \\ \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 & \lambda_5 \\ u & u & -u & -u & \beta_0 + \beta_1 b \end{array}\right\}. \tag{7.118}$$

Design (7.118) can be obtained from design (7.109) by the transformation

$$(x_1, x_2) = (y_1, b - y_2) \tag{7.119}$$

Table 7.7: D-optimal designs for selected values of parameters for model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, b] \times [0, b]$ and with parameters satisfying conditions $\beta_0^* < -1.5434$ with $-\frac{\beta_2}{\beta_{12}} < b$ and $-\frac{\beta_1}{\beta_{12}} > b$, or $-\frac{\beta_2}{\beta_{12}} > b$ and $-\frac{\beta_1}{\beta_{12}} < b$.

$\boldsymbol{\beta}$	D-optimal designs						
-7		x_1	1.675	2	2	1.125	0
5	ξ_1^*	x_2	0	0.541	1.459	0	2
1		λ^*	0.13391	0.1339	0.2411	0.2411	0.2500
-2		u^*	1.376	1.376	-1.376	-1.376	-5
-7		x_1	0	0.541	1.459	0	2
1	ξ_2^*	x_2	1.675	2	2	1.125	0
5		λ^*	0.1339	0.1339	0.2411	0.2411	0.2500
-2		u^*	1.376	1.376	-1.376	-1.376	-5

where (x_1, x_2) is a support point in design (7.118) and (y_1, y_2) is a support point in design (7.109). The transformation (7.119) which is similar to transformation (7.115) by interchanging the roles of x_1 and x_2. Therefore, the relationship between parameters can found from equations (7.116) by interchanging the roles of β_1 and β_2 to obtain

$$\left\{\begin{array}{lcl} \beta_0' & = & -\beta_0 - \beta_2 b \\ \beta_1' & = & -\beta_1 - \beta_{12} b \\ \beta_2' & = & \beta_2 \\ \beta_{12}' & = & \beta_{12} \end{array}\right. \iff \left\{\begin{array}{lcl} \beta_0 & = & -\beta_0' - \beta_2' b \\ \beta_1 & = & -\beta_1' - \beta_{12}' b \\ \beta_2 & = & \beta_2' \\ \beta_{12} & = & \beta_{12}'. \end{array}\right. \tag{7.120}$$

For example, the parameter values in Table 7.7 and equations (7.120) indicate that the support points of design ξ_2^* in Table can be obtained from those of design ξ_2^* in Table 7.6 using the transformation (7.119). The weights and the logit values remain unchanged. Equivalently, design ξ_2^* in Table 7.7 can be deduced from design ξ_1^* in the same table by exchanging x_1 and x_2 coordinates of support points.

7.6. Other cases of synergy and antagonism

Recall from Section 2.4.4 that in the two-variable binary logistic model

$$u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 \tag{7.121}$$

two cases of synergy and two cases of antagonism were possible. Synergy occurs when $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_{12} > 0$ or $\beta_0 > 0$, $\beta_1 < 0$, $\beta_2 < 0$ and $\beta_{12} < 0$. Antagonism occurs when $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_{12} < 0$ or $\beta_0 > 0$, $\beta_1 < 0$, $\beta_2 < 0$ and $\beta_{12} > 0$. The construction of D-optimal designs for synergy when $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_{12} > 0$ was discussed in Section 7.4, and that for antagonism when $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_{12} < 0$ in Section 7.5. This section briefly discusses the construction of the D-optimal designs for model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ for synergy when $\beta_0 > 0$, $\beta_1 < 0$, $\beta_2 < 0$ and $\beta_{12} < 0$, and for antagonism when $\beta_0 > 0$, $\beta_1 < 0$, $\beta_2 < 0$ and $\beta_{12} > 0$.

Consider model (7.121) for synergy on the design spaces $[0, \infty) \times [0, \infty)$ or $[0, b] \times [0, b]$ where $\beta_0 > 0$, $\beta_1 < 0$, $\beta_2 < 0$ and $\beta_{12} < 0$. In this case, $-\dfrac{\beta_2}{\beta_{12}} < 0$ and $-\dfrac{\beta_1}{\beta_{12}} < 0$ so that the asymptotes $x_1 = -\dfrac{\beta_2}{\beta_{12}}$ and $x_2 = -\dfrac{\beta_1}{\beta_{12}}$ of the hyperbola (7.121) intersect in the third quadrant of $\mathbb{R}^2$. The construction of D-optimal designs for this case is the same as that of Section 7.4 by multiplying each parameter β_0, β_1, β_2 and β_{12} by -1. In fact, multiplying both sides of (7.121) by -1 gives model

$$U = B_0 + B_1 x_1 + B_2 x_2 + B_{12} x_1 x_2 \tag{7.122}$$

where $U = -u$, $B_0 = -\beta_0 < 0$, $B_1 = -\beta_1 > 0$, $B_2 = -\beta_2 > 0$ and $B_{12} = -\beta_{12} > 0$. Clearly, model (7.122) is that of synergy discussed in Section 7.4. If the design space remains the same, the support points and weights of the D-optimal design remain unchanged, but the optimal logit u^* changes to $-u^*$.

Now, consider model (7.121) for antagonism on the design spaces $[0, \infty) \times [0, \infty)$ or $[0, b] \times [0, b]$ where $\beta_0 > 0$, $\beta_1 < 0$, $\beta_2 < 0$ and $\beta_{12} > 0$. In this case, $-\dfrac{\beta_2}{\beta_{12}} > 0$ and $-\dfrac{\beta_1}{\beta_{12}} > 0$ so that the asymptotes $x_1 = -\dfrac{\beta_2}{\beta_{12}}$ and $x_2 = -\dfrac{\beta_1}{\beta_{12}}$ of the hyperbola (7.121) intersect in the first quadrant of $\mathbb{R}^2$. The construction of D-optimal designs for this case in the same as that of Section 7.5 by changing the signs of the parameters β_0, β_1, β_2 and β_{12}. In fact, multiplying both sides of (7.121) by -1 gives model

$$U = B_0 + B_1 x_1 + B_2 x_2 + B_{12} x_1 x_2 \tag{7.123}$$

where $U = -u$, $B_0 = -\beta_0 < 0$, $B_1 = -\beta_1 > 0$, $B_2 = -\beta_2 > 0$ and $B_{12} = -\beta_{12} < 0$. Clearly, model (7.123) is that of antagonism discussed in Section 7.5. If the design space remains the same, the support and weights remain unchanged, but the optimal logit u^* changes to $-u^*$.

7.7. Practical examples

The objective of this section is to demonstrate the real world applications of the design theories and conjectures in this chapter to constructing D-optimal designs for detecting synergy or antagonism. "Real world" means that real life experimental data which is adequately described by the two-variable binary logistic model with interaction will be used to re-design the experiments using the estimated model parameters as the values of the parameters. The second objective of this section is to show that D-optimal designs discussed in this chapter may be more efficient than designs which are used in life because of their practicality and/or convenience.

Example 7.16. The data set in Table A.2 is from LePelly and Sullivan (1936) and was reported by Greco and Lawrence (1988). Two insecticides, rotenone and pyrethrin, were sprayed either alone or jointly to a species of flies, the chrysanthemum aphids. The total number of insects exposed to each drug combination was 1000. The response of interest was the number of insects killed. The explanatory variables, d_1 and d_2, were dose concentrations in mg/cc (grams per centimeters cubed) of the two insecticides. The design appears to be a 15-point ray design, on the design space $[0, 0.350] \times [0, 2]$, with 5 support points on each of the rays $d_1 = 0$ and $d_2 = 0$, and 5 support points on ray $d_2 \simeq 5d_1$. Greco and Lawrence (1988) fitted the binary response model (2.29) in Section 2.4.2 to the data, and found that the interaction parameter was significantly positive suggesting that the two drugs act synergistically.

PROC LOGISTIC in SAS Version 9.2 was used fit the two-variable logistic model with interaction $\text{logit}(p) = u = \beta_0 + \beta_1 d_1 + \beta_2 d_2 + \beta_{12} d_1 d_2$ to the data. The estimates of the model parameters are given in Table 7.8. Table 7.8 indicates that $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$ and

Table 7.8: Parameter estimates for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ for the data of Table A.2.

Drug/combination	Parameter	Estimate	p-value
Intercept	β_0	-2.2054	< 0.0001
Rotenone	β_1	13.5803	< 0.0001
Pyrethrin	β_2	2.2547	< 0.0001
Rotenone.Pyrethrins	β_{12}	1.6261	< 0.0001

$\beta_{12} = 1.6261 > 0$. Thus, rotenone and pyrethrin interact synergically. Also, the relative potency of rotenone to pyrethrin is $\rho = \dfrac{13.5803}{2.2547} \simeq 6$.

Consider transforming model $u = \text{logit}(p) = \beta_0 + \beta_1 d_1 + \beta_2 d_2 + \beta_{12} d_1 d_2$ on the design space $\mathcal{D} = [0, 0.35] \times [0, 2]$, with the estimates in Table 7.8 taken as the model parameter values, to model $u = \text{logit}(p) = \beta_0' + \beta_1' x_1 + \beta_2' x_2 + \beta_{12} x_1 x_2$ on the design space $\mathfrak{X} = [0, 2] \times [0, 2]$ using transformations (7.12) and (7.13). This is accomplished by letting

$(x_1, x_2) = (d_1, d_2)$, $(x_{1new}, x_{2new}) = (x_1, x_2)$ and $(\beta_{0new}, \beta_{1new}, \beta_{2new}, \beta_{12new}) = (\beta_0', \beta_1', \beta_2', \beta_{12}')$, and setting $a = c = a_1 = c_1 = 0$, $b = d = d_1 = 2$, and $b_1 = 0.35$. The process obtains $\beta_0' = -2.2054$, $\beta_1' = 2.3766$, $\beta_2' = 2.2547$ and $\beta_{12}' = 0.2846$ which satisfy condition (7.39). Thus a 6-point D-optimal design is expected with two support points on line $x_1 = 0$, two support points on line $x_2 = 0$ and two interior support points on ray $x_2 = \dfrac{\beta_1'}{\beta_2'} x_1 \simeq 1.054 x_1$. The D-optimal design constructed using the Gauss program in the Appendix C is the 6-point design

$$\xi^* = \left\{ \begin{array}{cccccc} (0.320, 0) & (1.536, 0) & (0, 0.337) & (0, 1.619) & (0.149, 0.157) & (0.911, 0.960) \\ 0.1149 & 0.2475 & 0.1149 & 0.2475 & 0.0252 & 0.2500 \\ -1.446 & 1.446 & -1.446 & 1.446 & -1.489 & 2.374 \\ 0.191 & 0.809 & 0.191 & 0.809 & 0.184 & 0.915 \end{array} \right\} \tag{7.124}$$

where the last row contains the probabilities of responses at the support points. Clearly, the support points are practically located on the 3 rays $x_1 = 0$, $x_2 = 0$ and $x_2 = 1.054$. Support points located on the same logit have equal weights. The optimal logits, and in turn the probabilities of responses increase with the increase of each insecticide given alone or given in combination with the other insecticide. Hence, confirmation of synergistic effect of the two insecticides. The location of the six support points of the D-optimal design is indicated by triangles and diamonds in Figure 7.29 (a). The graph of the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ is displayed in Figure 7.29 (b), and it suggests that design (7.124) is D-optimal. The D-optimal design (7.124) can be back-transformed to the original actual doses design space $\mathcal{D} = [0, 0.35] \times [0, 2]$ using (7.12). In this case, the back-transformation of the support points is $d_1 = \dfrac{0.35}{2} x_1 = 0.175 x_1$ and $d_2 = x_2$ to which gives the following 6-point design

$$\xi^* = \left\{ \begin{array}{cccccc} (0.056, 0) & (0.269, 0) & (0, 0.337) & (0, 1.619) & (0.026, 0.157) & (0.159, 0.960) \\ 0.1149 & 0.2475 & 0.1149 & 0.2475 & 0.0252 & 0.2500 \\ -1.446 & 1.446 & -1.446 & 1.446 & -1.489 & 2.374 \\ 0.191 & 0.809 & 0.191 & 0.809 & 0.184 & 0.915 \end{array} \right\}. \tag{7.125}$$

Note that on the original design space $\mathcal{D} = [0, 0.35] \times [0, 2]$, the two interior points $(0.026, 0.157)$ and $(0.159, 0.960)$ are practically located on the interior ray $d_2 = 6d_1$.

Now, consider calculating the efficiency of the equally weighed 15-point ray design, say ξ, used to generate the data in Table A.2 relative to the D-optimal design ξ^* given by (7.125). The determinant of the information for the parameter vector $\boldsymbol{\beta} = (-2.2054, 13.5803, 2.2547, 1.6261)^T$

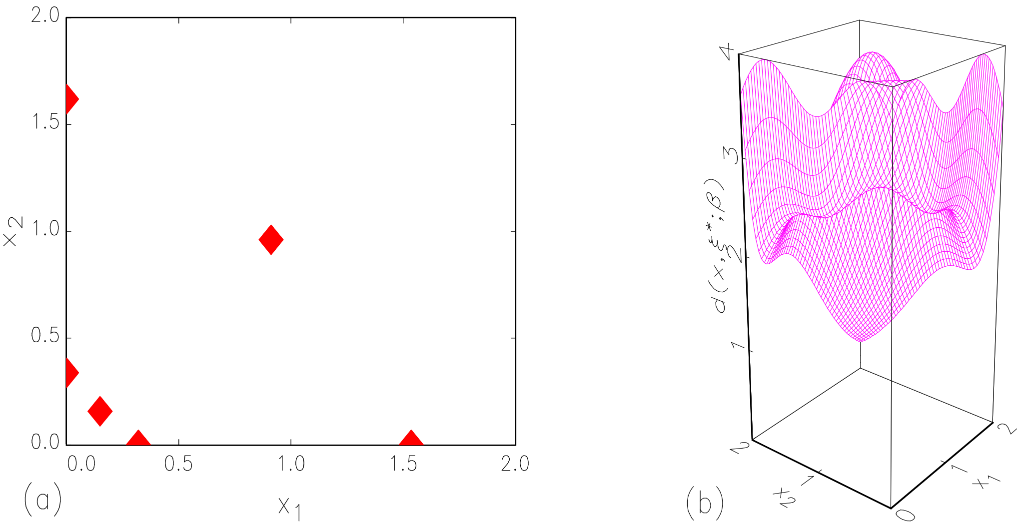


Figure 7.29: (a) Support points and (b) the standardized variance function of the D-optimal design for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0,2] \times [0,2]$ and with parameter vector $\boldsymbol{\beta} = (-2.2054, 2.3766, 2.2547, 0.2846)^T$ for the data set in Table A.2 reported by Le Pelly and Sullivan (1936).

evaluated at design ξ is $|M(\xi;\boldsymbol{\beta})| = 1.39341 \times 10^{-9}$, while the determinant of the information matrix of the same parameter vector at design ξ^* given in (7.125) is $|M(\xi^*;\boldsymbol{\beta})| = 3.18527 \times 10^{-9}$. Hence, the D-efficiency (3.38) of design ξ relative to the D-optimal design (7.125) is

$$D_{\text{eff}} = \left\{ \frac{|M(\xi;\boldsymbol{\beta})|}{|M(\xi^*;\boldsymbol{\beta})|} \right\}^{1/4} = 0.813.$$

As $D_{\text{eff}} = 0.813 < 1$, the 6-point D-optimal design (7.125) is more efficient than the 15-point ray design used to generate the data in Table A.2.

Example 7.17. The data set in Table A.3 is from Giltinan et al. (1988). Two insecticides, A and B, were sprayed either alone or jointly to tobacco budworms, known under the name of Heliothis virescens. The total number of insects exposed to each drug combination was 30. The response of interest was the number of killed budworms in 4 days after spray of insecticides. One worm in each of 3 batches was lost of follow up, making observations on 29 instead of 30 budworms in these groups. The explanatory variables, d_1 and d_2, were dose concentrations in ppm (parts per million) of respectively the two insecticides. The design appears to be a 20-point 5-ray design on the design space $[0,30] \times [0,30]$ with 4 support points on each of the rays $d_1 = 0$ and $d_2 = 0$, $d_2 = 3d_1$, $d_2 = d_1$ and $d_1 = 3d_2$. Giltinan et al. (1988)

fitted model (2.24) in Section 2.4.2 to the data, and found that the interaction parameter was significantly negative suggesting that the two insecticides act antagonistically.

PROC LOGISTIC in SAS Version 9.2 was used to fit the two-variable binary logistic model with interaction $\text{logit}(p) = u = \beta_0 + \beta_1 d_1 + \beta_2 d_2 + \beta_{12} d_1 d_2$ to the data. The estimates of the parameters are given in Table 7.9. Table 7.9 indicates that $\beta_0 < 0$, $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_{12} < -0.0092 < 0$. Therefore, A and B interact antagonistically.

Table 7.9: Parameter estimates for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ for the data of Table A.3.

Drug/Combination	Parameters	Estimates	p-value
Intercept	β_0	-2.5190	< 0.0001
A	β_1	0.1636	< 0.0001
B	β_2	0.1693	< 0.0001
$A \cdot B$	β_{12}	-0.0092	0.0005

Consider transforming the model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $\mathcal{D} = [0, 30] \times [0, 30]$, with parameter estimates in Table 7.9 taken as the model parameter values, to model $u = \text{logit}(p) = \beta_0' + \beta_1' x_1 + \beta_2' x_2 + \beta_{12} x_1 x_2$ on the design space $\mathfrak{X} = [0, 2] \times [0, 2]$ using transformations (7.12) and (7.13). This is accomplished by letting

$$(x_1, x_2) = (d_1, d_2),\ (x_{1new}, x_{2new}) = (x_1, x_2) \text{ and } (\beta_{0new}, \beta_{1new}, \beta_{2new}, \beta_{12new}) = (\beta_0', \beta_1', \beta_2', \beta_{12}'),$$

and setting $a = c = a_1 = c_1 = 0$, $b = 2$, $b_1 = d_1 = 30$. The process leads $\beta_0' = -2.5190$, $\beta_1' = 2.4540$, $\beta_2' = 2.5395$ and $\beta_{12}' = -2.0700$ to give $u = -25190 + 2.4540 x_1 + 2.5395 x_2 - 2.0700 x_1 x_2$. These parameters satisfy condition (7.79) for a candidate D-optimal design with at most 7 support points with vertex $(2, 2)$ as one of the support points. The candidate D-optimal design constructed using the Gauss program in the Appendix C is the 5-point design

$$\xi^* = \left\{ \begin{array}{ccccc} (0.450, 0) & (1.603, 0) & (0, 0.434) & (0, 1.549) & (2, 2) \\ 0.1300 & 0.2450 & 0.1300 & 0.2450 & 0.2500 \\ -1.416 & 1.416 & -1.416 & 1.416 & -0.812 \\ 0.195 & 0.805 & 0.195 & 0.805 & 0.307 \end{array} \right\}. \tag{7.126}$$

where the last row contains the probabilities of response at the support points. Clearly, the probability of response increases when the dose of each drug given alone increases. However, the probability of response is dramatically reduced with high doses of the two drugs when given together. Hence, confirmation of the antagonistic effect of the two insecticides A and B. The location of the five support points of the D-optimal design are indicated by triangles in Figure 7.30 (a). The graph of the standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ is displayed in Figure 7.30 (b), and it suggests that design (7.126) is D-optimal. The D-optimal design

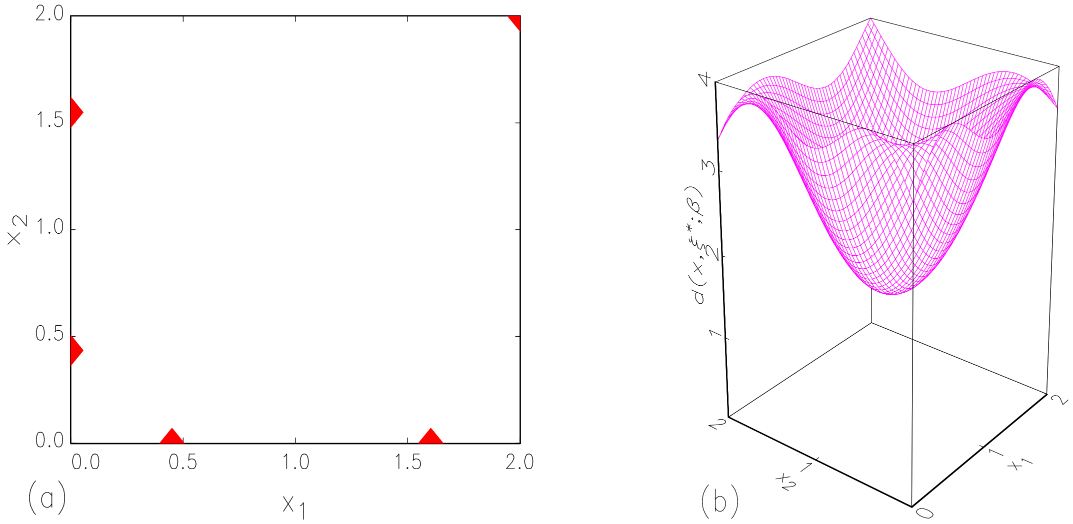


Figure 7.30: (a) Support points and (b) the standardized variance function of the D-optimal design for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $[0, 2] \times [0, 2]$ and with parameter vector $\boldsymbol{\beta} = (-2.5190, 2.4540, 2.5395, -2.0700)^T$ for the data set in Table A.3 reported by Giltinan et al. (1988).

(7.126) can be back-transformed to the original actual dose design space $\mathcal{D} = [0, 30] \times [0, 30]$ using (7.12). In this case, the back-transformation of the support points is $d_1 = \dfrac{30}{2} x_1 = 15x_1$ and $d_2 = \dfrac{30}{2} x_2 = 15x_2$ which gives the following 5-point design

$$\xi^* = \left\{ \begin{array}{ccccc} (6.750, 0) & (24.045, 0) & (0, 6.510) & (0, 23.235) & (30, 30) \\ 0.130 & 0.245 & 0.130 & 0.245 & 0.250 \\ -1.416 & 1.416 & -1.416 & 1.416 & -0.812 \\ 0.195 & 0.805 & 0.195 & 0.805 & 0.307 \end{array} \right\}. \tag{7.127}$$

Now, consider calculating the efficiency of the 20-point 5-ray design, say ξ, used to generate the data of Table A.3 relative to the D-optimal design ξ^* given by (7.127). The determinant of the information for the parameter vector $\boldsymbol{\beta} = (-2.5190, 0.1636, 0.1693, -0.0092)^T$ evaluated at design ξ is $|M(\xi; \boldsymbol{\beta})| = 2153.28$, while the determinant of the information matrix of the same parameter vector at design ξ^* given in (7.127) is $|M(\xi^*;\boldsymbol{\beta})| = 440410$. Hence, the D-efficiency (3.38) of design ξ relative to the D-optimal design (7.127) is

$$D_{\text{eff}} = \left\{ \frac{|M(\xi;\boldsymbol{\beta})|}{|M(\xi^*;\boldsymbol{\beta})|} \right\}^{1/4} = 0.264.$$

As $D_{\text{eff}} = 0.264$ is very small compared to 1, the 5-point D-optimal design (7.127) is more efficient than the 20-point ray design used to generate the data of Table A.3.

O'Brien (2004) discussed the assessment of drug interaction of the data in Table A.3 using separate ray models constructed on the basis on the Finney (1971) model for binary responses given by (2.22) in Section 2.4.1. O'Brien (2004) assessed the interaction effect of the two insecticides A and B for data on each interior ray and on $d_1 = 0$ and $d_2 = 0$ by evaluating combination indices given in Section 2.3. The results were that all the combination indices are significantly greater than 1, and hence the insecticides exhibit antagonism along all the 3 interior rays $d_2 = 3d_1$, $d_2 = d_1$ and $d_1 = 3d_2$.

Consider assessing interaction effect and constructing D-optimal designs for data in Table A.3 corresponding to each of the three interior rays supplemented by data on rays $d_1 = 0$ and $d_2 = 0$ with the two-variable binary logistic model with interaction (7.1) used in this thesis.

Ray $d_2 = 3d_1$

In this case, the estimates of the parameters of model $u$logit$(p) = \beta_0 + \beta_1 d_1 + \beta_2 d_2 + \beta_{12} d_1 d_2$ are given in Table 7.10. As the interaction parameter β_{12} is negative, but not significant (p-

Table 7.10: Parameter estimates for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 d_1 + \beta_2 d_2 + \beta_{12} d_1 d_2$ for the data of Table A.3 on ray $d_2 = 3d_1$.

Drug/Combination	Parameters	Estimates	p-value
Intercept	β_0	-2.100	< 0.0001
A	β_1	0.1415	< 0.0001
B	β_2	0.1404	< 0.0001
$A \cdot B$	β_{12}	-0.0038	0.3631

value $= 0.3631$), additivity can be concluded on ray $d_2 = 3d_1$. Estimates of parameters using the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 d_1 + \beta_2 d_2$ are given in Table 7.11. Since $\beta_0 = -2.064 > -1.5434$, $\beta_0 + 30\beta_1 = 2.06 > 1.5434$ and $\beta_0 + 30\beta_2 = 1.875 > 1.5434$, then condition (6.5) for a trapezium design is satisfied. Therefore, the D-optimal design is the 4-point trapezium design (6.5) where $b = 30$, and (λ^*, u^*) are

Table 7.11: Parameter estimates for model $u = \text{logit}(p) = \beta_0 + \beta_1 d_1 + \beta_2 d_2$ for the data of Table A.3 on ray $d_2 = 3d_1$.

Drug/Combination	Parameters	Estimates	p-value
Intercept	β_0	-2.064	< 0.0001
A	β_1	0.1370	< 0.0001
B	β_2	0.1313	< 0.0001

solutions of the simultaneous equations (5.13) and (5.18). For $\beta_0 = -2.064$, the optimal solution is $(\lambda^*, u^*) = (0.6618, 1.466)$, and hence the D-optimal corresponding to the parameters of Table 7.11 is

$$\xi^* = \left\{ \begin{array}{cccc} (4.367, 0) & (25.765, 0) & (0, 4.556) & (0, 26.884) \\ 0.1691 & 0.3309 & 0.1691 & 0.3309 \\ -1.466 & 1.466 & -1.466 & 1.466 \\ 0.188 & 0.812 & 0.188 & 0.812 \end{array} \right\}. \tag{7.128}$$

Ray $d_2 = d_1$

Parameter estimates are given in Table 7.12. Because the interaction parameter β_{12} is negative,

Table 7.12: Parameter estimates for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 d_1 + \beta_2 d_2 + \beta_{12} d_1 d_2$ for the data of Table A.3 on ray $d_2 = d_1$.

Drug/Combination	Parameters	Estimates	p-value
Intercept	β_0	-2.1080	< 0.0001
A	β_1	0.1382	< 0.0001
B	β_2	0.1414	< 0.0001
$A \cdot B$	β_{12}	-0.0102	0.0002

and significant (p-value $= 0.0002$), antagonism can be concluded on ray $d_2 = d_1$. Consider transforming the model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design space $\mathcal{D} = [0, 30] \times [0, 30]$, with parameter estimates in Table 7.12 taken as the model parameter values, to model $u = \text{logit}(p) = \beta_0' + \beta_1' x_1 + \beta_2' x_2 + \beta_{12} x_1 x_2$ on the design space $\mathfrak{X} = [0, 2] \times [0, 2]$ using transformations (7.12) and (7.13). This is accomplished by letting

$(x_1, x_2) = (d_1, d_2)$, $(x_{1new}, x_{2new}) = (x_1, x_2)$ and $(\beta_{0new}, \beta_{1new}, \beta_{2new}, \beta_{12new}) = (\beta_0', \beta_1', \beta_2', \beta_{12}')$, and setting $a = c = a_1 = c_1 = 0$, $b = 2$, $b_1 = d_1 = 30$. The process leads $\beta'_0 = -2.1080$, $\beta'_1 = 2.073$, $\beta'_2 = 2.121$ and $\beta'_{12} = -2.295$. The parameters of this model satisfy condition (7.77) of an 8-point D-optimal design. The D-optimal design constructed using the Gauss program in Appendix C for data on rays $d_1 = 0$, $d_2 = 0$ and $d_2 = d_1$ of the data in Table A.3 is indeed the following 8-point design.

$$\xi^* = \left\{ \begin{array}{cccccccc} (5.80, 0) & (24.71, 0) & (0, 5.67) & (0, 24.15) & (4.93, 30) & (30, 4.44) & (30, 20.32) & (20.51, 30) \\ 0.1165 & 0.0803 & 0.1341 & 0.0822 & 0.1618 & 0.1625 & 0.1285 & 0.1341 \\ -1.307 & 1.307 & -1.307 & 1.307 & 1.307 & 1.307 & -1.307 & -1.307 \\ 0.213 & 0.787 & 0.213 & 0.787 & 0.787 & 0.787 & 0.213 & 0.213 \end{array} \right\}. \tag{7.129}$$

Ray $d_1 = 3d_2$

The parameter estimates are given in Table 7.13. The interaction parameter β_{12} is negative,

Table 7.13: Parameter estimates for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 d_1 + \beta_2 d_2 + \beta_{12} d_1 d_2$ for the data of Table A.3 on ray $d_1 = 3d_2$.

Drug/Combination	Parameters	Estimates	p-value
Intercept	β_0	-1.9310	< 0.0001
A	β_1	0.1319	< 0.0001
B	β_2	0.1366	< 0.0001
$A \cdot B$	β_{12}	-0.0065	0.0976

but not quite significant (p-value = 0.0976). Hence, additivity can be concluded on ray $d_1 = 3d_2$. Estimates of parameters using the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 d_1 + \beta_2 d_2$ are given in Table 7.14. Since $\beta_0 = -1.8676 > -1.5434$,

Table 7.14: Parameter estimates for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 d_1 + \beta_2 d_2$ for the data of Table A.3 on ray $d_1 = 3d_2$.

Drug/Combination	Parameters	Estimates	p-value
Intercept	β_0	-1.8676	< 0.0001
A	β_1	0.1159	< 0.0001
B	β_2	0.1284	< 0.0001

$\beta_0 + 30\beta_1 = 1.6094 > 1.5434$, and $\beta_0 + 30\beta_2 = 1.9844 > 1.5434$, then condition (6.5) is verified and therefore, and the D-optimal design is the 4-point trapezium design (6.5) where

$b = 30$, and $(\lambda^*, u^*) = (0.6646, 1.492)$ solutions of the simultaneous equations (5.13) and (5.18). Hence, the D-optimal corresponding to the parameters of Table 7.14 is

$$\xi^* = \left\{ \begin{array}{cccc} (3.239, 0) & (28.989, 0) & (0, 2.923) & (0, 26.167) \\ 0.1677 & 0.3323 & 0.1677 & 0.3323 \\ -1.492 & 1.492 & -1.492 & 1.492 \\ 0.184 & 0.816 & 0.184 & 0.816 \end{array} \right\}. \tag{7.130}$$

7.8. Conclusions

In this chapter, the construction of the D-optimal designs for the two-variable binary logistic model with interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ on the design spaces $(-\infty, \infty) \times [c, d]$, $[0, \infty) \times [0, \infty)$ and $[0, b] \times [0, b]$ was discussed. The equally weighted 4-point, with support points on two pairs of hyperbolae $\pm u = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$, introduced by Jia and Myers (2001) was proved to be D-optimal. An alternative 4-point D-optimal design on the design space $\mathfrak{X} = (-\infty, \infty) \times [c, d]$, also with support points on the hyperbolae $\pm u$, was introduced in this chapter. The design was constructed from a candidate 4-point D-optimal whose support points were the points of intersection of the hyperbolae $u_1 = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ and $u_2 = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ and the boundary lines $x_2 = c$ and $x_2 = d$ of the design space $\mathfrak{X}$. This approach of constructing the 4-point D-optimal design was extended to constructing D-optimal designs on the $[0, \infty) \times [0, \infty)$ and the $[0, b] \times [0, b]$ design spaces. The number of support points of the D-optimal designs on the design $[0, b] \times [0, b]$ space, which can easily be transformed to a rectangular design space $[a, b] \times [c, d]$, varied from 4 to 6 for synergy and from 4 to 8 for antagonism depending on the size of the design space and the model parameter values. Because the interaction term in the model vanishes at design support points on the boundaries of the $[0, \infty) \times [0, \infty)$ and the $[0, b] \times [0, b]$ design spaces, D-optimal designs without interior support points are singular for estimating the interaction parameter β_{12}. This problem was resolved by choosing candidate D-optimal designs with one or two interior support points defined by the intersection of the interior ray $x_2 = \dfrac{\beta_1}{\beta_2} x_1$ and one or two branches of the hyperbolae u_1 and u_2. Simple D-optimal designs such as equally weighted 4-point and 8-point designs were constructed semi-analytically and semi-numerically in simple cases such as those of equally weighted 4-point and 8-point designs. Otherwise, conditions for the existence of the design patterns of the candidate D-optimal designs were derived, then the D-optimal designs for specific examples were constructed numerically, and the D-optimality of these designs was checked graphically.

Two practical examples, one on synergy and another on antagonism, were used to illustrate the usefulness of the design strategy adopted in this chapter in terms of their efficiency and cost reduction due to the use of fewer experimental runs over designs used to generate the data.

8

Conclusions

In this thesis, D-optimal designs for the two-variable binary logistic model without and with interaction for detecting drug interaction were investigated.

In Chapter 4, the D-optimal design for the one-variable binary logistic model $u = \text{logit}(p) = \beta_0 + \beta_1 x$ on the design space $\mathbb{R}$ which is known to have two support points on logit $u = \pm 1.5434$ was shown to be D-optimal. The D-optimality of the 2-point design for the one-variable binary logistic model on the design space $\mathbb{R}$ was derived since it is relevant to the construction of D-optimal designs for the two-variable binary logistic model with interaction.

In Chapter 5, theoretical constructions of D-optimal designs for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the design spaces $(-\infty, \infty) \times [-b, b]$ and $[0, \infty) \times [0, \infty)$ were discussed. For u unrestricted, the equally weighted 4-point parallelogram candidate D-optimal design, with support on the logit lines $u = \pm 1.22291$, of Sitter and Torsney (1995a), Jia and Myers (2001) also found in Myers et al. (2002), and in Atkinson and Haines (1996), was proved analytically to be D-optimal on the design space given by the (u_1, u_2)-space $(-\infty, \infty) \times [-b, b]$ where $u_1 = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ and $u_2 = b_0 + b_1 x_1 + b_2 x_2$. Two new D-optimal designs, for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the (x_1, x_2) design space $[0, \infty) \times [0, \infty)$, a 4-point trapezium design and a 3-point design respectively associated with the conditions about the intercept parameter β_0, $-\infty < \beta_0 < -1.5434$ and $-1.5434 \leq \beta_0 \leq 0$, were introduced and proved to be D-optimal.

In Chapter 6, D-optimal designs for the two-variable binary logistic model without interaction $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ on the bounded design space $[0, b] \times [0, b]$ were semi-analytically and semi-numerically investigated. In some cases, the D-optimal designs were found to be similar to the parallelogram and the trapezium D-optimal designs discussed in Chapter 5, and in other cases the design patterns were very different. The candidate D-optimal designs termed trapezium designs and non-trapezium designs satisfy the necessary condition that either $\beta_0 + \beta_1 b > 0$ and $\beta_0 + \beta_2 b > 0$ or $\beta_0 + \beta_1 b < 0$ and $\beta_0 + \beta_2 b < 0$. The candidate D-optimal designs termed parallelogram designs and non-parallelogram designs satisfy the necessary condition that the signs of $\beta_0+\beta_1 b$ and $\beta_0+\beta_2 d$ alternate. The case when $\beta_0+\beta_1 b = 0$ or $\beta_0 + \beta_2 b = 0$ lead to candidate D-optimal designs with 4 to 6 support points. A practical example on the joint effect of two insecticides was used to show that optimal designs can be less expensive in terms of experimental runs, and more efficient than some well known experimental designs.

In Chapter 7, D-optimal designs for the two-variable binary logistic model with interaction, $u = \text{logit}(p) = \beta_0+\beta_1 x_1+\beta_2 x_2+\beta_{12} x_1 x_2$, on the design spaces $(-\infty, \infty)\times[c, d]$ and $[0, b]\times[0, b]$ were investigated. In the case of the design space $(-\infty, \infty) \times [c, d]$, candidate equally weighted 4-point D-optimal designs were suggested and calculations indicated designs with support points located on the hyperbolic logit lines $u = \pm 1.5434$. Then, the design was proved to be optimal algebraically. In the case of the design space $[0, b] \times [0, b]$, the candidate D-optimal designs were found to have 4 to 6 support points for synergistic effect and 4 to 8 support points for antagonistic effect. The candidate D-optimal designs were proved to be optimal algebraically in the simplest cases of equally weighted 4- or 8-point designs. In complicated cases with unequally weighted design points, the patterns of various candidate D-optimal designs were conjectured and the conditions of existence of these design patterns were derived algebraically. Then D-optimality of these designs was checked graphically by plotting the standardized variance function. Real world examples were used to demonstrate the usefulness of the D-optimal designs and their superiority relative to some designs generally used in experimentation.

The results of this thesis on constructing D-optimal designs for the two-variable binary logistic models without and with interaction on a subset of the space $\mathbb{R}^2$ can be extended in further research, at least numerically, to binary logistic models with more than two explanatory variables on a subset of $\mathbb{R}^k$ where k is an integer greater than or equal to 3. Future studies can also be done by constructing optimal designs for other models containing two or more explanatory variables such as the Finney models reviewed in Chapter 2. Furthermore in order to avoid problems of parameter misspecification, the Bayesian or minimax methodologies can be used in constructing D-optimal designs for nonlinear and generalized linear models with two or more explanatory variables.

Bibliography

Abdelbasit, K. M. and Plackett, R. L. (1982). Experimental design for joint action, *Biometrics* **38**: 171–179.

Abdelbasit, K. M. and Plackett, R. L. (1983). Experimental design for binary data, *Journal of the American Statistical Association* **78**(381): 90–98.

Atkinson, A. C. (2006). Chap. 8: Generalized linear models and response transformation, *in* A. I. Khuri (ed.), *Response surface methodology and related topics*, World Scientific, New Jersey.

Atkinson, A. C., Bogacka, B. and Zhigljavsky, A. (2001). *Optimal design 2000*, Kluwer, Dordrecht.

Atkinson, A. C. and Donev, A. N. (1992). *Optimum Experimental Designs*, Oxyford University Press, New York.

Atkinson, A. C., Donev, A. N. and Tobias, R. D. (2007). *Optimum Experimental Designs, with SAS*, Oxyford University Press, New York.

Atkinson, A. C., Hackl, P. and Muller, W. G. (2001). *MODA 6 - Advances in Model-oriented Design and Analysis*, Physica-Verlag, Heidelberg.

Atkinson, A. C. and Haines, L. M. (1996). Chap. 11: Designs for nonlinear and generalized linear models, *in* S. Ghosh and C. R. Rao (eds), *Handbook of Statistics*, Vol. 13, B. V. Elsevier Science, Amsterdam, pp. 437–475.

Atkinson, A. C., Pronzato, L. and Wynn, H. P. (1998). *MODA 5 - Advances in Model-oriented Data Analysis and Experimental Design*, Physica-Verlag, Heidelberg.

Berenbaum, M. C. (1977). Synergy, additivity and antagonism in immunosuppression. a critical review, *Clin. Exp. Immunol.* **28**: 1–18.

Berenbaum, M. C. (1981). Criteria for analyzing interactions between biologically active agents, *Journal of the American Statistician Association* **78**: 91–98.

Berenbaum, M. C. (1989). What is synergy?, *The American Society for Pharmacology and Experimental Therapeutics* **41**: 93–141.

Bliss, C. I. (1939). The toxicity of poisons applied jointly, *Annals of Applied Biology* **26**: 585–615.

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

Bucchianico, A. D., Lauter, H. and Wynn, H. P. (2004). *MODA 7 - Advances in Model-oriented Design and Analysis*, Physica-Verlag, Heidelberg.

Burridge, J. and Sebastiani, P. (1992). Optimal designs for generalized linear models, *Journal of the Italian Statistical Society* **1**(2): 183–202.

Burridge, J. and Sebastiani, P. (1994). D-optimal designs for generalised linear models with variance proportional to the square of the mean, *Biometrika* **81**(2): 295–304.

Chaloner, K. and Lantz, K. (1989). Optimal bayesian design applied to logistic regression experiments, *Journal of Statistical Planning and Inference* **21**: 191–208.

Chaloner, K. and Verdinelli, I. (1995). Bayesian experimental design: a review, *Statistical Science* **10**: 273–304.

Chernoff, H. (1953). Locally optimal designs for estimating parameters, *Annals of Mathematical Statistics* **24**: 586–602.

Chernoff, H. (1979). *Sequential Analysis and Optimal Design*, SIAM, Philadelphia.

Chipman, H. A. and Welch, J. (1996). D-optimal design for generalized linear models. Unpublished report.

Chou, T. C. and Talalay, P. (1984). Quantitative analysis of dose-effect relationship: the combined effects of multiple drugs or enzyme inhibitors, *Adv. Enzyme Reg.* **22**: 27–55.

Collett, D. (2003). *Modelling Survival Data in Medical Research*, 2nd edn, Chapman and HALL/CRC, Boca Raton, Florida.

Dette, H. and Studden, W. J. (1997). *The theory of Canonical Moments with Applications in Statistics, Probability and Analysis*, Willey, New York.

Dobson, J. A. (2002). *An Introduction to Generalised Linear Models*, second edn, Chapman and HALL, Boca Raton, Florida.

Dodge, Y., Fedorov, V. V. and Wynn, H. P. (1988). *Optimal Design and Analysis of Experiments*, Elsevier, Amsterdam.

Elfving, G. (1952). Optimal allocation in linear regression theory, *Annals of Mathematical Statistics* **23**: 255–262.

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

Fedorov, V. V. and Hackl, P. (1997). *Model-oriented Design of Experiments. Lecture Notes in Statistics 125*, Springer, New York.

Fedorov, V. V., Müller, W. G. and Vuchnov, I. N. (1992). *MODA2 - Advances in Model-oriented Data Analysis*, Physica-Verlag, Heidelberg.

Finney, D. J. (1971). *Probit Analysis*, third edn, Cambridge University Press, Cambridge.

Flournoy, N., Rosenberger, W. F. and Wong, W. K. (1998). *New Developments and Applications in Experimental Designs*, IMS, Hayward.

Ford, I., Torsney, B. and Wu, C. F. (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.

Geric, T. M., Blum, U. and Meier, K. (1988). Statistical analysis of the joint inhibitory action of similar compounds, *Journal of Chemical Ecology* **15**(10): 2403–2412.

Ghosh, S. and Rao, C. R. (eds) (1996). *Handbook of Statistics 13: Design and Analysis of Experiments*, Elsevier, Amsterdam.

Giltinan, D. M., Capizzi, T. P. and Malani, H. (1988). Diagnostic tests for similar action of two compounds, *Applied Statistics* **37**: 39–50.

Graybill, A. F. (1983). *Matrices with applications in Statistics*, second edn, Wodworth, Belmont.

Greco, W. R., Bravo, G. and Parsons, J. C. (1995). The search for synergy: a critical review from a response surface perspective, *Pharmacological Reviews* **47**: 331–385.

Greco, W. R. and Lawrence, D. D. (1988). Assessment of the degree of drug interaction where the response variable is discrete, *American Statistician Association, Proceedings of the Biopharmaceutical Section*, pp. 226–231.

Greco, W. R. and Tung, Y. (1991). D-optimal experimental designs for quantifying synergy, *American Statistician Association, Proceedings of the Biopharmaceutical Section*, pp. 244–249.

Guest, P. G. (1958). The spacing of observations in polynomial regression, *Annals of Mathematical Statistics* **29**: 294–299.

Haines, L. M. (1987). The application of the annealing algorithm to the construction of exact d-optimal designs for linear regression models, *Technometrics* **29**: 439–447.

Haines, L. M. (1993). Optimal design for nonlinear regression models, *Communication statistics-Theory methodology* **22**(6): 1613–1627.

Haines, L. M. (1995). A geometric approach to optimal design for one-parameter non-linear models, *Journal of the Royal Statistical Society. Series B (Methodological)* **57**(3): 575–598.

Haines, L. M., Kabera, M. G., Ndlovu, P. and O'Brien, T. E. (2007). D-optimal designs for logistic regression in two variables, *in* J. Lopez-Fidalgo, J. M. Rodrigez-Diaz and B. Torsney (eds), *MODA8 - Advances in Model-Oriented Designs and Analysis*, Physica-Verlag, Heidelberg, pp. 91–98.

Hewlett, P. S. (1969). Measurement of the potencies of drug mixtures, *Biometrics* **25**: 477–486.

Hewlett, P. S. and Plackett, R. L. (1959). A unified theory for quantal responses to mixtures of drugs: non-interactive action, *Biometrics* **15**: 591–610.

Hoel, P. G. (1958). Efficiency problems in polynomial estimation, *Annals of Mathematical Statistics* **29**: 1134–1145.

Hosmer, D. W. and Lemeshow, S. (2000). *Applied Logistic Regression*, second edn, John Wiley and Sons, INC., New York.

Jia, Y. and Myers, R. (2001). Optimal experimental designs for two-variable logistic models. Preprint.

Joshi, M. C. and Moudgalya, K. M. (2004). *Optimization: Theory and Practice*, Morgan and Claypool, London.

Kiefer, J. (1958). On the nonrandomized optimality and randomized nonoptimality of symmetrical designs, *Annals of Mathematical Statistics* **29**: 675–699.

Kiefer, J. (1959). Optimum experimental designs, *Journal of the Royal Statistical Society, Series B* **21**: 272–319.

Kiefer, J. (1974). General equivalence theory for optimum designs (approximate theory), *Annals of Statistics* **2**: 849–879.

Kiefer, J. and Wolfowitz, J. (1960). The equivalence of two extremum problems, *Canadian Journal of Mathamatics* **12**: 363–366.

King, J. and Wong, W. K. (2000). Minimax d-optimal designs for the logistic model, *Biometrics* **56**: 1263–1267.

Kitsos, C. P. and Müller, W. G. (1995). *MODA 4 - Advances in Model-oriented Data Analysis*, Physica-Verlag, Heidelberg.

Kupchak, P. I. (2000). *Optimal Designs for the Detection of Drug Interaction*, PhD thesis, Graduate Department of Statistics, University of Toronto.

LePelly, R. H. and Sullivan, W. N. (1936). Toxicity of rotenone and pyrethrin, alone and in combination, *J. Econ. Ent.* **29**: 791–796.

Liski, E. P., Mandal, N. K., Shah, K. R. and Sinha, B. K. (2001). *Topics in Optimal Design*, Springer, New York.

Loewe, S. (1953). The problem of synergyism and antagonism of combined drugs, *Arzneittelforsch* **3**: 285–320.

López-Fidalgo, J., Rodriguez-Diaz, J. M. and Torsney, B. (2007). *MODA8 - Advances in Model-oriented Design and Analysis*, Physica-Verlag, Heidelberg.

Machado, S. and Robinson, G. A. (1994). A direct, general approach based on isobolograms for assessing the joint action of drugs in pre-clinical experiments, *Statistics in Medicine* **13**: 2289–2309.

Martin, J. T. (1942). The problem of the evaluation of rotenone-containing plants: The toxicity of l-elliptone and of poisons applied jointly, with further observations on the rotenone equivalent method of assigning the toxicity of derris root, *Annals of Applied Biology* **29**: 69–81.

McCullagh, P. and Nelder, J. A. (1989). *Generalised Linear Models*, second edn, Chapman and HALL, London.

Meadows, S. L., Gennings, C., Carter, W. H. J. and Bae, D. S. (2002). Experimental design for mixtures of chemicals along fixed ratio rays, *Environmental Health Perspectives* **110**(6): 979–983.

Mitchell, T. J. (1974). An algorithm for the construction of d-optimal experimental designs, *Technometrics* **16**: 203–210.

Müller, W. G. (1998). *Collecting Spatial Data. Optimal Design for Random Fields*, Heidelberg Physica-Verlag.

Müller, W. G., Wynn, H. P. and Zhigljavsky, A. A. (1993). *MODA 3 - Advances in Model-oriented Data Analysis and Experimental Design*, Physica-Verlag, Heidelberg.

Myers, R. H. and Montgomery, D. C. (1995). *Response Surface Methodology: Process and Product Optimization Using Design Experiments*, John Wiley and Sons, INC., New York.

Myers, R. H., Montgomery, D. C. and Vining, G. G. (2002). *Generalized Linear Models*, John Wiley and Sons, INC., New York.

Nelder, J. A. and Mead, R. (1965). A simplex method for function minimization, *Computer Journal* **7**: 308–313.

Nelder, J. A. and Wedderburn, R. W. M. (1972). Generalized linear models, *Journal of the Royal Statistical Sociaety, Series A* **135**(370–384).

O'Brien, T. E. (2004). Modeling and design to detect interaction of insecticides, herbicides and other similar compounds, *in* D. Johnson and G. Milliken (eds), *Proceedings of the 15th Conference on Applied Statistics in Agriculture*, pp. 303–321.

O'Brien, T. E. and Funk, G. M. (2003). A gentle introduction to optimal design for regression models, *American Statistician* **57**(4): 265–267.

Pázman, A. (1986). *Foundations of Optimum Experimental Design*, Reidel, Dordrecht.

Plackett, R. L. and Hewlett, P. S. (1952). Quantal responses to mixtures of poisons, *Journal of the Royal Statistical Society, Series B* **14**(2): 141–163.

Plackett, R. L. and Hewlett, P. S. (1963). A unified theory for quantal responses to mixtures of drugs: the fitting to data of certain models for two non-interactive drugs with complete positive correlation of tolerances, *Biometrics* **19**: 517–531.

Pukelsheim, F. (1993). *Optimal Design of Experiments*, John Wiley and Sons, New York.

Schwabe, R. (1996). *Optimal Designs for Multi-factor Models*, Springer, New York.

Sebastiani, P. and Settimi, R. (1997). A note on d-optimal designs for a logistic regression model, *Journal of Statistical Planning and Inference* **59**: 359–368.

Seber, G. A. F. and Wild, C. J. (2003). *Nonlinear Regression*, John Wiley and Sons, New Jersey.

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

Sibson, R. (1972). Discussion of paper by h. p. wynn, *Journal of the Royal Statistical Society, Series B* **34**: 181–183.

Sibson, R. (1974). d_a-optimality and duality. progress in statistics, *Colloq. Math.Soc.Janos Bolyai* **9**: 677–692.

Silvey, S. D. (1978). Optimal design measures with singular information matrices, *Biometrika* **65**(3): 553–559.

Silvey, S. D. (1980). *Optimal Design: An Introduction to the Theory for Parameter Estimation*, Chapman and HALL, London.

Silvey, S. D. and Titterington, D. M. (1972). A geometric approach to optimal design, *Biometrika* **60**(1): 21–32.

Sitter, R. R. (1992). Robust designs for binary data, *Biometrics* **48**: 1145–1155.

Sitter, R. R. and Fainaru, I. (1997). Optimal designs for logit and probit models for binary data, *The Canadian Journal of Statistics* **25**(2): 175–190.

Sitter, R. R. and Forbes, B. E. (1997). Optimal two-stage designs for binary response experiments, *Statistica Sinica* **7**: 941–955.

Sitter, R. R. and Torsney, B. (1995a). Optimal design for binary response experiments with two design variables, *Statistica Sinica* **5**: 405–419.

Sitter, R. R. and Torsney, B. (1995b). D-optimal designs for generalized linear models, *in* C. P. Kitsos and W. G. Müller (eds), *Proceedings of MODA 4*, Physica Verlag, Heidelberg, pp. 87–102.

Sitter, R. R. and Wu, C. F. J. (1993). Optimal designs for binary responses experiments: Fieller, d, and a criteria, *Scand. J. Statist.* **20**: 329–341.

Smith, D. M. and Ridout, M. S. (2003). Optimal designs for criteria involving log(potency) in comparative binary bioassys, *Journal of Statistical Planning and Inference* **113**: 617–632.

Smith, K. (1918). On the standard deviation of adjusted and interpolated values of an observed polynomial function and its constants and guidance they give towards a proper choice of the distribution of observations, *Biometrika* **64**: 1–85.

Stokes, M. E., Davis, C. S. and Koch, G. G. (2000). *Categorical Data Analysis Using the SAS System*, second edn, Cary, North Carolina.

Tallarida, J. R. (2000). *Drug Synergism and Dose-effect Data Analysis*, Chapman and Hall/CRC, Boca Raton, Florida.

Torsney, B. and Gunduz, N. (2001). Chap. 24: On optimal designs for high dimensional binary regression models, *in* A. C. Atkinson, B. Bogacka and A. Zhigljavsky (eds), *Optimal Design 2000*, Kluwer, Dordrecht.

Torsney, B. and Musrati, A. K. (1993). On the construction of optimal designs with applications to binary response and to weighted regression models, *in* W. G. Müller, H. P. Whynn and A. A. Zhigljavsky (eds), *Moda3 - Advances in Model-oriented Data Analysis and Experimental Design*, Physisca-Verlag, Heidelberg.

Wang, Y. (2002). *Optimal Experimental Designs for the Poisson Regression Model in Toxicity Studies*, PhD thesis, Faculty of the Virginia Polytechnic Institute and State University.

Wang, Y., Myers, R., Smith, E. and Ye, K. (2006). D-optimal designs for poisson regression models, *Journal of Statistical Planning and Inference* **136**: 2831–2845.

Welch, W. J. (1982a). Algorithmic complexity: three np-hard problems in computational statistics, *J. statist. comput. simulation* **15**: 17–25.

Welch, W. J. (1982b). Branch-and-bound search for experimental designs based on d-optimality and other criteria, *Technometrics* **24**(1): 41–48.

White, L. W. (1973). An extension of the general equivalence theorem to nonlinear models, *Biometrika* **60**(2): 345–348.

White, L. W. (1975). *The Optimal Design of Experiments for Estimation in Nonlinear Model*, PhD thesis, University of London.

Whittle, P. (1973). Some general points in the theory of optimal experimental design, *Journal of the Royal Statistical Society. Series B (Methodology)* **35**(1): 123–130.

Wu, C. F. J. (1988). Optimal design for percentile estimation of a quantal response curve, *in* Y. Dodge, V. V. Fedorov and H. P. Wynn (eds), *Optimal design and analysis of experiments*, Elsvier North Holland, Amsterdam, pp. 213–223.

A

Data sets

Table A.1: Data from Martin (1942): the total number and the number of dead chrysanthemum aphides due to joint application of rotenone and deguelin concentrates.

Rotenone (mg/l)	Deguelin (mg/l)	Total number of insects	Number of dead insects
10.2	0.0	50	44
7.7	0.0	49	42
5.1	0.0	46	24
3.8	0.0	48	16
2.6	0.0	50	6
0.0	50.5	48	48
0.0	40.4	50	47
0.0	30.3	49	47
0.0	20.2	48	34
0.0	10.1	48	18
5.1	20.3	50	48
4.0	16.3	46	43
3.0	12.2	48	38
2.0	8.1	46	27
1.0	4.1	46	22
0.5	2.0	47	7

Table A.2: Data from Le Pelly and Sullivan (1936): the total number and number of dead houseflies due to joint application of rotenone and pyrethrin concentrates.

Rotenone (mg/l)	Pyrethrin (mg/l)	Number of insects	Number of dead insects
0.100	0.0	1000	240
0.150	0.0	1000	440
0.200	0.0	1000	630
0.250	0.0	1000	810
0.350	0.0	1000	900
0.0	0.500	1000	200
0.0	0.750	1000	350
0.0	1.000	1000	530
0.0	1.500	1000	800
0.0	2.000	1000	880
0.050	0.250	1000	270
0.075	0.375	1000	530
0.100	0.500	1000	640
0.146	0.729	1000	820
0.196	0.979	1000	930

Table A.3: Data from Giltinan, Capizzi, and Malani (1988): Mortality in response to two insecticides A and B.

A (ppm)	B (ppm)	Number	Dead
0.000	30.000	30	26
0.000	15.000	30	19
0.000	7.500	30	7
0.000	3.750	30	5
6.500	019.500	30	23
3.250	9.750	30	11
1.625	4.875	30	3
0.813	2.438	30	0
13.000	13.000	30	15
6.500	6.500	30	5

A (ppm)	B (ppm)	Number	Dead
3.250	3.250	29	4
1.625	1.325	29	0
19.500	6.500	30	20
9.750	3.250	30	13
4.875	1.625	29	6
2.438	0.813	30	0
30.000	0.000	30	23
15.000	0.000	30	21
7.500	0.000	30	13
3.750	0.000	30	5

B

Gauss program for the construction of D-optimal designs for the two-variable binary logistic model without interaction on a rectangular design space

@ — D-optimal designs for model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ — @
@ — on the rectangular design space $[a, b] \times [c, d]$ — @
output file=optdes.out reset;
library optmum,pgraph;
include optmum.ext;
graphset;
format 15,6;

@—Rectangular design space—@
$a = 0$; $b = 2$; $c = 0$; $d = 2$;
" "; "Minimum and maximum x_1 values" $a \sim b$;
" "; "Minimum and maximum x_2 values" $c \sim d$;

@—Optimality criterion—@
" "; "Minimizing -ln(det)";
mult=1;

@ — Best guess of parameters and starting design — @
/*Martin's data*/
b0=-2.0287;

```
b1=2.1808;
b2=3.9239;
npts=4;
dmat={0.2 0,0 0.2,0 1,1.5 0};
wvec=ones(npts,1)/npts;
dwmat=dmat~wvec;
" ";"Best guess of the parameters b0 b1 b2" b0~b1~b2;
"Starting design";
"Number of support points = " npts;
"Support points and weights = " dwmat;

@ — Grid over the design space [a,b] × [c,d] — @
ng=51;
dx1=(b-a)/(ng-1);
dx2=(d-c)/(ng-1);
gridx1=seqa(a,dx1,ng);
gridx2=seqa(c,dx2,ng);
grid=gridgen(ng);
ntop=16;


@ — Optimization process — @
y0=constr(dwmat);
{ y,f,g,h } = optmum(&fun,y0);
dwmat=unconstr(y);
y=constr(dwmat);
crit=fun(y);
infomatopt=infomat(dwmat);
invmat=inv(infomatopt);

@ — results — @
" ";"optimal design";
"Number of support points = " npts;
"Support points and weights = " dwmat;
" ";"Criterion value " crit;
" ";"Optimum information matrix " infomatopt;
" ";"Inverse of the optimum information matrix " invmat;
```

@ — Standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ at the support points — @
stdvgrid=stdvarf(infomatopt,dwmat[.,1:2]);
expu=exp(b0+b1*dwmat[.,1]+b2*dwmat[.,2]);
pvec=expu./(1+expu);
lvec=ln(pvec./(1-pvec));
" ";"Optimal design, p and logit(p) at support points" stdvgrid $\sim$ pvec $\sim$ lvec;

@ — Plot of $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ over a grid — @
stdvgrid=stdvarf(infomatopt,grid);
matcont=reshape(stdvgrid[.,3],ng,ng);
_pdate = 0;
fonts("simplex complex microb simgrma");
ztics(0,3,1,0.5);
xtics(0,2,0.5,0.1);
ytics(0,2,0.5,0.1);
zlabel("$d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$");
xlabel("x]1[");
ylabel("x]2[");
_pnum = 2;
_pnumht = 0.25;
_paxht = 0.2;
_pmcolor = $\{0, 0, 0, 0, 0, 0, 0, 0, 15\}$;
_pmsgctl = $\{0.8\ 0.7\ 0.3\ 0\ 1\ 0\ 0\}$;
/*_pmsgstr="(b)";
/ surface(gridx1',gridx2,matcont');
stdvgrid=rev(sortc(stdvgrid,3));
ddmax=stdvgrid[1:ntop,.];
" ";"maximum $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ values over the rectangle $[a, b] \times [c, d]$ " ddmax;

@ — Procedure - Function for minimization – @
proc fun(y);
local dwmat,imat,crit;
dwmat=unconstr(y);
imat=infomat(dwmat);
crit=-mult*ln(det(mmat));

```
retp(crit);
endp;
```

@ — Procedure to generate grid on the rectangle $[a, b] \times [c, d]$ — @

```
proc gridgen(ng);
local grid,k,i,j,toler;
toler=0.0000000001;
grid=zeros(ng*ng,2);
k=1;
i=1;
do until i > ng;
j=1;
do until j > ng;
if(abs(gridx1[i]) < toler);
gridx1[i]=0;
endif;
if(abs(gridx2[j]) < toler);
gridx2[j]=0;
endif;
grid[k,.]=gridx1[i]~gridx2[j];
k=k+1;
j=j+1;
endo;
i=i+1;
endo;
retp(grid);
endp;
```

@ — Procedure - Information matrix — @

```
proc infomat(dwmat);
local n,dmat,wvec,x1,x2,expu,vterm,xmat,wmat,imat;
n=rows(dwmat);
dmat=dwmat[.,1:2];
wvec=dwmat[.,3];
x1=dmat[.,1];
x2=dmat[.,2];
```

```
expu=exp(b0+b1*x1+b2*x2);
vterm=sqrt(expeta)./(1+expeta);
xmat=vterm.*(ones(n,1)~x1~x2);
wmat=diagrv(zeros(n,n),wvec);
imat=xmat'.*wmat.*xmat;
retp(imat);
endp;

@ — Procedure - Unconstrained x and λ to constrained y— @
proc constr(dwmat);
local y,dmat,i,wvec,den;
y=zeros(3*npts-1,1);
/* transform support points */
dmat=dwmat[.,1:2];
i=1;
do until i>npts;
y[i,1]=arcsin(sqrt((dmat[i,1]-a)/(b-a)));
y[i+npts,1]=arcsin(sqrt((dmat[i,2]-c)/(d-c)));
i=i+1;
endo;
/* transform weights */ wvec=dwmat[.,3];
den=1.0;
i=1;
do until i>npts-1;
y[i+2*npts,1]=arcsin(sqrt(wvec[i]/den));
den=den-wvec[i];
i=i+1;
endo;
retp(y);
endp;

@ — procedure - constrained y to unconstrained x and λ— @
proc unconstr(y);
local dmat,i,wvec,yw,mult,dwmat;
/* transform support points */
dmat=zeros(npts,2);
```

```
i=1;
do until i>npts;
dmat[i,1]=x1min+(x1max-x1min)*sin(y[i,1])*sin(y[i,1]);
dmat[i,2]=x2min+(x2max-x2min)*sin(y[i+npts,1])*sin(y[i+npts,1]);
i=i+1;
endo;
/* transform weights */
wvec=zeros(npts,1);
yw=y[2*npts+1:3*npts-1,.];
mult=1;
i=1;
do until i>npts-1;
wvec[i,1]=sin(yw[i,1])*sin(yw[i,1])*mult;
mult=mult*cos(yw[i,1])*cos(yw[i,1]);
i=i+1;
endo;
wvec[npts,1]=mult;
dwmat=dmat~wvec;
retp(dwmat);
endp;

@ — Procedure to calculate the standardized variance function — @
proc stdvarf(mopt,grid);
local ngrid,stdvgrid,i,matx;
ngrid=rows(grid);
stdvgrid=grid zeros(ngrid,1);
i=1;
do until i > ngrid;
matx=infomat(grid[i,.]~1);
ddgrid[i,3]=sumc(diag(inv(mopt)*matx));
i=i+1;
endo;
retp(stdvgrid);
endp;
```

C

Gauss program for the construction of D-optimal designs for the two-variable binary logistic model with interaction on a rectangular design space

@ — D-optimal designs for model $u = \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$ — @
@ — on the rectangular design space $[a, b] \times [c, d]$ — @
output file=optdes.out reset;
library optmum,pgraph;
include optmum.ext;
graphset;
format 15,6;

@—Rectangular design space—@
$a = 0$; $b = 2$; $c = 0$; $d = 2$;
“ ”; “Minimum and maximum x_1 values” $a \sim b$;
“ ”; “Minimum and maximum x_2 values” $c \sim d$;

@—Optimality criterion—@
“ ”; “Minimizing -ln(det)”;
mult=1;

@ — Best guess of parameters and starting design — @
/*LePelly and Sullivan data*/
b0=-2.2054;

```
b1=2.3766;
b2=2.2547;
b12=0.2846;
npts=4;
dmat={0.3 0,1.5 0,0 0.3,0 1.5,0.2 0.2,0.9 0.9};
wvec=ones(npts,1)/npts;
dwmat=dmat~wvec;
" ";"Best guess of the parameters b0 b1 b2 b12" b0 ~b1~b2~b12;
"Starting design";
"Number of support points = " npts;
"Support points and weights = " dwmat;

@ — Grid over the design space [a, b] × [c, d] — @
ng=51;
dx1=(b-a)/(ng-1);
dx2=(d-c)/(ng-1);
gridx1=seqa(a,dx1,ng);
gridx2=seqa(c,dx2,ng);
grid=gridgen(ng);
ntop=16;

@ — Optimization process — @
y0=constr(dwmat);
{ y,f,g,h } = optmum(&fun,y0);
dwmat=unconstr(y);
y=constr(dwmat);
crit=fun(y);
infomatopt=infomat(dwmat);
invmat=inv(infomatopt);

@ — results — @
" ";"optimal design";
"Number of support points = " npts;
"Support points and weights = " dwmat;
" ";"Criterion value " crit;
" ";"Optimum information matrix " infomatopt;
" ";"Inverse of the optimum information matrix " invmat;
```

@ — Standardized variance function $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ at the support points — @

stdvgrid=stdvarf(infomatopt,dwmat[.,1:2]);

expu=exp(b0+b1*dwmat[.,1]+b2*dwmat[.,2]+b12*dwmat[.,1].*dwmat[.,2]);

pvec=expu./(1+expu);

lvec=ln(pvec./(1-pvec));

" ";"Optimal design, p and logit(p) at support points" stdvgrid ~ pvec ~ lvec;

@ — Plot of $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ over a grid — @

stdvgrid=stdvarf(infomatopt,grid);

matcont=reshape(stdvgrid[.,3],ng,ng);

_pdate = 0;

fonts("simplex complex microb simgrma");

ztics(0,4,1,0.5);

xtics(0,2,0.5,0.1);

ytics(0,2,0.5,0.1);

zlabel("$d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$");

xlabel("x]1[");

ylabel("x]2[");

_pnum = 2;

_pnumht = 0.25;

_paxht = 0.2;

_pmcolor = $\{0, 0, 0, 0, 0, 0, 0, 0, 15\}$;

_pmsgctl = {0.8 0.7 0.3 0 1 0 0};

/*_pmsgstr="(b)";

/ surface(gridx1',gridx2,matcont');

stdvgrid=rev(sortc(stdvgrid,3));

ddmax=stdvgrid[1:ntop,.];

" ";"maximum $d(\boldsymbol{x}, \xi^*; \boldsymbol{\beta})$ values over the rectangle $[a, b] \times [c, d]$ " ddmax;

@ — Procedure - Function for minimization – @

proc fun(y);

local dwmat,imat,crit;

dwmat=unconstr(y);

imat=infomat(dwmat);

crit=-mult*ln(det(mmat));

retp(crit);

```
endp;
@ — Procedure to generate grid on the rectangle [a, b] × [c, d] — @
proc gridgen(ng);
local grid,k,i,j,toler;
toler=0.0000000001;
grid=zeros(ng*ng,2);
k=1;
i=1;
do until i > ng;
j=1;
do until j > ng;
if(abs(gridx1[i]) < toler);
gridx1[i]=0;
endif;
if(abs(gridx2[j]) < toler);
gridx2[j]=0;
endif;
grid[k,.]=gridx1[i] ~ gridx2[j];
k=k+1;
j=j+1;
endo;
i=i+1;
endo;
retp(grid);
endp;
@ — Procedure - Information matrix — @
proc infomat(dwmat);
local n,dmat,wvec,x1,x2,x3,expu,vterm,xmat,wmat,imat;
n=rows(dwmat);
dmat=dwmat[.,1:2];
wvec=dwmat[.,3];
x1=dmat[.,1];
x2=dmat[.,2];
x3=x1 .*x2;
```

```
expu=exp(b0+b1*x1+b2*x2+b12*x3);
vterm=sqrt(expeta)./(1+expeta);
xmat=vterm.*(ones(n,1)~ x1 ~ x2 ~ x3);
wmat=diagrv(zeros(n,n),wvec);
imat=xmat'.*wmat.*xmat;
retp(imat);
endp;

@ — Procedure - Unconstrained x and λ to constrained y— @
proc constr(dwmat);
local y,dmat,i,wvec,den;
y=zeros(3*npts-1,1);
/* transform support points */
dmat=dwmat[.,1:2];
i=1;
do until i>npts;
y[i,1]=arcsin(sqrt((dmat[i,1]-a)/(b-a)));
y[i+npts,1]=arcsin(sqrt((dmat[i,2]-c)/(d-c)));
i=i+1;
endo;
/* transform weights */ wvec=dwmat[.,3];
den=1.0;
i=1;
do until i>npts-1;
y[i+2*npts,1]=arcsin(sqrt(wvec[i]/den));
den=den-wvec[i];
i=i+1;
endo;
retp(y);
endp;

@ — procedure - constrained y to unconstrained x and λ — @
proc unconstr(y);
local dmat,i,wvec,yw,mult,dwmat;
/* transform support points */
dmat=zeros(npts,2);
```

```
i=1;
do until i>npts;
dmat[i,1]=x1min+(x1max-x1min)*sin(y[i,1])*sin(y[i,1]);
dmat[i,2]=x2min+(x2max-x2min)*sin(y[i+npts,1])*sin(y[i+npts,1]);
i=i+1;
endo;
/* transform weights */
wvec=zeros(npts,1);
yw=y[2*npts+1:3*npts-1,.];
mult=1;
i=1;
do until i>npts-1;
wvec[i,1]=sin(yw[i,1])*sin(yw[i,1])*mult;
mult=mult*cos(yw[i,1])*cos(yw[i,1]);
i=i+1;
endo;
wvec[npts,1]=mult;
dwmat=dmat~wvec;
retp(dwmat);
endp;

@ — Procedure to calculate the standardized variance function — @
proc stdvarf(mopt,grid);
local ngrid,stdvgrid,i,matx;
ngrid=rows(grid);
stdvgrid=grid zeros(ngrid,1);
i=1;
do until i > ngrid;
matx=infomat(grid[i,.]~1);
ddgrid[i,3]=sumc(diag(inv(mopt)*matx));
i=i+1;
endo;
retp(stdvgrid);
endp;
```