OPTIMAL EXPERIMENTAL DESIGNS FOR TWO-VARIABLE LOGISTIC REGRESSION MODELS

by

Yan Jia

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY
in
Statistics

APPROVED:

Raymond H. Myers, Chairman

Jeffrey B. Birch | Marvin Lentner

Marion R. Reynolds | Robert S. Schulman

July, 1996
Blacksburg, Virginia

Key Words: Design, Regression, Logistic

OPTIMAL EXPERIMENTAL DESIGNS FOR TWO-VARIABLE LOGISTIC REGRESSION MODELS

by

Yan Jia

Raymond H. Myers, Chairman

Statistics

(ABSTRACT)

Binary response data is often modeled using the logistic regression model. Experimental design theory for the logistic model appears to be increasingly important as experimentation becomes more complex and expensive. The optimal design work is extremely valuable in areas such as biomedical and environmental applications.

Most design research dealing with the logistic model has been concentrated on the one-variable case. Relative little has been done for the two-variable model. The primary goal of this research is to develop and study efficient and practical experimental design procedures for fitting the logistic model with two independent variables. Optimal designs are developed addressing D optimality, Q optimality, and the estimation of interaction between the design variables. The two-variable models with and without interaction usually have to be handled separately. The equivalence theory concerning D optimal designs is studied. The designs are compared using their relative efficiencies in the presence of interaction. Robustness to parameter misspecification is investigated. Bayesian design procedures are explored to provide relatively more robust experimental plans.

ACKNOWLEDGMENTS

I would like to extend my deepest appreciation to my advisor, Dr. Raymond Myers. His invaluable guidance and extraordinary insight has made my studies both challenging and rewarding. With his encouragement and friendship, the course of my research has been a truly enjoyable experience.

I wish to thank the members of my advisory committee: Jeffrey Birch, Marvin Lentner, Marion Reynolds, and Robert Schulman for their full support and interest in this research. Sincere gratitude is due to Michele Marini for her diverse computing advice. Special appreciation goes to Dr. Klaus Hinkelmann for his assistance. In general, I am grateful to the Department of Statistics at Virginia Tech, who has provided a strong academic environment and has made this achievement possible.

My parents and my husband, Alan, have been a great part of this work. Thanks to my family whose love and inspiration have kept me going with conscious endeavor.

Table of Contents

List of Figures

List of Tables

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

§1.1 Introduction

The primary focus of this work will be the experimental design aspect of response surface methodology (RSM), more specifically optimal-design theory for modeling the mean of binary response data using two-variable logistic regression models. As experimental situations become more complex and thus more expensive, researchers seek ways to maximize the information gained from a finite-size experiment. Optimal design theory has been thoroughly researched in the context of the linear model due to its simplicity and applicability. As the need for nonlinear statistical modeling increases in areas such as the chemical, biological, and clinical sciences, design research in the area of nonlinear models has gained momentum. In particular, optimal designs for fitting the logistic regression model with binary data have been the subject of numerous papers over the past two decades. However, the majority of those works consider only one design variable. Modeling using two design variables is more complicated and less structured in the context of design theory. Nevertheless, the usefulness and practicality of the two-variable logistic models make the development of its design theory an important subject. This research seeks to extend the optimal-design theory from the one-variable framework to the situation of fitting a logistic regression model containing two design variables and their possible interaction.

§1.2 Linear Model and Design Optimality

The theory of design optimality for the linear model has been well documented in the statistical literature. The standard linear model has the following form:

$$y = X\beta + \varepsilon$$

where

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{nk} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}.$$

In this model, y is a vector of responses from n experimental runs, X is a matrix containing k+1 regressor variables across the n runs and is determined by the design variables and the model form, β is a vector of unknown parameters, and ε is a vector of random errors. A general assumption is that $E[\varepsilon] = 0$ and $Var[\varepsilon] = \sigma^2 I$ where σ^2 is a common variance, usually unknown. Under this assumption, the best linear unbiased estimator (BLUE) for β is the ordinary least squares (OLS) estimator of β given by $\hat{\beta} = (X'X)^{-1}X'y$ with $E[\hat{\beta}] = \beta$ and $Var[\hat{\beta}] = \sigma^2(X'X)^{-1}$. Normality is often assumed for ε, i.e. $\varepsilon \sim N(0, \sigma^2 I)$. This will result in the convenient property that the maximum likelihood estimator (MLE) of β is identical to the OLS estimator $\hat{\beta}$, which is distributed as $N(\beta, \sigma^2(X'X)^{-1})$. The normality assumption also leads to the fact that the Fisher information matrix of β, $I(\beta)$, is equal to the inverse of the variance-covariance matrix of $\hat{\beta}$, or $I(\beta) = \frac{1}{\sigma^2}(X'X)$. In fact, the matrix $(X'X)^{-1}$ plays a key role in the design optimality criteria as will be seen shortly.

Kiefer and Wolfowitz (1959) introduced the foundation of optimal design criteria in the early development. Representing the design by a probability measure assigning the sample allocation over the design space, they established the D and E optimality criteria, from which evolved the term "alphabetic optimality". Additional performance criteria were created subsequently including A, G, and Q. These design criteria address either parameter estimation or prediction of the mean responses.

The most well-known and commonly used design criterion is D optimality. The D optimal design minimizes the determinant of $(X'X)^{-1}$, the generalized variance of the coefficient estimates $\hat{\beta}$. It can be shown that minimization of $|(X'X)^{-1}|$ results in the smallest volume of the confidence ellipsoid for β. A scaling factor $\frac{N}{\sigma^2}$ is adopted in most criteria to remove the effects of the total resource consumed, represented by the sample size N, and the magnitude of the error variance, represented by σ^2. Thus minimizing the determinant of $\frac{N}{\sigma^2}\text{Var}[\hat{\beta}]$, or $|N(X'X)^{-1}|$, leads to the D optimality criteria expressed in the following form:

$$\underset{\mathcal{D}}{\text{Max}} \left|\frac{X'X}{N}\right|$$

where $\mathcal{D}$ is the set of possible designs.

A common criterion addressing the quality of prediction is Q optimality. The Q criterion aims at improving prediction stability in an average sense over a specific region of the design variables. Box and Draper (1959, 1963) first considered the idea of averaging or integrating prediction variance over a region of interest R. The prediction variance $\text{Var}[\hat{y}(x)]$ at the point x in the model space is given by $\text{Var}[x'\hat{\beta}]=\sigma^2 x'(X'X)^{-1}x$. Thus the Q optimality defined by minimizing the average scaled prediction variance (APV) is expressed as

$$\underset{\mathcal{D}}{\text{Min}} \frac{N}{K} \int_R x'(X'X)^{-1}x \, dx$$

where K is the volume of the region and is given by $\int_R dx$.

Another criterion motivated by enhancing prediction is G optimality. This criterion addresses the same target of prediction variance as Q optimality does. However the approach is minimax rather than averaging. Seeking the best among the worst, G optimality minimizes the maximum scaled prediction variance over a region R according to the following expression:

$$\underset{\mathcal{D}}{\text{Min}}\ \underset{x \in R}{\text{Max}}\ N\, x'(X'X)^{-1}x\ .$$

Myers, Khuri, and Carter (1989) provided more detailed discussion on the design criteria and their use in RSM. Since these criteria are functions of only the design-point locations, an optimal design in the linear case can be expressed explicitly as a placement of design points and thus can be implemented independently of the model parameters. It is apparent that the above criteria are all driven entirely or partially by the matrix $X'X$. In the linear model situation, one can recognize the matrix $(X'X)^{-1}$ directly as an active part of $Var[\hat{\beta}]$ without referring to the information matrix. However for a nonlinear model, the notion of an information matrix becomes a necessity since one has to rely on asymptotic properties of the MLE for β which are in fact governed by the information matrix $I(\beta)$. The information matrix $I(\beta)$ in the nonlinear case usually involves unknown parameters in addition to the design layout. This is one reason that design optimization becomes more complicated and parameter dependent for nonlinear models. The criteria mentioned in this section were originally developed for the linear model. As the application of statistics broadens, much work has been done to adapt them to nonlinear situations. The nature of a nonlinear model though can significantly complicate the way in which one implements the criteria.

§1.3 Nonlinear and Logistic Models

A nonlinear model usually takes the form

$$y = f(x; \beta) + \varepsilon \tag{1.3.1}$$

where x is a vector of regressor variables, y is the response at the point x, β is a vector of unknown parameters, ε is the random error, and f is a known function nonlinear in β. A usual assumption is that $E[y] = f(x; \beta)$ when modeling the mean response.

The distribution of the random response y in nonlinear modeling often comes from one of the exponential families such as Poisson, binomial, exponential, or gamma, etc. The general form of the distribution functions for the exponential families is given by

$$p(y; \theta) = \exp\{ r(\phi)[y\theta - g(\theta)] + h(y, \phi) \} \tag{1.3.2}$$

where y is the random variable, θ is the unknown parameter of interest, and the function $r(\phi)$ is called the scale parameter. Based on (1.3.2), a subclass of nonlinear models called the generalized linear models (GLM) is developed by using nonlinear functions of a linear predictor $x'\beta$.

The binary response data under study in this research have Bernoulli distributions. The mean response is modeled by the so called **logistic function** written as

$$f(x; \beta) = E[y] = P = \frac{1}{1+e^{-x'\beta}}, \tag{1.3.3}$$

which corresponds to the **logit** link function

$$\text{logit}(P) = \log(\frac{P}{1-P}) = x'\beta \tag{1.3.4}$$

according to the theory of generalized linear models. For an experiment with a total of N observations over m design points with sample size n_i at the i^{th} point, i = 1, ... , m, the **logistic regression model** is given by

$$y_{ij} = f(x_i ; \beta) + \varepsilon_{ij} = P_i + \varepsilon_{ij} = \frac{1}{1+e^{-x_i'\beta}} + \varepsilon_{ij} \tag{1.3.5}$$

where y_{ij} and ε_{ij} are respectively the binary response and error for the j^{th} observation at the i^{th} design point, and $x_i'\beta$ is the linear predictor in p parameters at the i^{th} point. The information matrix for β is given by

$$I(\beta) = X'\Sigma X \tag{1.3.6}$$

where X is an m×p matrix with the i^{th} row x_i' corresponding to the i^{th} design point, and Σ is an m×m diagonal matrix with the i^{th} diagonal element being the binomial variance $n_iP_i(1-P_i)$ at the i^{th} design point, i = 1, ... , m.

The logistic model in (1.3.5) is not only a result of an elegant mathematical derivation in the framework of GLM, but it is also a well suited nonlinear function which often closely describes binary data in practice, more specifically in areas such as biological assays of drugs, material fatigue to failure studies, and stress and fracture studies. In a drug testing experiment for instance, the response y_{ij} in (1.3.5) can be a cure or lack thereof (1 or 0), and the mean response P_i is the probability of a cure (1) at the dose level(s) determined by x_i. Due to its nondecreasing and nonnegative nature along with being bounded by 0 and 1, the logistic function in (1.3.5) can appropriately serve as the cumulative distribution function of the linear predictor $x'\beta$. This gives rise to the concepts of tolerance and tolerance distribution. Each subject in a bioassay is assumed to have a given tolerance level which is the maximum amount of drug that the subject can sustain before producing a response. With this tolerance corresponding to the linear predictor $x'\beta$, the logistic function in (1.3.5) can be used as the cumulative distribution of the tolerance over the population of subjects under study. Hence Prob(tolerance$\leq x'\beta$) = $\frac{1}{1+e^{-x'\beta}}$. Since the population proportion with tolerance lower than or equal to a particular level $x'\beta$ is nothing but the proportion of responses in the population, or the probability of response for a single subject, when a dose at level x is administered, the two ways of interpreting the probability in (1.3.5) logically fit each other. In other words, Prob(y=1|x) = P(tolerance$\leq x'\beta$) = $\frac{1}{1+e^{-x'\beta}}$. This probability as a cumulative distribution further leads to the density function of the tolerance distribution

$$f(x'\beta) = \frac{e^{-x'\beta}}{(1+e^{-x'\beta})^2} \ . \tag{1.3.7}$$

§1.4 Design Optimality for One-Variable Logistic Model

Research in design optimality for the logistic model began with the simplest situation in which a single design variable is considered. The logistic probability function in (1.3.5) for the one-variable model is written as

$$f(x;\, \beta) = E[y] = P = \frac{1}{1+e^{-(\beta_0+\beta_1 x)}}$$

where x is a scalar. In this work x will be a vector unless otherwise indicated. An example of one-variable logistic curve with $\beta_0 = -3$ and $\beta_1 = 2$ is shown in figure 1.4.1.

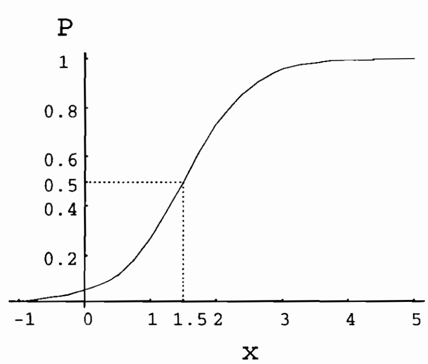


Figure 1.4.1 One-variable logistic curve, $\beta_0 = -3$, $\beta_1 = 2$

Again, design considerations center around parameter estimation and prediction of the response, involving the information matrix $X'\Sigma X$ as a key factor. Naturally, design work attempts to apply the ideas of the alphabetic criteria (D, A, Q, G, E) developed for the linear model to the logistic situation. In addition, the F criterion which is unique to the one-variable logistic case has been created. Unlike the linear case, the scaling factor applied to variances in the criteria is N rather than $\frac{N}{\sigma^2}$ because the common variance σ^2 no longer exists in the logistic case and the variance is in fact a function of the mean. Some important works are summarized in the following.

Kalish and Rosenberger (1978) derived the D optimal designs. Since the asymptotic variance-covariance matrix of the MLE for β is the inverse of the information matrix $X'\Sigma X$, the D optimality concerning $Var(\hat{\beta})$ becomes

$$\underset{\mathcal{D}}{Max} \left| \frac{I(\beta)}{N} \right| .$$

They also developed the G optimal designs addressing $Var[\hat{P}]$, the variance of predicted response. It can be shown that $Var[\hat{P}]$ asymptotically approaches $P^2(1-P)^2x'I(\beta)^{-1}x$. Thus the G optimality is given by

$$\underset{\mathcal{D}}{Min} \underset{x \in R}{Max} \; N \, P^2(1-P)^2 \, x' \, I(\beta)^{-1}x.$$

Myers, Myers, and Carter (1994) derived the Q and G optimal designs in which the variance of predicted logit is addressed rather than the variance of predicted response. Since the logit is defined by $logit(P) = log(\frac{P}{1-P}) = x'\beta$, its prediction is $logit(\hat{P}) = x'\hat{\beta}$ with the asymptotic variance $Var[logit(\hat{P})] = x'I(\beta)^{-1}x$. Based on logit, the Q criterion is

$$\underset{\mathcal{D}}{Min} \; \frac{N}{K} \int_R x' \, I(\beta)^{-1} x \, dx ,$$

and similarly the G criterion is

$$\underset{\mathcal{D}}{Min} \underset{x \in R}{Max} \; N \, x' \, I(\beta)^{-1} x .$$

Myers (1991) and Letsinger (1995) investigated optimal designs using the E optimality which is given by

$$\underset{\mathcal{D}}{Min} \underset{i}{Max} \; \lambda_i$$

where λ_i is an eigenvalue of $I(\beta)^{-1}$, i = 1, 2.

The F optimality is a criterion of different interest. It addresses the estimation of the dose level required to produce a certain probability of response, P. This dose is usually referred to as ED_{100P} where ED stands for "effective dose". An F optimal design minimizes the length of the Fieller interval on ED_{100P} for a specific P. Several F optimal as well as D and A optimal designs were given for the logistic model as well as other nonlinear models by Sitter and Wu (1994). Letsinger (1995) obtained various F optimal designs and studied their relationship to other criteria.

As previously noted, since these criteria are functions of both the design layout and the unknown parameters, the optimal designs obtained can only be expressed to the best extent as certain response probabilities, which implicitly point to the actual design levels via the parameter values. For instance, the D optimal design for the one-variable logistic model is a two point design with equal numbers of design points at $ED_{17.6}$ and $ED_{82.4}$. The actual doses corresponding to the probabilities of response being 0.176 and 0.824 remain unknown unless true values of the parameters are available. In this sense, the optimal designs based on the above criteria depend on the values of the parameters. In practice, parameters must be "guessed" prior to implementing a design. This brings up the issue of design robustness to parameter misspecification.

Various optimal designs were further evaluated upon their robustness to parameter misspecification through efficiency studies. These results accompany much of the literature dealing with optimal design procedures. It is felt that designs with multiple levels tend to be more robust than the two-level designs, especially in the presence of severe parameter misspecification. Related discussions are found with Abdelbasit and Plackett (1983) and Myers (1991).

Moreover, several other statistical techniques including sequential experimentation, Bayesian and minimax approaches have been incorporated into design optimization procedures in order to produce more robust designs. Bayesian analogs to D and A optimal designs were given by Chaloner and Larntz (1989). Their procedure takes into account the parameter uncertainty by utilizing prior distributions of parameters specified by the experimenter. The minimax procedure, also based on prior knowledge about the parameters, has been studied by Sitter (1992), whose work dealt with the D, F, and asymptotic confidence interval criteria. Letsinger (1995) explored the Bayesian F and D criteria in greater detail. Sequential designs allow experiments to proceed in stages, where a later stage could benefit from the information gained earlier. Two-stage designs were discussed by Abdelbasit and Plackett (1983) and Minkin (1987) using the D optimality criterion. Myers (1991) developed a two-stage D-Q design with the first stage being D optimal and the second stage Q optimal conditioned on the first stage. Letsinger (1995) developed a two-stage D-D design which involves a Bayesian procedure in the first stage.

§1.5 Design Optimality for Two-Variable Logistic Models

Due to the complexity of the design issue, the vast majority of design optimality research for logistic models thus far has been dealing with the simplest situation with the one-variable model. However, as experimental studies of binary data become highly sophisticated, experimentation with two or more design factors receives increasing use in practice. In clinical sciences for instance, development of a combination drug by studying the effects of the two component compounds and their possible interaction is a frequent topic in the current literature of biopharmaceutical science. The prevailing research interest in the joint effect of two variables on binary responses motivates investigation of optimal-design theory for the two-variable logistic models. This research studies optimal designs for the two-variable logistic models in two basic forms: the simple first-order model and the model with both the first-order and interaction terms.

For the simple first-order model, the logistic probability function in (1.3.5) is written as

$$f(x; \beta) = E[y] = P = \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\beta_2 x_2)}} \qquad (1.5.1)$$

The probability of response P remains constant when the design point in the (x_1, x_2) space stays on a straight line defined by

$$\beta_0 + \beta_1 x_1 + \beta_2 x_2 = L \qquad (1.5.2)$$

where L is the logit of the constant probability P, given by $L = \text{logit}(P) = \log(\frac{P}{1-P})$. The line expressed in (1.5.2) can be viewed as ED_{100P}. Graphical features are illustrated with an example where $\beta_0 = -1$, $\beta_1 = 1$, and $\beta_2 = 2$. Figure 1.5.1 presents the parallel-line contours for the logit and the response probability. Figure 1.5.2 shows the surfaces of the logit and the response probability respectively over the (x_1, x_2) space. The logit forms a plane. Applying the logistic function to the logit bends the plane into a curved surface for the response probability, where the curvature parallels that of a typical logistic curve in the one-variable case.

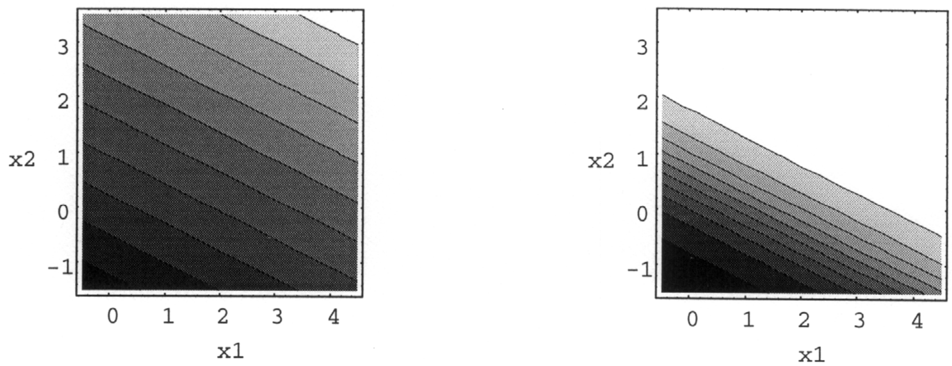
For the model with the additional interaction term, the logistic probability function in (1.3.5) is written as

$$f(x; \beta) = E[y] = P = \frac{1}{1+\exp^{-(\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_{12} x_1 x_2)}} . \qquad (1.5.3)$$

A constant probability P falls on a pair of hyperbolae in the design space represented by

$$\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 = L \qquad (1.5.4)$$

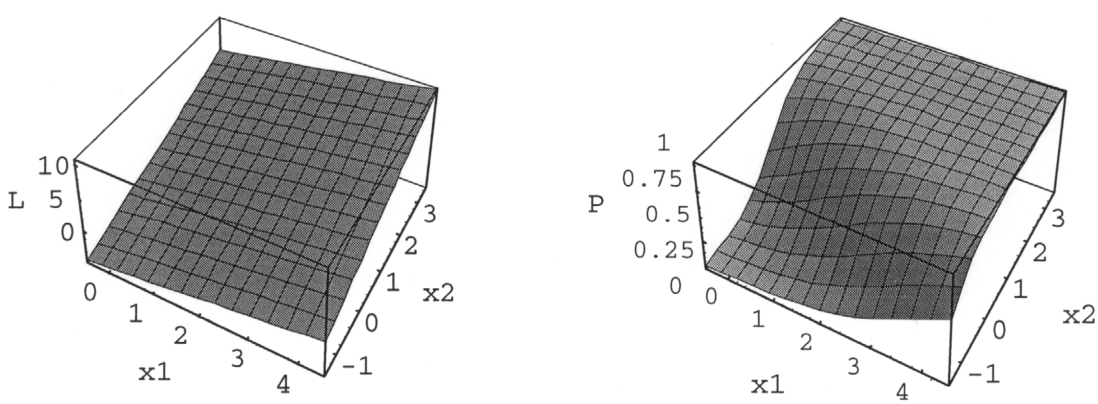
where L is the logit given by $L = \text{logit}(P) = \log(\frac{P}{1-P})$. The pair of hyperbolae expressed in (1.5.4) can be viewed as ED_{100P}. Graphical features are illustrated in figures 1.5.3 and 1.5.4 for an example with $\beta_0 = -1$, $\beta_1 = 1$, $\beta_2 = 2$, and $\beta_{12} = -1$. Note that except β_{12}, the other parameters are the same as in the example for the no-interaction model. Figure 1.5.3



(a) Logit (b) Probability

Figure 1.5.1 Contours of logit and probability for two-variable logistic model,

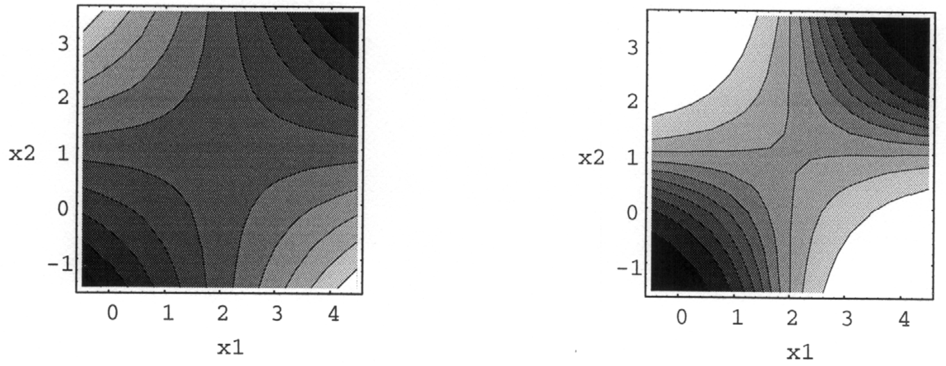
$$\beta_0 = -1,\ \beta_1 = 1,\ \beta_2 = 2,\ \beta_{12} = 0$$



(a) Logit (b) Probability

Figure 1.5.2 Surfaces of logit and probability for two-variable logistic model,

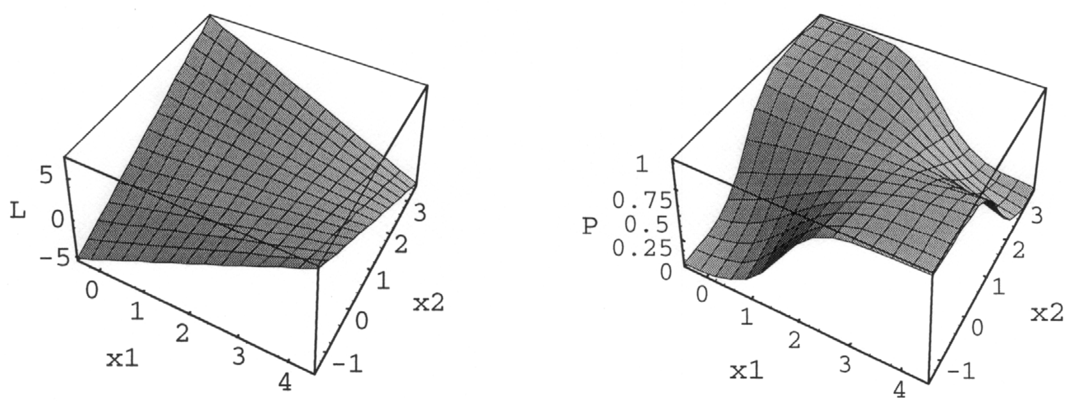
$$\beta_0 = -1,\ \beta_1 = 1,\ \beta_2 = 2,\ \beta_{12} = 0$$



(a) Logit (b) Probability

Figure 1.5.3 Contours of logit and probability for two-variable logistic model,

$\beta_0 = -1, \beta_1 = 1, \beta_2 = 2, \beta_{12} = -1$



(a) Logit (b) Probability

Figure 1.5.4 Surfaces of logit and probability for two-variable logistic model,

$\beta_0 = -1, \beta_1 = 1, \beta_2 = 2, \beta_{12} = -1$

shows the hyperbolic contours of the logit and those of the probability. Figure 1.5.4 plots the surfaces of the logit and probability respectively over the (x_1, x_2) space. Both the logit and the response probability appear in curved surfaces.

Obviously, the presence of interaction considerably changes the behaviors of the logit and the response probability over the design space. This could substantially complicate the design optimality studies for the model containing interaction. Completely different design geometries may have to be employed in handling the situations with and without the interaction.

Relatively little work has been attempted in seeking optimal designs for two-variable logistic models. To study two drugs used in combination, Mantel (1958) and others first advocated an idea which later was termed a "ray" design by Brunden, Vidmar, and McKean (1988). The design received its name from the fact that the design points fall on rays emanating from the origin onto the first quadrant. Using the D optimality criterion, Brunden *et al.* proposed the two-ray and three-ray designs as well as some 2×2 and 3×3 factorial designs. A 2×2 factorial design was also created through minimization of Var($\hat{\beta}_{12}$).

Theoretically, the design criteria introduced in the previous section for the one-variable logistic model also apply to the two-variable models discussed here. However, additional considerations and necessary constraints must be given as the design space expands from one to two dimensions. In the two-dimensional space, a design with an arbitrary geometric pattern may very likely result in optimal ED's that are dependent on the parameter values. The ray designs by Brunden *et al.* are of this type. The current research pays special attention to the natural ED pattern, which appears as parallel lines for the no-interaction model and pairs of hyperbolae for the interaction model. Allowing the natural ED pattern to dominate the design geometry, optimal designs are obtained in this work where the optimal ED's are independent of the parameter values.

Several D optimal designs are developed for the no-interaction and interaction models. The equivalence property is studied for certain D optimal designs. The D efficiencies of various designs, including some by Brunden *et al.*, are evaluated in the presence of interaction.

Designs for optimizing the estimation of the interaction coefficient are given and their performances compared using the relative efficiencies. Q optimal designs are created for both the no-interaction and interaction models. The relative efficiencies indicate that the logit-based Q optimal designs and the probability-based ones achieve quite similar performances.

As design implementation remains parameter dependent with the two-variable logistic model, the robustness property to parameter misspecification is investigated for various D optimal designs. The Bayesian design procedure is explored with D optimality in an effort to better cope with poor parameter knowledge. Robustness of the Bayesian designs is also studied and they appear to have overall improvement over the conventional designs. In general, the Bayesian design approach is found to be successful in building more robust experimental plans for the two-variable logistic model.

CHAPTER 2

D OPTIMAL DESIGNS

The D optimality criterion focuses on achieving best parameter estimation by addressing the variance-covariance structure of the coefficient estimates. The asymptotic variance-covariance matrix of the MLE for β is the inverse of the Fisher information matrix $I(\beta)$, which is given by $X'\Sigma X$ in the case of a logistic model. A D optimal design maximizes $|X'\Sigma X|$. Unlike the one-variable logistic model for which an optimal design is expressed purely through ED's, designs for the two-variable models come in different geometric patterns in addition to the optimal ED's.

Based on the ED patterns, a parallel-line design is given for the no-interaction model and a hyperbola-based design for the interaction model. The factorial and ray designs initially found by Brunden *et al.* (1988) are rederived to illustrate additional insight. The D criterion used by these authors is unusual in that the information matrix $I(\beta)$ comes from fitting an interaction model while the dose-response relationship involved in $I(\beta)$ is assumed to be the no-interaction model. Obviously, they tried to involve the interaction into the optimization while still making use of the simple straight line ED's. To experience such an approach, an alternative parallel-line design is generated in a similar fashion. Still, the effectiveness of Brunden's criterion will be investigated through efficiency studies. On the other hand, modified ray designs are also created using the regular D criterion for the interaction model as opposed to Brunden's approach, i.e. assuming that the information matrix from fitting the interaction model does contain the interaction in the true dose-response relationship.

§2.1 Parallel-Line Design

The no-interaction two-variable logistic model takes the form

$$y = P + \varepsilon = \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\beta_2 x_2)}} + \varepsilon . \qquad (2.1.1)$$

The ED's appear as parallel lines in the two-dimensional design space. In particular, ED_{100P} is a line defined by $\beta_0 + \beta_1 x_1 + \beta_2 x_2 = L$, where $L = \text{logit}(P) = \log(P/(1-P))$. The D optimal design in the one-variable case is given by two levels: $ED_{17.6}$ and $ED_{82.4}$. In the two-variable case, the two ED points are now relaxed into two parallel lines as the design space extends from one to two dimensions. An analogy to the one-variable design in the two-variable situation should naturally lie on two parallel ED lines.

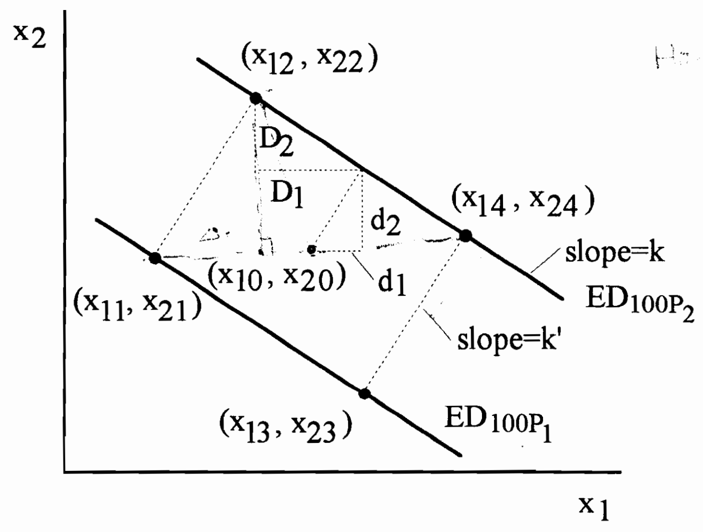


Figure 2.1.1 Parallel-line design

Consider a preliminary design arrangement as shown in figure 2.1.1. The four-point design falls on two parallel lines denoted ED_{100P_1} and ED_{100P_2}. The distance between the two points on the same ED is equal to that on the other ED. The slopes k' and k, as indicated in the figure, are not restricted to be perpendicular. The center of the design is denoted by (x_{10}, x_{20}) where $x_{10} = \frac{x_{11}+x_{14}}{2} = \frac{x_{12}+x_{13}}{2}$, and $x_{20} = \frac{x_{21}+x_{24}}{2}$ $= \frac{x_{22}+x_{23}}{2}$. The sample sizes are assumed equal for the points on the same ED but not necessarily so for points on different ED's. Let r and 1-r denote the sample proportions allocated to ED_{100P_1} and ED_{100P_2} respectively.

Similar to the one-variable case, the probabilities P_1 and P_2 as well as the sample proportion r in the parallel-line design are to be optimized through the D optimality criterion, which is given by

$$|I(\beta)| = N^3 \left(\frac{D_1^2}{\beta_2^2}\right) \left[\log\left(\frac{P_2}{1-P_2}\right) - \log\left(\frac{P_1}{1-P_1}\right)\right]^2$$

$$\cdot\ rP_1(1-P_1)\ (1-r)P_2(1-P_2)\cdot[rP_1(1-P_1) + (1-r)P_2(1-P_2)]\ . \qquad (2.1.2)$$

Maximizing $|I(\beta)|$ leads to the optimal solutions

$$P_1 = 0.227425, \qquad P_2 = 0.772575, \qquad \text{and} \qquad r = 0.5\ , \qquad (2.1.3)$$

which give an equal sample size design with symmetric ED's. The optimal ED's ($ED_{22.7}$, $ED_{77.3}$) draw closer together than those of the one-variable design ($ED_{17.6}$, $ED_{82.4}$). The D criterion for a one-variable two-level design is given by

$$|I(\beta)| = \frac{N^2}{\beta_1^2} \left[\log\left(\frac{P_2}{1-P_2}\right) - \log\left(\frac{P_1}{1-P_1}\right)\right]^2 rP_1(1-P_1)(1-r)P_2(1-P_2)\ . \qquad (2.1.4)$$

The two-variable criterion differs from the one-variable one with the additional factor $[rP_1(1-P_1) + (1-r)P_2(1-P_2)]$, which provides an extra force driving P_1 and P_2 both toward ED_{50}. This causes the smaller dispersion in the two-variable design.

The parallel-line design allows user customization in several ways as illustrated in figure 2.1.2. First, one can choose his or her own design location since the D criterion remains constant as the design moves along the two ED lines. Secondly, the two non-ED edges can be rotated to a desired angle without changing the criterion value. Thirdly, one can always improve the design by making it wider in the direction of the ED lines as this will yield a greater criterion value.

(a) (b)

Figure 2.1.2 Properties of parallel-line design
(a) Invariance of D criterion to location and angle
(b) Increase in D criterion as the design widens

The parallel-line design is a four-point design assuming a main effect model and hence offers one lack-of-fit degree of freedom for testing interaction. Adding the design center as another design point would further provide one extra degree of freedom designated to the detection of quadratic trend. Similar to designs with center runs for the linear model, the single degree of freedom due to the center point can only be used to test the sum of the quadratic coefficients as the individual quadratic effects in x_1 and x_2 are confounded. As some of the observations are assigned as center runs, the D criterion of the design is expected to become less than optimal. The D efficiency of design A relative to design B can be evaluated with

$$\text{D-EFF} = \left(\frac{|I(\beta)| \text{ from design A}}{|I(\beta)| \text{ from design B}} \right)^{\frac{1}{p}}, \qquad (2.1.5)$$

where p is the number of model parameters.

It is assumed that the center point is given the same number of runs as any other point of the parallel-line design. When fitting a no-interaction model, the D efficiency of a parallel-line design with center runs relative to the design at the same symmetric ED's but without center runs is

$$\frac{4}{5}\left(1+\frac{1}{16P(1-P)}\right)^{\frac{1}{3}}, \tag{2.1.6}$$

where P and 1-P are the response probabilities that the design resides on. For the optimal design where P = 0.772575, the efficiency due to the center runs is 88.54%, which is not severely low. It is the balance between an acceptable D criterion and the flexibility in being able to check model adequacy that leads to one's decision as to whether the center runs should be used.

Design implementation with a nonlinear model normally requires parameter knowledge. As in the one-variable logistic case, the parameters must be known to construct a parallel-line design for the two-variable model. Design robustness to parameter misspecification will be dealt with in a later chapter. Assuming the parameter values are supplied, the design points of a rectangular parallel-line design can be found using the design matrix

$$\begin{bmatrix} x_{11} & x_{21} \\ x_{12} & x_{22} \\ x_{13} & x_{23} \\ x_{14} & x_{24} \end{bmatrix} = \begin{bmatrix} x_{10} - D_1 - \frac{\beta_1 L}{\beta_1^2+\beta_2^2} & -\frac{1}{\beta_2}[\beta_0 + \beta_1(x_{10} - D_1)] - \frac{\beta_2 L}{\beta_1^2+\beta_2^2} \\ x_{10} - D_1 + \frac{\beta_1 L}{\beta_1^2+\beta_2^2} & -\frac{1}{\beta_2}[\beta_0 + \beta_1(x_{10} - D_1)] + \frac{\beta_2 L}{\beta_1^2+\beta_2^2} \\ x_{10} + D_1 - \frac{\beta_1 L}{\beta_1^2+\beta_2^2} & -\frac{1}{\beta_2}[\beta_0 + \beta_1(x_{10} + D_1)] - \frac{\beta_2 L}{\beta_1^2+\beta_2^2} \\ x_{10} + D_1 + \frac{\beta_1 L}{\beta_1^2+\beta_2^2} & -\frac{1}{\beta_2}[\beta_0 + \beta_1(x_{10} + D_1)] + \frac{\beta_2 L}{\beta_1^2+\beta_2^2} \end{bmatrix} \tag{2.1.7}$$

where L = logit(P_2) = 1.22291. The points on the lower ED are (x_{11}, x_{21}) and (x_{13}, x_{23}) while those on the upper ED are (x_{12}, x_{22}) and (x_{14}, x_{24}). The experimenter must choose the location and size of the design through specifying the central level x_{10} and the distance D_1, which are indicated in figure 2.1.1. If a total deviation Δ_1 from the central level to the level at the outermost point is specified, then D_1 can be obtained by $D_1 = \Delta_1 - \frac{|\beta_1|L}{\beta_1^2 + \beta_2^2}$.

For instance, suppose that the parameters are given by $\beta_0 = -3$, $\beta_1 = 0.5$, and $\beta_2 = 1$. Suppose that the variable x_1 is operable within ±1 from its central level. This implies $\Delta_1 = 1$ and thus $D_1 = 1 - \frac{|\beta_1|L}{\beta_1^2 + \beta_2^2} = 0.49$. Applying the design matrix in (2.1.7) yields the design points (1.00, 1.28) and (2.02, 0.76) on $ED_{22.7}$ and (1.98, 3.24) and (3.00, 2.72) on $ED_{77.3}$. The equal sample size design appears in figure 2.1.3.

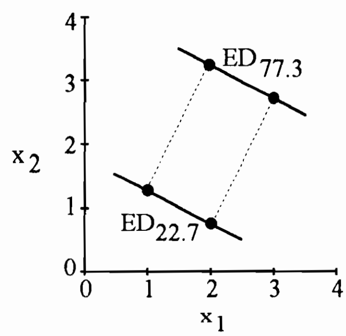


Figure 2.1.3 Example of parallel-line D optimal design

§2.2 Hyperbola-Based Design

The two-variable logistic model with interaction takes the form

$$y = P + \varepsilon = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2)}} + \varepsilon . \tag{2.2.1}$$

A constant probability P falls on a pair of hyperbolae in the design space defined by $\beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_{12}x_1x_2 = L$, where L = logit(P). A four-point design situated on two pairs of hyperbolae is considered and shown in figure 2.2.1. The points are placed in such a way that any two points on the same ED are symmetric about the hyperbola center. The points on a common ED are assumed to have equal sample sizes. The sample proportions for ED_{100P_1} and ED_{100P_2} are denoted r and 1-r respectively.

The hyperbolic ED curves center at (x_{10}, x_{20}) where $x_{10} = -\frac{\beta_2}{\beta_{12}}$ and $x_{20} = -\frac{\beta_1}{\beta_{12}}$. The centered variables can be obtained as $z_1 = x_1 - x_{10}$ and $z_2 = x_2 - x_{20}$. With the centered variables, the logit expression simplifies to $L = \beta_0^* + \beta_{12}z_1z_2$, where β_0^* represents the logit value anywhere on the centered axes and is given by

$$\beta_0^* = \beta_0 - \frac{\beta_1\beta_2}{\beta_{12}}. \qquad (2.2.2)$$

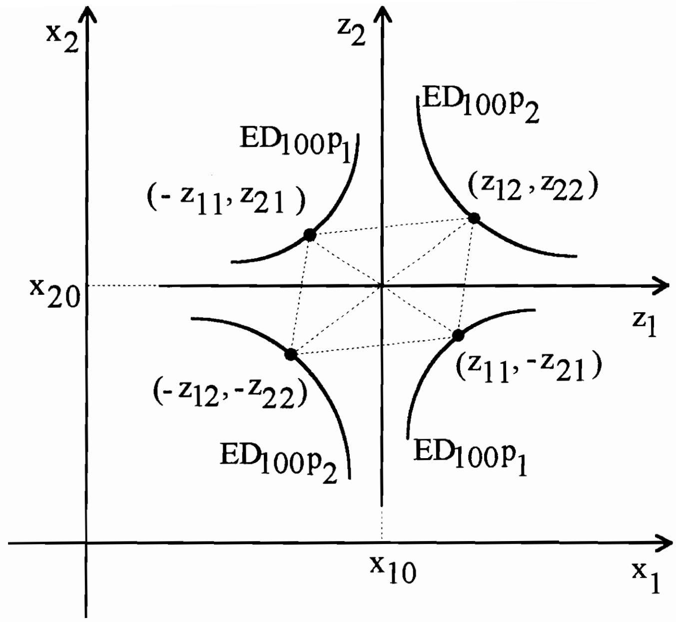
Different possible situations of a hyperbola-based design are also shown in figure 2.2.1 along with their respective conditions.

Given the hyperbola-based design structure, the probabilities P_1 and P_2 as well as the sample proportion r will be optimized using the D optimality criterion, which is shown to be

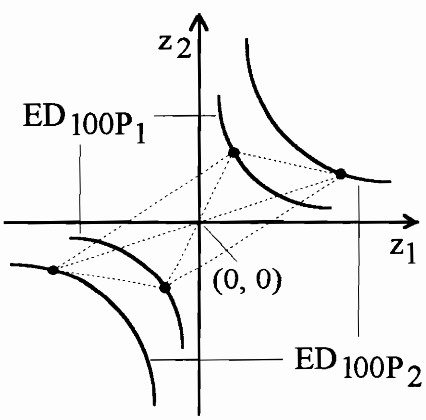
$$|I(\beta)| = \frac{N^4}{\beta_{12}^4} [rP_1(1-P_1)(1-r)P_2(1-P_2)]^2 [\log(\frac{P_2}{1-P_2}) - \log(\frac{P_1}{1-P_1})]^2$$

$$\cdot [t(\log(\frac{P_2}{1-P_2}) - \beta_0^*) - \frac{1}{t}(\log(\frac{P_1}{1-P_1}) - \beta_0^*)]^2, \qquad (2.2.3)$$

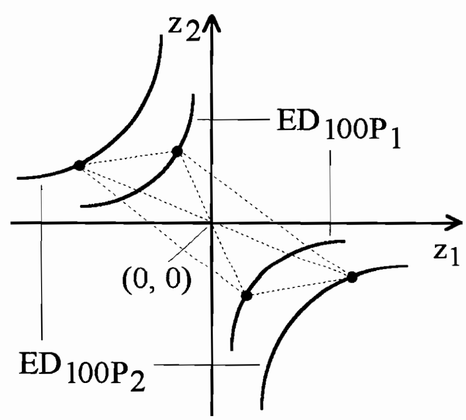
where t is a ratio of the centered design levels given by either $\left|\frac{z_{11}}{z_{12}}\right|$ or $\left|\frac{z_{21}}{z_{22}}\right|$. Since $t \in (0, \infty)$ and $|I(\beta)| \to \infty$ as $t \to 0_+$ or $t \to \infty$ for any values of P_1, P_2 and r, an optimal t that is



(a)



(b)



(c)

Figure 2.2.1 Hyperbola-based design in different patterns

(a) $\text{logit}(P_1) < \beta_0^* < \text{logit}(P_2)$ for $\beta_{12} > 0$, or $\text{logit}(P_2) < \beta_0^* < \text{logit}(P_1)$ for $\beta_{12} < 0$

(b) $\beta_0^* < \text{logit}(P_1) < \text{logit}(P_2)$ for $\beta_{12} > 0$, or $\text{logit}(P_2) < \text{logit}(P_1) < \beta_0^*$ for $\beta_{12} < 0$

(c) $\text{logit}(P_2) < \text{logit}(P_1) < \beta_0^*$ for $\beta_{12} > 0$, or $\beta_0^* < \text{logit}(P_1) < \text{logit}(P_2)$ for $\beta_{12} < 0$

bounded, or practical, does not exist. Therefore the optimal P_1, P_2 and r may have to depend on the parameter β_0^* and a user selected ratio t. However, it is felt that using t = 1 is an option which is both mathematically neat and practically convenient. With t = 1, β_0^* is canceled from (2.2.3) and thus the D criterion reduces to

$$|I(\beta)| = \frac{N^4}{\beta_{12}^4}[rP_1(1-P_1)\,(1-r)P_2(1-P_2)]^2\,[\log(\frac{P_2}{1-P_2}) - \log(\frac{P_1}{1-P_1})]^4 . \qquad (2.2.4)$$

The condition t = 1 corresponds to certain design symmetry as can be seen in figure 2.2.2. For both situations given in the figure, the symmetry refers to the fact that the points on different ED's are equidistant to one of the centered axes.

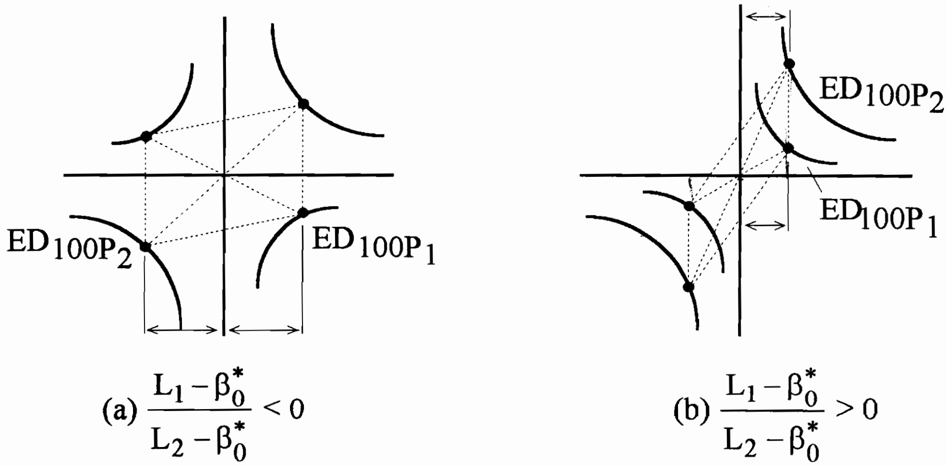


Figure 2.2.2 Symmetry of hyperbola-based design with t = 1

Maximizing $|I(\beta)|$ in (2.2.4) gives an optimal design with symmetric ED's and equal sample sizes, i.e.

$$P_1 = 0.176041, \qquad P_2 = 0.823959, \qquad \text{and} \qquad r = 0.5 . \qquad (2.2.5)$$

It is realized that these optimal ED's are in fact identical to the one-variable D optimal solutions. Examining the D criterion expressions for the one- and two-variable designs as given in (2.1.4) and (2.2.4) tells that apart from the constants, the two-variable expression is exactly the square of the one-variable expression. Such coincidence causes the identical D optimal solutions.

If observations are also taken at the center of the design, these center runs would offer one lack-of-fit degree of freedom for a check of quadratic effect. Assuming equal number of center runs are used as for each corner point, the D efficiency of the design involving center runs relative to the one using the corner points only is expressed as

$$\frac{4}{5}\left(1+\frac{1}{16P(1-P)}(1+\frac{\beta_0^{*2}}{L^2})\right)^{\frac{1}{4}} \qquad (2.2.6)$$

where P and 1-P are the design probabilities and L = logit(P). For the optimal design where P = 0.823959, the efficiencies are evaluated in table 2.2.1. Higher efficiencies occur when the central logit β_0^* is far from zero. In general, using center runs does not seem to result in substantially less efficient designs. This gives more justification to the usual practice of applying center runs in almost all real experiments.

Table 2.2.1 D efficiency of hyperbola-based design, center runs vs. no center runs

$\lvert\beta_0^*\rvert$	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8	2.0
EFF	.876	.879	.885	.892	.901	.912	.923	.938	.953	.969

As with any other design in the logistic case, implementation of the hyperbola-based design also requires parameter knowledge. The effect of poor parameter guesses will be investigated later in the discussion of robustness. To construct a hyperbola-based design with t = 1, one must choose one of the design variables to be equidistant to the central level. Assuming that x_1 is the equidistant variable and the parameter values are supplied, the design points can be found using the design matrix

$$\begin{bmatrix} x_{11} & x_{21} \\ x_{12} & x_{22} \\ x_{13} & x_{23} \\ x_{14} & x_{24} \end{bmatrix} = \begin{bmatrix} -\frac{\beta_2}{\beta_{12}} - z_1 & -\frac{\beta_1}{\beta_{12}} - \frac{L-\beta_0^*}{\beta_{12} z_1} \\ -\frac{\beta_2}{\beta_{12}} - z_1 & -\frac{\beta_1}{\beta_{12}} + \frac{L-\beta_0^*}{\beta_{12} z_1} \\ -\frac{\beta_2}{\beta_{12}} + z_1 & -\frac{\beta_1}{\beta_{12}} - \frac{L-\beta_0^*}{\beta_{12} z_1} \\ -\frac{\beta_2}{\beta_{12}} + z_1 & -\frac{\beta_1}{\beta_{12}} + \frac{L-\beta_0^*}{\beta_{12} z_1} \end{bmatrix} \tag{2.2.7}$$

where $L = \text{logit}(0.823959) = 1.54341$ and z_1 is the user specified common deviation of x_1 from the central level.

Suppose that the parameters are given by $\beta_0 = -2.5$, $\beta_1 = 2$, $\beta_2 = 1.5$, and $\beta_{12} = -1$, and the variable x_2 is desired to be equidistant at ±1.46 from its central level, i.e. $z_2 = 1.46$. Then using (2.2.7) while paying attention to the position switch of the two variables, the design points can be obtained as (0.10, 0.54) and (2.90, 3.46) on $ED_{17.6}$ and (0.79, 3.46) and (2.21, 0.54) on $ED_{82.4}$. The hyperbola-based design with equal sample sizes appears in figure 2.2.3.

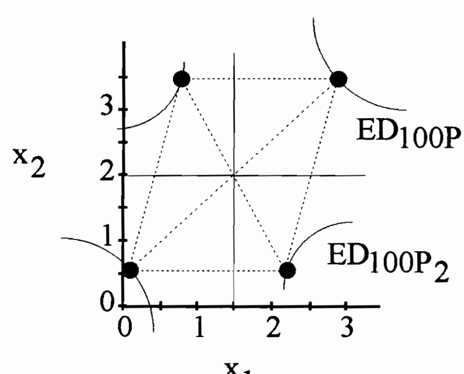


Figure 2.2.3 Example of hyperbola-based D optimal design

§2.3 Ray Design

Brunden *et al.* (1988) proposed the two-ray design as shown in figure 2.3.1. The design makes use of two parallel ED lines in the no-interaction situation as well as two rays extending from the origin. The four intersections are taken as design points. The two points on the same ED are assumed to have equal sample sizes. The two rays are assumed to be symmetric about the 45° line and thus their slopes can be denoted by g and 1/g.

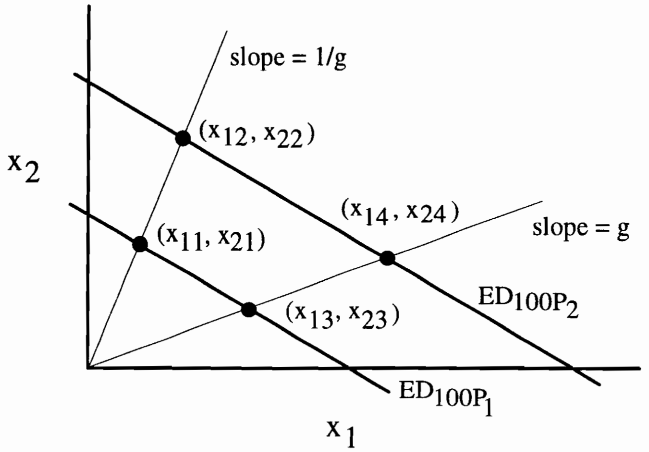


Figure 2.3.1 Ray design

The design criterion used by the authors is the determinant of the information matrix for fitting an interaction model while assuming no interaction in the true relationship. Under such a condition, the criterion is shown to be the following expression

$$|I(\beta)| = \frac{N^4}{16\beta_1^4\beta_2^4} f(k)\, r^2 P_1^2 (1-P_1)^2 (1-r)^2 P_2^2 (1-P_2)^2$$

$$\cdot[\log(\frac{P_1}{1-P_1}) - \log(\frac{P_2}{1-P_2})]^4 [\log(\frac{P_1}{1-P_1}) - \beta_0]^2 [\log(\frac{P_2}{1-P_2}) - \beta_0]^2, \qquad (2.3.1)$$

where r is the sample proportion for P_1 and

$$f(k) = \frac{k^4(1-k)^2(1+k)^2(1-g)^4(1+g)^4}{g^4(1+\frac{k}{g})^6(1+gk)^6} \qquad (2.3.2)$$

where $k = \frac{\beta_1}{\beta_2}$.

Maximizing f(k) with respect to g yields the optimal ray slope expressed as a function of k:

$$g(k) = \frac{1}{4}\left(\sqrt{(\sqrt{k}+\frac{1}{\sqrt{k}})^2+\sqrt{12+(k+\frac{1}{k})^2}} - \sqrt{(\sqrt{k}-\frac{1}{\sqrt{k}})^2+\sqrt{12+(k+\frac{1}{k})^2}}\right)^2. \qquad (2.3.3)$$

Maximizing the balance of the criterion expression in (2.3.1) gives an optimal sample proportion of r = 0.5 as well as the optimal ED's, which are however dependent upon the parameter β_0. Selected results are given in table 2.3.1. The parameter β_0 can also be interpreted as the logit at the origin. The response probability at the origin, given by $P_0 = \frac{1}{1+e^{-\beta_0}}$, is also listed in the table to give an idea how far the design lies away from the origin in terms of the response probability.

Table 2.3.1 Optimal probabilities for ray designs

β_0	P_0	P_1	P_2
0	0.5	0.710619	0.977885
-0.5	0.377541	0.633708	0.969554
-1.0	0.268941	0.560989	0.960133
-2.0	0.119203	0.445324	0.940530
-5.0	6.69285×10^{-3}	0.293432	0.897616
-10.0	4.53979×10^{-5}	0.231674	0.867236
-10^6	0.	0.176041	0.823960

In table 2.3.1, the extreme negative value -10^6 for β_0 is intended to show that as the response probability at the origin tends to 0, the optimal ED's of a ray design will approach those of the one-variable D optimal design, namely $ED_{17.6}$ and $ED_{82.4}$. In fact, the effective part of $|I(\beta)|$ with respect to P_1 and P_2 approaches the square of the one-variable D criterion as $\beta_0 \to \pm\infty$.

Construction of a ray design requires parameter values as in other designs for the logistic models. With the specified parameters, the ED lines can be found as $\beta_0 + \beta_1 x_1 + \beta_2 x_2 = L$, where L corresponds to logit(P_1) and logit(P_2) respectively. The ray slope can be computed from (2.3.3). Then some simple geometry will produce the design points. A potential drawback of the ray design lies in its total failure when k = 1, or $\beta_1 = \beta_2$. The condition k = 1 causes $|I(\beta)| = 0$, or a singular information matrix. This may imply that fitting an interaction model with proportional doses may not be statistically viable under equal linear effects.

To avoid repetition of the original work and to provide further insight, we derived the ray design in a more organized way, which parallels the presentation in the development of other designs in this research. Most of the derived expressions and discussed issues in this section are not found in the original work by Brunden *et al.* For instance, the original authors denoted the ray slope as kT and provided numerical results for the optimal T, while in this work an analytical solution as expressed in (2.3.3) is given for the optimal ray slope g. The results from using (2.3.3) match the original numerical results.

§2.4 Factorial Design

Brunden *et al.* (1988) also developed a 2×2 factorial design as shown in figure 2.4.1. The design forms a rectangle falling on three parallel ED lines. The sample proportions for each point on P_1, P_2, and P_3 are denoted r_1, r_2, and r_3 respectively. The total allocation on P_2 is $2r_2$.

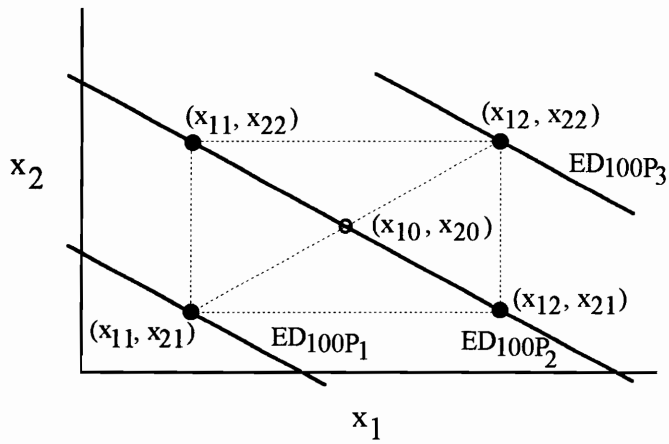


Figure 2.4.1 Factorial design

The rectangular design implies that

$$P_2 = \frac{\sqrt{P_1P_3}}{\sqrt{P_1P_3} + \sqrt{(1-P_1)(1-P_3)}} . \tag{2.4.1}$$

The design criterion is derived similarly to the ray design situation, namely fitting an interaction model while assuming $\beta_{12} = 0$. The D criterion so defined is shown to be

$$|I(\beta)| = \frac{N^4}{4^5\beta_1^4\beta_2^4} r_1r_3(1 - r_1 - r_3)^2$$

$$\cdot \frac{P_1^2P_3^2(1-P_1)^2(1-P_3)^2}{\left(\sqrt{P_1P_3} + \sqrt{(1-P_1)(1-P_3)}\right)^4} [\log(\frac{P_3}{1-P_3}) - \log(\frac{P_1}{1-P_1})]^2 . \tag{2.4.2}$$

Maximizing $|I(\beta)|$ and then making use of the relationship in (2.4.1) leads to the optimal design given by

$$P_1 = 0.015818, \qquad P_2 = 0.5, \qquad P_3 = 0.984182,$$

$$\text{and} \qquad r_1 = r_2 = r_3 = 0.25 . \tag{2.4.3}$$

The design has symmetric ED's and equal sample sizes.

Assuming that the parameters are supplied, a factorial design can be constructed using the design matrix

$$\begin{bmatrix} x_{11} & x_{21} \\ x_{11} & x_{22} \\ x_{12} & x_{21} \\ x_{12} & x_{22} \end{bmatrix} = \begin{bmatrix} x_{10} - \frac{L}{2\beta_1} & -\frac{\beta_0 + \beta_1 x_{10}}{\beta_2} - \frac{L}{2\beta_2} \\ x_{10} - \frac{L}{2\beta_1} & -\frac{\beta_0 + \beta_1 x_{10}}{\beta_2} + \frac{L}{2\beta_2} \\ x_{10} + \frac{L}{2\beta_1} & -\frac{\beta_0 + \beta_1 x_{10}}{\beta_2} - \frac{L}{2\beta_2} \\ x_{10} + \frac{L}{2\beta_1} & -\frac{\beta_0 + \beta_1 x_{10}}{\beta_2} + \frac{L}{2\beta_2} \end{bmatrix} \qquad (2.4.4)$$

where $L = \text{logit}(P_3) = 4.13068$ and x_{10} is a user selected central level for x_1.

In this section we derived the 2×2 factorial design in a way that better fits into the current research and meanwhile provided supplementary discussion to the original work.

§2.5 Alternative Parallel-Line Design

A parallel-line design can also be optimized using the type of criterion adopted by the authors of the ray design. In such a criterion, the information matrix comes from an interaction model while the true relationship is assumed to contain no interaction effect. Given the parallel-line design structure, the D criterion formed in this fashion is shown to be

$$|I(\beta)| = N^4 \left(\frac{D_1^4}{\beta_2^4}\right) \left(\frac{k+k'}{k-k'}\right)^2 [rP_1(1-P_1)(1-r)P_2(1-P_2)]^2 \left[\log\left(\frac{P_2}{1-P_2}\right) - \log\left(\frac{P_1}{1-P_1}\right)\right]^2 \qquad (2.5.1)$$

where $k = -\frac{\beta_1}{\beta_2}$ and k' is the slope of a non-ED edge as indicated in figure 2.1.1.

Normally the design would be made rectangular with $k' = -1/k$ and thus the factor $\frac{k+k'}{k-k'}$ in (2.5.1) becomes $\frac{k^2-1}{k^2+1}$. However, when $k = \pm 1$, the rectangular design will lead to a singular information matrix with $|I(\beta)| = 0$. An alternative is to make the non-ED edges either horizontal or vertical, i.e. to use $k' = 0$ or ∞. Using the convention that $k' = -\frac{1}{k}$ if $k \neq \pm 1$ while $k' = 0$ or ∞ if $k = \pm 1$, the factor $\frac{k+k'}{k-k'}$ in (2.5.1) can be replaced by

$$f(k) = \begin{cases} \dfrac{k^2-1}{k^2+1}, & k \neq \pm 1 ; \\ 1 , & k = \pm 1 . \end{cases} \tag{2.5.2}$$

Apart from a constant and f(k), the rest of the $|I(\beta)|$ expression coincides with the square of the one-variable D criterion. Therefore the same optimal results as in the one-variable case are obtained below:

$$P_1 = 0.176041, \qquad P_2 = 0.823959, \qquad \text{and} \qquad r_1 = r_2 = 0.5 . \tag{2.5.3}$$

As the interaction effect is added to the fitted model, the dispersion of the design becomes greater than that of the regular parallel-line design.

Center runs can also be used to check for quadratic effect. Here to evaluate the D efficiency, the D criterion is based on fitting an interaction model while assuming a true relationship that involves no interaction, namely the same condition that yielded the optimal design. If the center point still receives the same sample size as other design points, the D efficiency of a design with center runs relative to the corresponding one without center runs is expressed as

$$\frac{4}{5}\left(1+\frac{1}{16P(1-P)}(1+f(k,g))\right)^{\frac{1}{4}} \tag{2.5.4}$$

where P and 1-P are the response probabilities that the design resides on. The function f(k, g) in (2.5.4) is given by

$$f(k, g) = \begin{cases} \left((\frac{L_D}{L})\frac{k^2+1}{k^2-1} - (\frac{L}{L_D})\frac{k^2}{(k^2+1)(k^2-1)} \right)^2, & k \neq \pm 1; \\ (\frac{L_D}{L})^2, & k = \pm 1, \end{cases} \quad (2.5.5)$$

where $k = \frac{\beta_1}{\beta_2}$, $L_D = |\beta_1 D_1| = |\beta_2 D_2|$, and $L = \text{logit}(P)$.

Table 2.5.1 D efficiency of alternative parallel-line design, center runs vs. no center runs

	L_D									
|k|	.2	.4	.6	.8	1.0	1.2	1.4	1.6	1.8	2.0
0.1	0.88	0.88	0.88	0.89	0.90	0.91	0.93	0.94	0.96	0.97
0.3	0.89	0.88	0.88	0.89	0.90	0.92	0.93	0.95	0.97	0.99
0.5	1.04	0.90	0.88	0.88	0.90	0.93	0.96	0.99	1.03	1.06
0.7	1.44	1.03	0.89	0.88	0.92	0.99	1.06	1.13	1.19	1.26
0.9	2.68	1.70	1.12	0.89	1.15	1.41	1.63	1.81	1.97	2.12
1.0	0.88	0.88	0.88	0.89	0.90	0.91	0.92	0.94	0.95	0.97

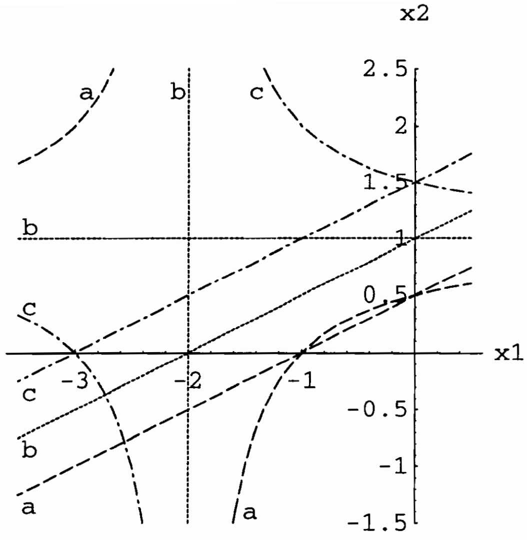
Table 2.5.1 evaluates the efficiency due to center runs for the optimal design where P = 0.823959. The figures in this table indicate that the efficiency of using center runs tends to be higher when the length of the design in the direction of the ED lines is greater in terms of the scale free, or logit type, distance L_D. The efficiency also appears to be increasing when the linear coefficients β_1 and β_2 are more alike in their values, which is indicated by a value of k closer to 1.

§2.6 Modified Ray Design

When interaction enters the model as an additional effect, each ED line is turned into a pair of hyperbolae. Figure 2.6.1 illustrates the geometric nature of the one-to-one correspondence between a line and a pair of hyperbolae with identical logit levels. A nice feature is that a hyperbola curve, whenever crossing a natural axis of x_1 or x_2, will cross the axis at the same point(s) as the original line with the equal logit value has passed by. The line with a constant logit equal to β_0^*, or $\beta_0 - \frac{\beta_1\beta_2}{\beta_{12}}$, separates all lines of constant logits into two categories. This separation line crosses the x_1 and x_2 axes at the same points as the centered axes for the hyperbolae do. The lines on one side of the separation line, the side that holds the origin, are turned into pairs of hyperbolae such that only one curve in each pair passes the x_1 and x_2 axes as well as the original line, certainly at common points. Those lines on the other side of the separation line correspond to pairs of hyperbolae such that both curves in a pair each intersect one of the x_1 and x_2 axes and the original line in the meantime.

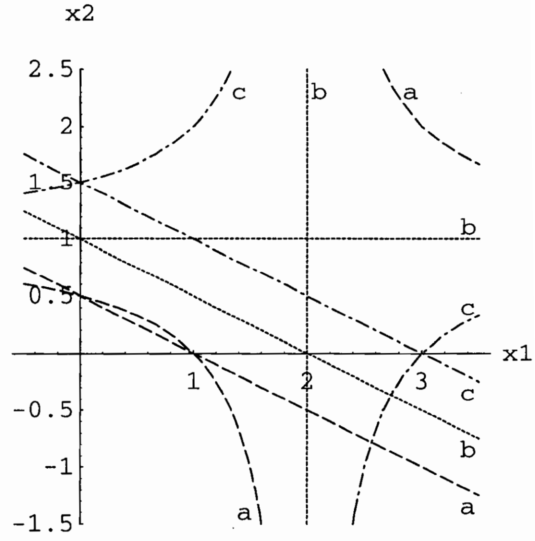
Each graph in figure 2.6.1 shows three matches of a line and the hyperbolae, with the logits indicated by a, b, and c respectively. The logit b represents β_0^* and corresponds to the separation line as well as the centered axes.

Figure 2.6.1 also shows the position of the hyperbola center in different quadrants depending on the signs of the parameters. Graph (3) illustrates a scenario of synergism, where the hyperbola center is found in the third quadrant. Antagonism is shown in graph (2) with the hyperbola center located in the first quadrant.



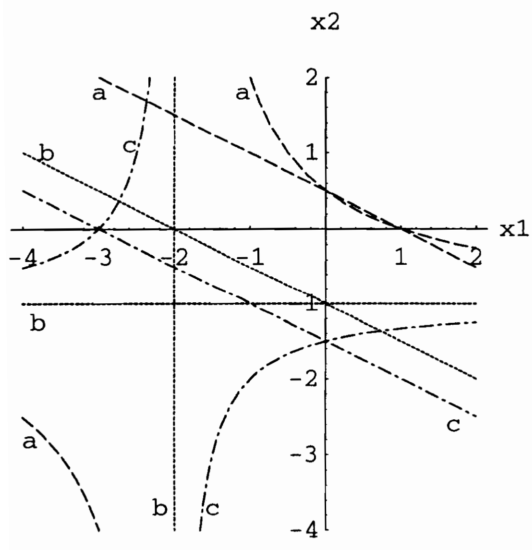
(1)

β1	β2	β12
−	+	+
+	−	−



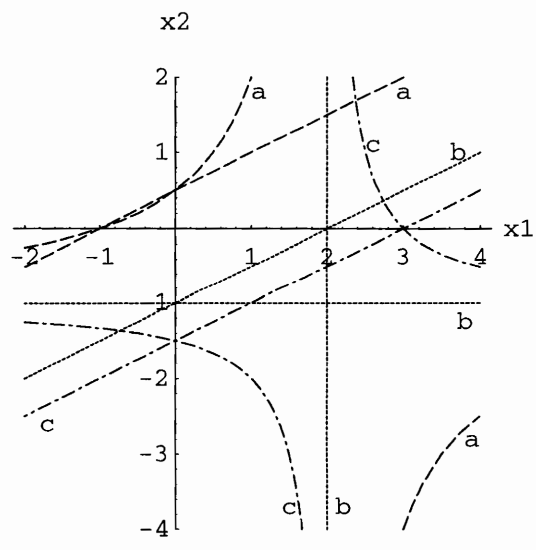
(2)

β1	β2	β12
+	+	−
−	−	+



(3)

β1	β2	β12
+	+	+
−	−	−



(4)

β1	β2	β12
−	+	−
+	−	+

Figure 2.6.1 Geometric interpretation of interaction effect

The graphs in figure 2.6.1 are intended to show the four general situations that an interaction circumstance can fall in. The specific parameter values used to create these examples are given in table 2.6.1. Also listed are the hyperbola centers (x_{10}, x_{20}) and the selected logit values a, b, and c indicated in the graphs.

Table 2.6.1 Examples illustrated in figure 2.6.1

Graph	β_0	β_1	β_2	β_{12}	(x_{10}, x_{20})	a	b(β_0^*)	c
(1)	-1	-1	2	1	(-2, 1)	0	1	2
(2)	-1	1	2	-1	(2, 1)	0	1	2
(3)	-1	1	2	1	(-2, -1)	0	-3	-4
(4)	-1	-1	2	-1	(2, -1)	0	-3	-4

The ray designs developed by Brunden *et al.* are intended to handle situations where potential interaction is concerned. In a ray design scenario, the interaction can be either synergistic or antagonistic. The ray design is created by assuming no interaction in the true relationship, though the dimension of the information matrix comprises the interaction coefficient. With such designs, there seems to be an inconsistency between the goal dealing with interaction and the optimization environment assuming a no-interaction state of nature. Therefore it might be suspected that the ray design may not perform as well compared to the hyperbola-based design in the presence of interaction. In fact, it will be shown later that when interaction exits, the D efficiencies of ray designs relative to the hyperbola-based design are often unsatisfactory. The efficiency falls extremely low when the interaction is synergistic, possibly because the ray design locates farther away from the hyperbola center in the third quadrant under synergism than from the one in the first quadrant under antagonism.

The ineffectiveness of the ray design under synergism is also revealed in the work by Brunden *et al.* They computed an "efficiency" which is the ratio of the D criterion of a ray design under interaction relative to the D criterion of the same design under the no-interaction situation. Though these ratios are not proper efficiencies that compare two different designs, they do indicate that the D criterion of a ray design usually goes down quickly when synergism forms and becomes more severe. However if the interaction starts to grow toward the antagonism direction, the D criterion of a ray design will be increasing and will begin to decrease at some point when the antagonism becomes too strong.

In spite of its difficulty with synergism, the ray design seems attractive to practitioners for some practical reasons. First, design points on the same ray receive proportional doses since the ratio x_1/x_2 stays constant anywhere on the ray. Once the experimental mixture at one point is administered, the same x_1/x_2 mixture will also apply to the other point on the ray and usually it is just a matter of adjusting the amount of mixture. The entire experimental task is therefore greatly reduced. The ray type designs are also useful in some experiments by offering all positive design levels. In a drug experiment, if the natural doses are to be used as regressor variables, then negative design levels are not allowed. This may not always be an issue as centered variables or log doses are often used in practice. Using log doses in a logistic model can sometimes produce better fit and more closely represent the unknown relationship. Finney (1978) pointed out that in bioassays, an analysis in terms of the log doses may be preferable to one in absolute units, and that normality of tolerance distribution is likely to be better approximated by log doses.

Of the two types of interaction effect, antagonism does not pose as much concern since the efficiency of the ray design with antagonism is less inferior and meanwhile a hyperbola-based design can often be found even if positive doses are required. Synergism however can be problematic due to poor efficiencies of the ray designs and lack of a practical hyperbola-based design alternative with all positive design levels.

The goal of this section is to better handle the synergistic condition while still pursuing the practical advantages of a ray type design, which are desired not in all but in certain experimental situations as mentioned above. The proposed strategy will be referred to as a modified ray design, which still makes use of symmetric rays as in the original ray design. The modification focuses on removal of the inconsistency explained earlier between the assumption and the goal of the ray design. This is achieved by using the hyperbolic ED curves rather than the ED lines, namely by assuming the presence of interaction in the true state of nature. Figure 2.6.2 shows the design structure. The slopes of the rays are denoted by g and 1/g. The probabilities of response on the two curves are P_1 and P_2. The sample sizes for the two points on a common ED are assumed equal.

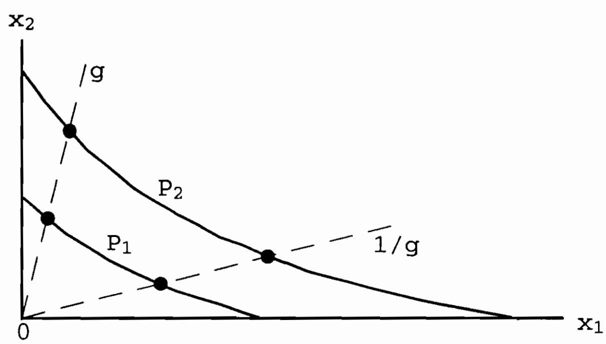


Figure 2.6.2 Modified ray design

Define a function S as follows:

$$S(g, L) = \sqrt{(1+gk)^2 + 4gk\frac{L-\beta_0}{\Delta_\beta}} - (1 + g\,k) \qquad (2.6.1)$$

where $k = \dfrac{\beta_1}{\beta_2}$ and Δ_β is the difference in logit between the origin and the hyperbola center, which is given by $\Delta_\beta = \beta_0 - \beta_0^* = \dfrac{\beta_1\beta_2}{\beta_{12}}$. The model matrix for a modified ray design can then be written as

$$X = \begin{bmatrix} 1 & \frac{\beta_2}{2\beta_{12}} S(\frac{1}{g}, L_1) & \frac{g\beta_2}{2\beta_{12}} S(\frac{1}{g}, L_1) & g\left(\frac{\beta_2}{2\beta_{12}} S(\frac{1}{g}, L_1)\right)^2 \\ 1 & \frac{\beta_2}{2\beta_{12}} S(\frac{1}{g}, L_2) & \frac{g\beta_2}{2\beta_{12}} S(\frac{1}{g}, L_2) & g\left(\frac{\beta_2}{2\beta_{12}} S(\frac{1}{g}, L_2)\right)^2 \\ 1 & \frac{\beta_2}{2\beta_{12}} S(g, L_1) & \frac{\beta_2}{2g\beta_{12}} S(g, L_1) & \frac{1}{g}\left(\frac{\beta_2}{2\beta_{12}} S(g, L_1)\right)^2 \\ 1 & \frac{\beta_2}{2\beta_{12}} S(g, L_2) & \frac{\beta_2}{2g\beta_{12}} S(g, L_2) & \frac{1}{g}\left(\frac{\beta_2}{2\beta_{12}} S(g, L_2)\right)^2 \end{bmatrix} \tag{2.6.2}$$

where L_1 and L_2 are the logits of P_1 and P_2. The determinant of X is given by

$$|X| = \left(\frac{\Delta_\beta}{4k\beta_{12}}\right)^2 [S(g, L_1) - S(g, L_2)][S(\frac{1}{g}, L_1) - S(\frac{1}{g}, L_2)]$$

$$\cdot [(\frac{1}{g^2} - 1)\, S(g, L_1)S(g, L_2) + (g^2 - 1)S(\frac{1}{g}, L_1)S(\frac{1}{g}, L_2)]\,. \tag{2.6.3}$$

Due to $|I(\beta)| = |X'\Sigma X|$, the D criterion of a modified ray design is given by

$$|I(\beta)| = \left(\frac{N}{2}\right)^4 [\, r(1-r)\, P_1(1-P_1)\, P_2(1-P_2)\, |X|\,]^2 \tag{2.6.4}$$

where N is the total sample size and r is the sample proportion assigned to P_1.

The D criterion in (2.6.4) is a function of the model parameters β_0, Δ_β, and k as well as the design indices P_1, P_2, r, and g. The optimal P_1, P_2, r, and g will then depend on β_0, Δ_β, and k. Since $|I(\beta)|$ is invariant when k is replaced by 1/k, only the $k \leq 1$ case will be investigated. As for the original ray designs, only negative values of β_0 will be addressed based on the practical consideration that the probability of response at the origin is normally below 0.5. The assumption of positive linear effects was implicitly used in the work of Brunden *et al.* on the original ray designs and must have been based on practical appropriateness. Since currently we are in the situation of synergism, using the assumption of positive linear effects implies that the parameter Δ_β takes on positive values. The optimal sample proportion is always 50% at each probability level in all cases. The optimal probabilities and ray slope given values of β_0, Δ_β, and k are listed in table 2.6.2.

Table 2.6.2 indicates that both the lower and upper design probabilities decrease and their dispersion widens as the probability of response at zero doses goes down. This is quite intuitive since by assumption the design probabilities have to lie above the probability level at zero doses. Slight decrease of the design levels also occurs when Δ_β becomes smaller. Recall that Δ_β is the difference in logit between the origin and the hyperbola center. This implies that as the hyperbola center (in the third quadrant) moves closer to the origin in terms of the logit distance, the design will also draw closer towards the origin, though the movement appears much smaller than that caused by the variation in β_0. The design probabilities are quite invariant to the ratio $k = \frac{\beta_1}{\beta_2}$. However, the impact of k is fairly obvious on the position of the rays through the ray slope g. As the linear parameters become more unbalanced, indicated by values of k remotely away from 1 in either direction, the rays are required to be more open, or nearer to the x_1 and x_2 axes. Assuming that the parameters are known, the design can be implemented by using the second and third columns of the model matrix given in (2.6.2) as the design matrix to compute the design points.

Table 2.6.2 Modified ray designs

β_0	Δ_β	k	P_1	P_2	g
0	5	.1	.697953	.973191	24.7503
		.3	.697487	.972940	10.6253
		.5	.697144	.972749	8.45559
		.7	.696974	.972653	7.78838
		.9	.696911	.972617	7.58271
	2	.1	.687681	.969538	29.8512
		.3	.687105	.969069	13.3580
		.5	.686736	.968745	10.8683
		.7	.686565	.968588	10.1045
		.9	.686504	.968531	9.86915
	1	.1	.678292	.966311	37.5012
		.3	.677935	.965674	17.5635
		.5	.677762	.965273	14.5994
		.7	.677695	.965085	13.6911
		.9	.677673	.965018	13.4112
	.5	.1	.669025	.963215	51.7133
		.3	.669293	.962488	25.5569
		.5	.669526	.962071	21.7252
		.7	.669650	.961883	20.5516
		.9	.669697	.961816	20.1897
-1	5	.1	.540937	.951946	25.4451
		.3	.540215	.951523	10.9871
		.5	.539693	.951208	8.77299
		.7	.539435	.951049	8.09269
		.9	.539341	.950991	7.88306
	2	.1	.525777	.945976	31.5896
		.3	.524942	.945244	14.2769
		.5	.524420	.944751	11.6759
		.7	.524181	.944514	10.8791
		.9	.524097	.944428	10.6337
	1	.1	.512789	.940992	41.0379
		.3	.512323	.940078	19.4541
		.5	.512110	.939516	16.2639
		.7	.512031	.939256	15.2879
		.9	.512005	.939163	14.9873
	.5	.1	.500841	.936494	58.9948
		.3	.501239	.935545	29.4789
		.5	.501573	.935012	25.1800
		.7	.501749	.934774	23.8661
		.9	.501814	.934689	23.4614

Table 2.6.2 (continued)

β_0	Δ_β	k	P_1	P_2	g
-2	5	.1	.420300	.928965	26.5083
		.3	.419426	.928395	11.5422
		.5	.418806	.927979	9.26018
		.7	.418503	.927772	8.55983
		.9	.418392	.927696	8.34410
	2	.1	.403245	.921308	34.2826
		.3	.402312	.920417	15.7046
		.5	.401746	.919833	12.9312
		.7	.401491	.919557	12.0830
		.9	.401401	.919457	11.8219
	1	.1	.389889	.915414	46.5642
		.3	.389413	.914408	22.4129
		.5	.389207	.913808	18.8677
		.7	.389132	.913534	17.7854
		.9	.389109	.913437	17.4523
	.5	.1	.378669	.910513	70.4217
		.3	.379045	.909575	35.6260
		.5	.379374	.909059	30.5880
		.7	.379542	.908830	29.0513
		.9	.379605	.908749	28.5783
-5	5	.1	.268844	.882230	31.1186
		.3	.268076	.881584	13.9684
		.5	.267567	.881144	11.3922
		.7	.267325	.880932	10.6064
		.9	.267239	.880855	10.3623
	2	.1	.257492	.875206	46.0443
		.3	.256853	.874459	21.9914
		.5	.256496	.874009	18.4647
		.7	.256339	.873803	17.3907
		.9	.256285	.873730	17.0606
	1	.1	.250718	.871021	70.6319
		.3	.250426	.870360	35.3723
		.5	.250303	.869991	30.2763
		.7	.250258	.869826	28.7265
		.9	.250243	.869768	28.2501
	.5	.1	.246165	.868209	119.624
		.3	.246261	.867693	62.1822
		.5	.246357	.867419	53.9525
		.7	.246410	.867299	51.4497
		.9	.246430	.867256	50.6803

CHAPTER 3

EQUIVALENCE THEORY FOR D OPTIMAL DESIGNS

The equivalence of certain properties associated with a D optimal design leads to deeper insight into design optimality criteria, especially for logistic models. Quite often a D optimal design may have to be found numerically or with restriction. This is seen in the designs given in chapter 2 as well as in many other cases not involved in this research. In such a case, the exact D optimal property of the design may not be clear enough. The equivalence theory can often provide alternative ways for checking D optimality which are much more convenient than the D criterion itself. The design equivalence theory reveals the relationship between the parameter oriented D criterion and other prediction oriented performance measures.

§3.1 Design Equivalence Overview

Kiefer and Wolfowitz (1960) introduced the equivalence theorem for linear models. Federov (1972) developed the equivalence theorem in the general framework of design optimization, which accommodates nonlinear models as well.

In Federov's work, a design is expressed as a probability measure $\xi(x)$ defined on a sample space R, satisfying the conditions

$$\int_R d\xi(x) = 1, \qquad \xi(x) \geq 0, \qquad x \in R . \qquad (3.1.1)$$

In this definition, $\xi(x)$ could represent the proportion of sample size assigned to the point x for a design with isolated points, as usually seen in practical design optimization, though in the broad sense the probability measure $\xi(x)$ could characterize any design, not necessarily one with isolated points.

The D optimal design for a preselected region R and a model with parameter β is defined as the probability measure $\xi^*(x)$ which maximizes $|I(\xi; \beta)|$, the determinant of the information matrix for β due to design ξ, among all probability measures defined on the sample space R. This definition of D optimality takes the form

$$|I(\xi^*; \beta)| = \max_{\xi \in \Xi_R} |I(\xi; \beta)| \tag{3.1.2}$$

where Ξ_R is the set of all probability measures defined on the sample space R.

A generalized scaled prediction variance $V(\xi; x)$ is defined as

$$V(\xi; x) = N \, Tr[I(\xi; \beta)^{-1} J(x; \beta)] \tag{3.1.3}$$

where $I(\xi; \beta)$ is the information matrix for β due to the design ξ, $J(x; \beta)$ is the information matrix for β due to a single observation at the point x, and N is the total number of observations in the design ξ. Both $I(\xi; \beta)$ and $J(x; \beta)$ are p×p matrices, where p is the number of parameters in the vector β.

The **equivalence theorem** proved by Federov states that the following assertions are equivalent:

(1) the design ξ^* maximizes $|I(\xi; \beta)|$: $|I(\xi^*; \beta)| = \max_{\xi \in \Xi_R} |I(\xi; \beta)|$,

(2) the design ξ^* minimizes $\max_{x \in R} V(\xi; x)$: $\max_{x \in R} V(\xi^*; x) = \min_{\xi \in \Xi_R} \max_{x \in R} V(\xi; x)$,

(3) $\max_{x \in R} V(\xi^*; x) = p$. (3.1.4)

The equivalence guarantees that any one of the three properties implies the other two. The design ξ^* in condition (1) defines a D optimal design for the region R.

Condition (2) is related to G optimality but not guaranteed to be the same. Condition (3) involves only the design itself and thus can be verified without examining any competing designs. When the equivalence theorem is applied later in this chapter, condition (3) will play an important role. The equivalence implies that conceptually a D optimal design can also be solved by property (2) or (3), as can be seen later with several designs.

§3.2 Design Equivalence for Linear Model

For the linear model, since $I(\xi;\ \beta) = \frac{1}{\sigma^2}X'X$ and $J(x;\ \beta) = \frac{1}{\sigma^2}xx'$, the generalized scaled prediction variance is given by

$$V(\xi;\ x) = Nx'(X'X)^{-1}x\ . \qquad (3.2.1)$$

The generalized scaled prediction variance $V(\xi;\ x)$ coincides with the scaled prediction variance addressed in the G optimality. Therefore property (2) of the equivalence theorem corresponds to G optimality and thus properties (1) through (3) can be put as

$$(1)\ \xi^* \text{ is D optimal,} \qquad (3.2.2)$$

$$(2)\ \xi^* \text{ is G optimal,} \qquad (3.2.3)$$

$$(3)\ \underset{x \in R}{\mathrm{Max}}\, V(\xi^*;\ x) = p\ . \qquad (3.2.4)$$

The equivalence theorem indicates that for the standard linear model, a D optimal design for a given region is also G optimal for that region.

For instance, when the linear model is first order with two design variables and the region R is rectangular, it is well known that a four-point design placed at the region vertices with equal sample sizes would be both D and G optimal. It may be less obvious to recognize properties (1) and (2), which involves comparisons among all possible designs on the region R. Property (3) though can be seen very easily and D and G optimalities would then follow from the equivalence theorem.

For a design ξ^* with equal sample sizes placed at the corners of a rectangular region bounded by $x_1 \in [x_{11}, x_{12}]$ and $x_2 \in [x_{21}, x_{22}]$, the condensed model matrix X and the vector x due to one observation can be expressed as

$$X = ZA' \qquad \text{and} \qquad x = Az \tag{3.2.5}$$

respectively where

$$A = \begin{bmatrix} 1 & 0 & 0 \\ \frac{x_{11}+x_{12}}{2} & \frac{x_{12}-x_{11}}{2} & 0 \\ \frac{x_{21}+x_{22}}{2} & 0 & \frac{x_{22}-x_{21}}{2} \end{bmatrix} \quad \text{and} \quad Z = \begin{bmatrix} 1 & -1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix}. \tag{3.2.6}$$

From (3.2.1), it follows that

$$V(\xi^*; x) = Nx'(\frac{N}{4}X'X)^{-1}x = 4z'A'(AZ'ZA')^{-1}Az = 1 + z_1^2 + z_2^2 . \tag{3.2.7}$$

Obviously, $\underset{x \in R}{\text{Max}}\, V(\xi^*; x) = p$, where p = 3 for the two-variable first-order model. And the maximums occur at the design points. It is not important where the maximum is achieved. What matters is that the maximum scaled prediction variance over the entire region is equal to p, i.e. condition (3) holds for the design ξ^*. By means of the equivalence theorem, once property (3) gets verified, the design ξ^* can then be said to be both D and G optimal.

The scaled prediction variance at a design point is related to the hat diagonal h_{ii} by

$$V(\xi; x_i) = N\, h_{ii} \tag{3.2.8}$$

where $h_{ii} = x_i'(X'X)^{-1}x_i$. Since $\sum_{i=1}^{N} h_{ii} = p$, a design with equal hat diagonals $h_{ii} = p/N$, i = 1, ..., N, always satisfies $V(\xi; x_i) = p$ at any design point. However, such a design has nothing to do with D optimality since the maximum $V(\xi^*; x)$ over the region can still exceed p. For example, in the previous situation with the rectangular region R, a design

ξ_1 with four evenly weighted design points forming a smaller rectangle lying inside and concentric with the region R corresponds to

$$Z = \begin{bmatrix} 1 & -a & -b \\ 1 & -a & b \\ 1 & a & -b \\ 1 & a & b \end{bmatrix} \tag{3.2.9}$$

where $|a| < 1$ and $|b| < 1$. Equation (3.2.1) implies that

$$V(\xi_1; x) = 1 + \frac{z_1^2}{a^2} + \frac{z_2^2}{b^2} . \tag{3.2.10}$$

It follows that $V(\xi_1; x) = p = 3$ at the design points. However, the design ξ_1 can not be D or G optimal since at least, $V(\xi^*; x)$ is above 3 anywhere outside the design rectangle but within the region.

§3.3 Design Equivalence for Logistic Model

For logistic models, the information matrix due to the design ξ is $I(\xi; \beta) = X'\Sigma X$. The information matrix due to one observation at the point x is $J(x; \beta) = \sigma^2 xx'$, where $\sigma^2 = P(1-P)$ and P is the probability that the observation is equal to 1 at the point x. Hence the scaled prediction variance is given by

$$V(\xi; x) = N\sigma^2 x'(X'\Sigma X)^{-1} x . \tag{3.3.1}$$

The equivalence theorem applied to a logistic model then says that a D optimal design ξ^* for the region R achieves the following simultaneously:

$$(1)\ |I(\xi^*; \beta)| = \underset{\xi \in \Xi_R}{\text{Max}}\, |X'\Sigma X| , \tag{3.3.2}$$

$$(2)\ \underset{x \in R}{\text{Max}}\, V(\xi^*; x) = \underset{\xi \in \Xi_R}{\text{Min}}\ \underset{x \in R}{\text{Max}}\ N\sigma^2 x'(X'\Sigma X)^{-1} x , \tag{3.3.3}$$

$$(3)\ \underset{x \in R}{\text{Max}}\, V(\xi^*; x) = p . \tag{3.3.4}$$

In the logistic case, G optimality based on the logit addresses $\mathrm{NVar}[\mathrm{logit}(\hat{y})] = \mathrm{N}x'(X'\Sigma X)^{-1}x$ while G optimality based on the probability addresses $\mathrm{NVar}[\hat{y}] = \mathrm{N}(\sigma^2)^2 x'(X'\Sigma X)^{-1}x$. Neither of the scaled prediction variances involved in the two forms of G optimality corresponds to $V(\xi; x)$ as in (3.3.1). Therefore property (2) of the equivalence theorem does not represent G optimality in the above senses. In fact, $V(\xi; x)$ is the geometric average $\sqrt{(\mathrm{NVar}[\hat{y}])(\mathrm{NVar}[\log\mathrm{it}(\hat{y})])}$.

If h_{ii} denotes a hat diagonal based on the hat type matrix $H = \Sigma^{1/2}X(X'\Sigma X)^{-1}X'\Sigma^{1/2}$, then $V(\xi; x_i) = N\, h_{ii}$ at a design point. Since $\sum_{i=1}^{N} h_{ii} = p$, a design with equal hat type diagonals where $h_{ii} = p/N$, $i = 1, \ldots, N$, would have $V(\xi; x_i) = p$ at all design points. However the condition $V(\xi; x_i) = p$ at the design points will not bring the design any closer to being D optimal since D optimality requires $\underset{x \in R}{\mathrm{Max}}\, V(\xi; x) = p$, which is not guaranteed in any way by such a design.

§3.4 Parameterization and Design Equivalence

The equivalence theorem discussed in the previous section essentially enunciated three design properties that occur simultaneously. The entire proposition is based on the same parameterization. Reparameterization for a given model, sometimes in order to accommodate practical concerns or specific analyses, also bears its implication to design equivalence. Raising the issue of parameterization opens another dimension in the study of design equivalence.

Suppose that β is the initial parameterization for a given model. Let another parameterization α for the same model be defined as a function of β with the Jacobian matrix

$$C = \frac{\partial \alpha}{\partial \beta'} = [\frac{\partial \alpha_i}{\partial \beta_j}]_{p \times p} , \tag{3.4.1}$$

where p is the number of parameters. For a proper reparameterization α, the matrix C is of full rank and is independent of the design characteristics involved in the optimization (not necessarily independent of parameters). The information matrix $I(\xi; \alpha)$ for the transformed parameter α due to the design ξ will be different from but related to the information matrix $I(\xi; \beta)$ for the initial parameter β due to the same design. Let ℓ be the joint log likelihood for all observations of the design. The following relationship holds:

$$I(\xi; \beta) = E[\frac{\partial \ell}{\partial \beta} \frac{\partial \ell}{\partial \beta'}] = E[\frac{\partial \alpha'}{\partial \beta} \frac{\partial \ell}{\partial \alpha} \frac{\partial \ell}{\partial \alpha'} \frac{\partial \alpha}{\partial \beta'}] = C' \, I(\xi; \alpha) \, C ,$$

and thus

$$I(\xi; \alpha) = (C')^{-1} I(\xi; \beta) C^{-1} \tag{3.4.2}$$

with the determinant

$$|I(\xi; \alpha)| = |C|^{-2} \, |I(\xi; \beta)| . \tag{3.4.3}$$

More insight about the information matrix under reparameterization is discussed in Lehmann (1983).

The relationship in (3.4.3) indicates that the same design will yield different amounts of information regarding different parameterizations in terms of the determinant of the information matrix, unless the reparameterization maintains an orthogonal matrix C. That is, for a given design, the D criteria for different parameterizations will not have the same value unless C is orthogonal. However, since the matrix C under proper reparameterization does not involve any design parameters subject to D optimization, the D optimal design for β, which maximizes $|I(\xi; \beta)|$, will also be a D optimal design for α which maximizes $|I(\xi; \alpha)| = |C|^{-2} \, |I(\xi; \beta)|$.

The result in (3.4.2) implies that

$$J(z; \alpha) = (C')^{-1}J(x; \beta)C^{-1} \tag{3.4.4}$$

where $J(x; \beta)$ and $J(z; \alpha)$ are the information matrices for β and α respectively due to one observation at the common point identified by x with parameter β and by z with parameter α. From (3.4.2) and (3.4.4), it follows that the scaled prediction variance evaluated under parameterization α, denoted $V_\alpha(\xi; z)$, proves to be identical to the result evaluated under parameterization β, denoted $V_\beta(\xi; x)$:

$$\begin{aligned} V_\alpha(\xi; z) &= NTr[I(\xi; \alpha)^{-1}J(z; \alpha)] \\ &= NTr[C\, I(\xi; \beta)^{-1}C'(C')^{-1}J(x; \beta)C^{-1}] \\ &= NTr[I(\xi; \beta)^{-1}J(x; \beta)] \\ &= V_\beta(\xi; x)\ . \end{aligned} \tag{3.4.5}$$

It is found from the above that under proper reparameterization, i.e. $|C| \neq 0$ and C being independent of the design optimization,

(1) the D criterion value ($|I(\xi; \beta)|$) for a given design is invariant only up to reparameterizations such that the derivative matrix C is orthogonal,

(2) the solution of a D optimal design is invariant to any reparameterization,

(3) the scaled prediction variance for a given design is invariant to any reparameterization.

These invariance properties to reparameterizations imply that a D optimal design can be found by properly choosing a convenient parameterization. Once the design is found, it will remain D optimal and will thereby enjoy the equivalence property for any proper reparameterization, although the value of the D criterion will vary with the parameterization. The above discussion of parameterization and design equivalence assumes a general situation. Thus the result applies to both linear and logistic models.

In the following sections concerning logistic models, linear transformations of natural design variables into centered and scaled variables are often used to ease analyses. Such transformations in general take the form $Z = XA^{-1}$, where Z is the centered and scaled model matrix. This implies $\alpha = A\beta$. The logit is given by $L = x'\beta = z'\alpha$. One must realize that the vector α here is not a reparameterization since the matrix A heavily involves design characteristics subject to optimization. This implies that the true value of α changes from design to design. In fact, a set of model parameters in its statistical sense should be a set of unknown constants regardless of where observations are taken. D optimality aims at seeking a design that brings about the most information towards a set of unknown constants, which are parameterization dependent but not design dependent.

From (3.4.3), it follows that

$$|I(\xi; \alpha)| = |A|^{-2}\, |I(\xi; \beta)| \,. \tag{3.4.6}$$

This might project the illusion that transformations with different matrix A will result in different D optimal designs for various parameterization $\alpha = A\beta$ since the matrix A participates in the optimization. Then by (3.4.5), two different designs, each being D optimal only for its own parameterization (not for the other), should achieve minimax of the scaled prediction variance equivalently defined under the other parameterization. By the equivalence theorem, each of the two designs which is D optimal only for its initial parameterization should also be D optimal for the other parameterization. It seems as if a contradiction had occurred.

The key reason causing the above misunderstanding can be explained as follows. It is true that a design can be found by maximizing (3.4.6), the determinant of the information matrix pertaining to some improper parameterization α. Since $|I(\xi; \beta)|$ in (3.4.6) stays constant under transformations of β into various α, the solution of such a design will vary according to the matrix A that provides a specific formulation of α from the initial parameter β. Since α=Aβ and A depends on the designs, for a given transformation A, the design obtained by maximizing (3.4.6) has the largest information among all competing designs only in the sense that the information offered by each competing design, which identifies its competitiveness, is the information pertaining to a specific set of unknown parameter values determined by this design. So the competition is really judged by levels of information, each for a different set of parameters, due to various candidate designs. The winner of such competition is not a D optimal design in the proper sense because the competing informations are not for any common set of parameter values, whereas a true D optimal design is found by comparing informations all for a set of common parameter values among various designs. When a design is not D optimal, there will not be the equivalence property. In conclusion, a contradiction really has never happened.

A caution drawn from the above is that despite the invariance of a D optimal design to parameterizations, one should not attempt to obtain a D optimal design by maximizing the information for a transformed parameter when the transformation is design dependent.

§3.5 Equivalence for Two-Level D Optimal Design, One-Variable Logistic Model

Despite the moderate amount written about the one-variable logistic design problem, design equivalence has been discussed very little. Equivalence of D optimal designs for a one-variable logistic model is mentioned briefly for the situation of a limited design space in White (1973).

The determinant of the information matrix for the one-variable logistic model, due to a design with N observations, is given by

$$|I(\beta)| = \frac{1}{\beta_1^2}[\sum_{i=1}^{N} \sigma_i^2 \sum_{i=1}^{N} \sigma_i^2 L_i^2 - (\sum_{i=1}^{N} \sigma_i^2 L_i)^2], \tag{3.5.1}$$

where $\sigma_i^2 = P_i(1-P_i)$ and $L_i = \text{logit}(P_i)$, where P_i is the probability of response for the i^{th} observation. Maximizing $|I(\beta)|$ in (3.5.1) under no restriction as to the number of design points yields a D optimal design with two equally weighted levels: $ED_{17.6}$ and $ED_{82.4}$. The D criterion restricted with two design levels, $|I(\beta)| = \frac{1}{\beta_1^2} \sigma_1^2 \sigma_2^2 (L_2 - L_1)^2$, would of course lead to the same optimal solution.

For an equal sample size design placed at the symmetric logits $-L_0$ and L_0, where $L_0 > 0$, the scaled prediction variance at the point x as given in (3.3.1) can be expressed as a function of the logit L at the point x, where $L = \beta_0 + \beta_1 x$, in the following form:

$$V(L_0; L) = \frac{(1+e^{-L_0})(1+e^{L_0})}{(1+e^{-L})(1+e^{L})}[\frac{L^2}{L_0^2} + 1]. \tag{3.5.2}$$

Define z to be a centered and scaled design variable given by the transformation

$$z = \frac{L}{L_0} = \frac{\beta_0}{L_0} + \frac{\beta_1}{L_0}x. \tag{3.5.3}$$

The transformation is done in such a way that z equals 0 at ED_{50} and ± 1 at the design points. The accordingly transformed vector of parameters is given by $\alpha = A\beta$, where

$$A = \begin{bmatrix} 1 & -\frac{\beta_0}{\beta_1} \\ 0 & \frac{L_0}{\beta_1} \end{bmatrix}, \tag{3.5.4}$$

so that the logit can be equivalently written as $L = x'\beta = z'\alpha$. However, the matrix A involves design specification L_0 which participates in the optimization. The discussion on parameterization in §3.4 indicates that the parameter α, resulted from centering and scaling of the natural design variable, does not constitute a proper parameterization as far as design optimization is concerned and will not lend itself to a meaningful D optimal design. The sole purpose of the transformation is to benefit from the centered and scaled design variable z in the evaluation of a design generated for the initial parameter β.

The scaled prediction variance $V(L_0; L)$ appearing in (3.5.2) can be reexpressed as a function of z in the form

$$W(L_0; z) = \frac{(1+e^{-L_0})(1+e^{L_0})}{(1+e^{-L_0 z})(1+e^{L_0 z})}[\,1 + z^2\,]\,. \tag{3.5.5}$$

For the one-variable logistic model, p = 2. Since $W(L_0; z) = p = 2$ at $z = \pm 1$ for any L_0, any symmetric two-level design with equal sample sizes would have a scaled prediction variance equal to p at the design points. Nevertheless, for a design to be D optimal, according to the equivalence theorem, it has to satisfy $\underset{z \in R}{\mathrm{Max}}\, W(L_0; z) = 2$, where R is the entire real line $(-\infty, \infty)$. The above implies that if $\underset{z \in R}{\mathrm{Max}}\, W(L_0; z)$ occurs at the design points, or $z = \pm 1$, for some $L_0 = L^*$, then the design identified by the logit L^* will be a D optimal design for the region R. Therefore, the condition

$$W(L^*; c) = \underset{z \in R}{\mathrm{Max}}\, W(L^*; z)\,, \tag{3.5.6}$$

where $c = \pm 1$, may offer another way to solve for a D optimal design.

To solve the D optimal logit L^* from condition (3.5.6), obtain the following derivative from (3.5.5):

$$\frac{\partial W(L_0; z)}{\partial z} = \frac{P_z(1-P_z)}{P_0(1-P_0)}[\,L_0(1 - 2P_z)(1 + z^2) + 2z\,]\,, \tag{3.5.7}$$

where $P_0 = \dfrac{1}{1+e^{-L_0}}$ and $P_z = \dfrac{1}{1+e^{-L_0 z}}$. Evaluating (3.5.7) at z = 1 and setting it to zero leads to

$$L_0(1 - 2P_0) + 1 = 0\,. \tag{3.5.8}$$

Numerical solution of (3.5.8) yields $L_0 = L^* = 1.5434056$, or $P_0 = P^* = 0.823959$, which agrees with the result from maximizing $|I(\beta)|$ in (3.5.1). The second order derivative is shown to be

$$\frac{\partial^2 W(L_0;z)}{\partial z^2} = \frac{P_z(1-P_z)}{P_0(1-P_0)}\left[L_0^2(1-2P_z)^2(1+z^2) - 2L_0^2 P_z(1-P_z)(1+z^2) + 4zL_0(1-2P_z) + 2\right] \tag{3.5.9}$$

and it reduces to $-4L^{*2}P^*(1-P^*)$ at $L_0 = L^*$ and $z = \pm 1$, which is negative and thereby confirms the maximization of $W(L^*; z)$ at $z = \pm 1$.

The above has shown that $L^* = 1.5434$, originally coming from Max $|I(\beta)|$, is also a solution of (3.5.6) and equivalently a solution of $\max_{z \in R} W(L^*; z) = p$ (condition (3)). This indicates that the two-level D optimal design simultaneously meets conditions (1) and (3) in the equivalence theorem. Condition (2) will not be illustrated to the fullest extent here since it involves comparison of all possible designs. Direct verification of condition (2) has been partially done. Since $W(L_0; z) = p$ at $z = \pm 1$ for any L_0 while $\max_{z \in R} W(L_0; z) = p$ only for $L_0 = L^*$, the D optimal design does minimize the maximum scaled prediction variance among all symmetric, equally weighted two-level designs.

In figure 3.5.1 on page 64, the scaled prediction variance $W(L_0; z)$ is plotted as a function of z for the designs with the upper logits at $L_0 = 1$, $L_0 = L^* = 1.5434$, and $L_0 = 2$. Obviously, all three designs have $W(L_0; z) = 2$ at $z = \pm 1$, while only the D optimal design has $\max_{z \in R} W(L_0; z) = 2$ (condition (3)) and maintains the smallest maximum $W(L_0; z)$ (condition (2)) among the three designs.

§3.6 Equivalence for Three-Level D Optimal Design,One-Variable Logistic Model

The determinant of the information matrix for a three-level design in the one-variable logistic case can be expressed as

$$|I(\beta)| = \frac{1}{\beta_1^2}\,[\sigma_1^2\,\sigma_2^2\,(L_1 - L_2)^2 + \sigma_1^2\,\sigma_3^2\,(L_1 - L_3)^2 + \sigma_2^2\,\sigma_3^2\,(L_2 - L_3)^2]\,, \qquad (3.6.1)$$

where L_i = logit(P_i) and $\sigma_i^2 = n_iP_i(1\text{-}P_i)$, i = 1, 2, 3. Maximizing $|I(\beta)|$ constrained on equal sample sizes yields an optimal design given by $ED_{13.6}$, ED_{50}, and $ED_{86.4}$. The three-level D optimal design is in fact the result of constrained optimization of the $|I(\beta)|$ in (3.5.1). Hence the three-level D optimal design is not truly D optimal among all possible designs, as implied in the equivalence theorem, but rather conditionally D optimal among all equally weighted three-level designs. As a result, the three-level D optimal design will not have any of the three properties in the equivalence theorem. It does not minimize $|I(\beta)|$ among all designs (losing property (1)) because unconstrained maximization of $|I(\beta)|$ yields the two-level design. Its maximum scaled prediction variance, as can be seen later, is greater than p = 2 (losing property (3)) and hence is not the smallest among all designs (losing property (2)) since the two-level D optimal design keeps its maximum scaled prediction variance as low as p = 2.

Although the three-level D optimal design is not truly D optimal and hence is deprived of all the properties of the equivalence theorem, its conditional D optimality still earns itself respective properties in equivalence in a restricted scale. The two properties in equivalence for the three-level D optimal design are as follows:

(1) the design achieves $\underset{\xi\in\Xi_0}{\text{Max}}\,|I(\beta)|$,

(2) the design achieves $\underset{\xi\in\Xi_0}{\text{Min}}\ \underset{x\in R}{\text{Max}}\,V(\xi;\,x)$,

where Ξ_0 is the set of all equal sample size three-level designs defined on the entire real line.

For an equal sample size three-level design with symmetric logit levels $-L_0$, 0, and L_0 ($L_0 > 0$), the scaled prediction variance $V(\xi; x)$ can be expressed as a function of the logit $L = \beta_0 + \beta_1 x$ in the form

$$V(L_0; L) = \frac{3}{2}\frac{(1+e^{-L_0})(1+e^{L_0})}{(1+e^{-L})(1+e^{L})}\left[\frac{L^2}{L_0^2} + \frac{8}{8+(1+e^{-L_0})(1+e^{L_0})}\right]. \tag{3.6.2}$$

Transforming the natural variable x into the centered and scaled variable z by equation 3.5.3, where z equals 0 at ED_{50} and ± 1 at the design points, leads to the scaled prediction variance expressed as a function of z:

$$W(L_0; z) = \frac{3}{2}\frac{(1+e^{-L_0})(1+e^{L_0})}{(1+e^{-L_0 z})(1+e^{L_0 z})}\left[z^2 + \frac{8}{8+(1+e^{-L_0})(1+e^{L_0})}\right]. \tag{3.6.3}$$

Again, one should be cautioned that the modified parameters accompanying the transformed variable z do not form a proper parameterization for the purpose of design optimization and therefore cannot support the regular definition of a D optimal design.

In figure 3.6.1 on page 64, the scaled prediction variance $W(L_0; z)$ is plotted for the three-level designs with the upper logit at $L_0 = 1.5$, 1.85, and 2.63. The design with $L_0 = 1.85$ is the three-level D optimal design and the design with $L_0 = 2.63$ has its scaled prediction variance equal to $p = 2$ at the design points. Clearly, the three-level D optimal design has the lowest maximum $W(L_0; z)$, which equals 2.226 and occurs at the lower and upper design points. The design with $L_0 = 2.63$, though having $W(L_0; z) = 2$ at all design points, has higher maximum $W(L_0; z)$ than the three-level D optimal design.

It turns out that the three-level D optimal design can also be solved by the condition that $\underset{z \in R}{\text{Max}}\, W(L_0; z)$ occurs at the design points. To find the three-level D optimal design by requiring $\underset{z \in R}{\text{Max}}\, W(L_0; z) = W(L_0; 1)$, obtain the derivative

$$\frac{\partial W(L_0; z)}{\partial z} = \frac{3}{2} \frac{P_z(1-P_z)}{P_0(1-P_0)} \left[L_0(1 - 2P_z)(z^2 + \frac{8}{8 + \frac{1}{P_0(1-P_0)}}\right] + 2z \,], \qquad (3.6.4)$$

where $P_0 = \frac{1}{1+e^{-L_0}}$ and $P_z = \frac{1}{1+e^{-L_0 z}}$. Evaluating (3.6.4) at $z = 1$ and setting it to zero leads to

$$L_0(1 - 2P_0)(2 - \frac{1}{8P_0(1-P_0)+1}) + 1 = 0 \, . \qquad (3.6.5)$$

Numerical solution of (3.6.5) yields $L_0 = L^* = 1.850123$, or $P_0 = P^* = 0.864142$, which agrees with the result from maximizing $|I(\beta)|$ in (3.6.1). It can be verified that a negative second order derivative $\frac{\partial^2 W(L_0; z)}{\partial z^2}$ is concurrent with the maximum $W(L_0; z)$ at $z = \pm 1$ for the optimal design with $L_0 = L^*$.

§3.7 Equivalence for Parallel-Line D Optimal Design,Two-Variable Logistic Model

The parallel-line D optimal design is developed for the two-variable, no-interaction logistic model. If the region R is taken to be the entire two-dimensional design space, there would not exist a D optimal design since $|I(\beta)|$ could be made infinitely large by stretching the design points to the outer extremes along the direction of a constant probability. For the two-variable logistic model, one can only talk about the D optimal design for some constrained region R. This is equivalent to the usual practical restriction in the linear model that the variables are confined to the scaled interval [-1, 1]. Design levels allowed to wander beyond this interval will result in larger values of |X'X|.

The situation is shown in figure 3.7.1. Consider an elongated region R that spans unbounded from one extreme ED_0 to the other ED_{100}, but within a limited distance of $\pm\sqrt{D_1^2 + D_2^2}$ in the direction of constant probabilities, where the lengths D_1 and D_2 are indicated in figure 2.1.1. Consider a subset of designs placed on this region, each having four design points located at the intersections of two constant probability lines (P_1 and P_2 specific to the design) and the two common lines of region boundaries. Equal sample sizes are assumed for the two points on the same probability level. All possible combinations of P_1 and P_2 along with various sample weightings constitute the complete subset of parallel-line designs.

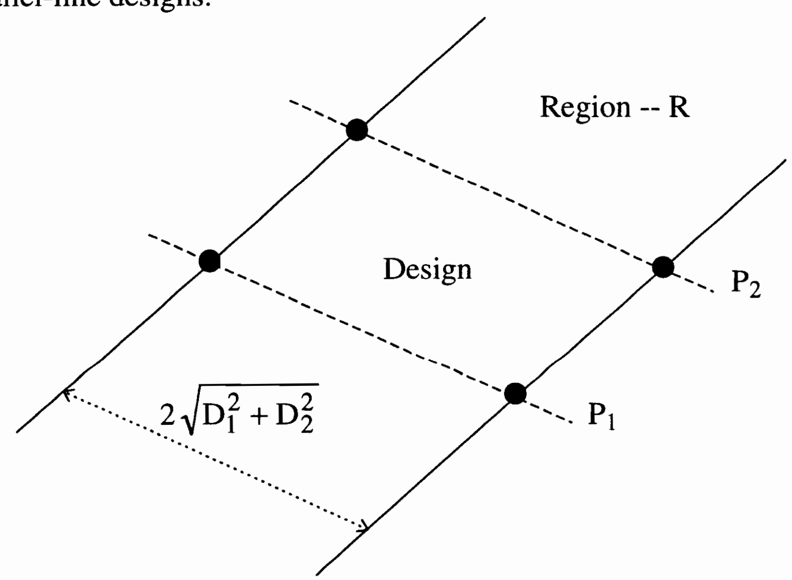


Figure 3.7.1 Region and design layout for parallel-line design

In fact, the parallel-line D optimal design, given by $ED_{22.7}$ and $ED_{77.3}$ with equal sample sizes, is obtained by finding the optimal P_1 and P_2 along with the best sample allocation that maximize $|I(\beta)|$ among the subset of all parallel-line designs on the region R. Up to this point the parallel-line D optimal design is known to be the best only in the set of parallel-line designs. The following work verifies that property (3) of the equivalence theorem holds for the parallel-line D optimal design and thereby confirms its D optimality among all possible designs in the region R.

The centered and scaled variables z_1 and z_2 can be created through the relationship

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_{10} \\ x_{20} \end{bmatrix} + \begin{bmatrix} D_1 & d_1 \\ D_2 & d_2 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \tag{3.7.1}$$

where D_i and d_i, i = 1, 2, are quantities characterizing the design and region geometry as displayed in figure 2.1.1, and x_{10} and x_{20} identify the design center also as shown in figure 2.1.1. Once again, since d_1 and d_2 involve the design probability levels that need to be optimized, the set of the transformed parameters accompanying variables z_1 and z_2 will not represent a proper parameterization that can lend itself to the normal definition of a D optimal design. Using variables z_1 and z_2, the design center is represented by (0, 0), and the design points are located at levels ±1. Therefore, the design identified by z_1 and z_2 is a square with the four corners being the design points.

The quantity $x'(X'\Sigma X)^{-1}x$ contained in the expression for the scaled prediction variance as in (3.3.1) can be expressed as

$$x'(X'\Sigma X)^{-1}x = \frac{1}{8}\left[\frac{(z_2-1)^2}{\sigma_1^2} + \frac{(z_2+1)^2}{\sigma_2^2}\right] + \frac{z_1^2}{2(\sigma_1^2+\sigma_2^2)} \tag{3.7.2}$$

where $\sigma_i^2 = n_iP_i(1-P_i)$, i = 1, 2. For a design with symmetric ED's at $1-P_0$ and P_0 as well as equal sample sizes, $\sigma_1^2 = \sigma_2^2 = \frac{N}{4}P_0(1-P_0)$. Substituting (3.7.2) into (3.3.1), the scaled prediction variance for such a design is given by

$$V(\xi; x) = \frac{P(1-P)}{P_0(1-P_0)}[\,1 + z_1^2 + z_2^2\,] \tag{3.7.3}$$

where $P = \dfrac{1}{1+e^{-x'\beta}}$. The logit L, where $L = \text{logit}(P) = x'\beta$, can be expressed in terms of z_1 and z_2 via (3.7.1) and turns out to be a function of only z_2:

$$L = \frac{L_2 - L_1}{2} z_2, \tag{3.7.4}$$

where L_1 and L_2 are the logit levels of the parallel-line design. For a symmetric design indicated by P_0,

$$L = L_0 z_2 \tag{3.7.5}$$

where $L_0 = \text{logit}(P_0)$. With (3.7.5), the scaled prediction variance of a symmetric design indicated by L_0 can be written as a function of z in the form

$$V(\xi; x) = W(L_0; z) = \frac{(1+e^{-L_0})(1+e^{L_0})}{(1+e^{-L_0 z_2})(1+e^{L_0 z_2})}[\,1 + z_1^2 + z_2^2\,]\,. \tag{3.7.6}$$

Clearly, for any symmetric parallel-line design with equal sample sizes, the scaled prediction variance would be equal to the number of parameters p = 3 at the design points, where $z_i = \pm 1$, i = 1, 2. However, only for the optimal design is the maximum scaled prediction variance over the entire region equal to 3. This can be seen by inspecting the contours of $W(L_0; z)$ in figures 3.7.2 through 3.7.4. In the (z_1, z_2) space, the initially defined region R is now the area constrained by $|z_1| \leq 1$. Figures 3.7.2 and 3.7.3 on page 65 show the contours of $W(L_0; z)$ for the designs with the below-optimal logit $L_0 = 1$ and the above-optimal logit $L_0 = 1.5$ respectively. Within the region R, areas where $W(L_0; z)$ exceeds 3 are found in both situations. Such areas lie outside of the design square for the design with lower-than-optimal L_0 and inside for the design with higher-than-optimal L_0. The contours of $W(L_0; z)$ for the optimal design, where L_0 =

1.2229, appears in figure 3.7.4 on page 66. The contour plot for the optimal design shows that except at the design points, where $W(L_0; z) = 3$, the scaled prediction variance is smaller than 3 throughout the region.

The contour plot for the optimal design has confirmed that the parallel-line D optimal design possesses equivalence property (3): $\underset{x \in R}{\text{Max}}\, V(\xi; x) = 3$. This design will then be entitled to properties (1) and (2). Meanwhile, property (2) has been partially verified as it has been shown that the optimal parallel-line design has the smallest maximum scaled prediction variance among all symmetric, equal sample size parallel-line designs.

Equation (3.7.6) indicates that for a fixed z_2, $W(L_0; z)$ is maximized at $z_1 = \pm 1$, or the region edges. This implies that the maximum $W(L_0; z)$ within the region R is always to be achieved somewhere along the region boundaries. The scaled prediction variance on the region boundaries is a function of only z_2, denoted $w(L_0; z)$, and is given by

$$w(L_0; z) = \frac{(1+e^{-L_0})(1+e^{L_0})}{(1+e^{-L_0 z})(1+e^{L_0 z})}[\,2 + z^2\,] \tag{3.7.7}$$

where z is a scalar variable equivalent to z_2. The functional form of $w(L_0; z)$ largely resembles that of $W(L_0; z)$ for a symmetric two-level design in the case of a one-variable model. The function $w(L_0; z)$ is plotted in figure 3.7.5 on page 66 for the same designs inspected earlier using the contour plots. The plot of $w(L_0; z)$ in figure 3.7.5, with a pattern similar to the one-variable plot of $W(L_0; z)$, depicts the behavior of the scaled prediction variance along the region boundaries for both the optimal and other than optimal designs. All three curves cross at $z = \pm 1$, giving $w(L_0; z) = 3$, but only the one for the optimal design is able to refrain itself from exceeding 3.

Similar to the cases in the previous sections, the function $w(L_0; z)$ offers an alternative in solving for the optimal design. Again, equivalence property (3) suggests that the design for which $\underset{z \in R}{\text{Max}}\, w(L_0; z)$ is achieved at the design points $z = \pm 1$ is the optimal solution. Consider the derivative

$$\frac{\partial w(L_0;z)}{\partial z} = \frac{P_z(1-P_z)}{P_0(1-P_0)} [\, L_0(1 - 2P_z)(2 + z^2) + 2z \,] , \tag{3.7.8}$$

where $P_0 = \dfrac{1}{1+e^{-L_0}}$ and $P_z = \dfrac{1}{1+e^{-L_0 z}}$. Evaluating (3.7.8) at $z = 1$ and setting it to zero leads to

$$3L_0(1 - 2P_0) + 2 = 0 . \tag{3.7.9}$$

Numerical solution of (3.7.9) yields $L_0 = L^* = 1.22291$, or $P_0 = P^* = 0.772575$, which agrees with the result from maximizing $|I(\beta)|$ in §2.1. The second derivative $\dfrac{\partial^2 w(L_0;z)}{\partial z^2}$ is found to be negative at $z = \pm 1$ given $L_0 = L^*$, confirming that $w(L_0; z)$ has indeed been maximized at $z = \pm 1$ for $L_0 = L^*$.

In summary, we have verified equivalence property (3), and partially property (2), for the parallel-line D optimal design and thereby confirmed all three equivalence properties associated with the design. The parallel-line D optimal design was originally intended to be the best only among the set of all parallel-line designs over the selected region. Its global D optimality among all possible designs in the region remained unknown until the discovery of property (3) for the design, i.e. $\underset{x \in R}{\text{Max}}\, V(\xi; x) = p$. Since property (3) has been verified, it can be concluded that the parallel-line D optimal design is not only the best in the set of parallel-line designs but also truly D optimal among all possible designs that can be constructed in the selected region. The set of all possible design measures includes designs of any other patterns with any number of design points.

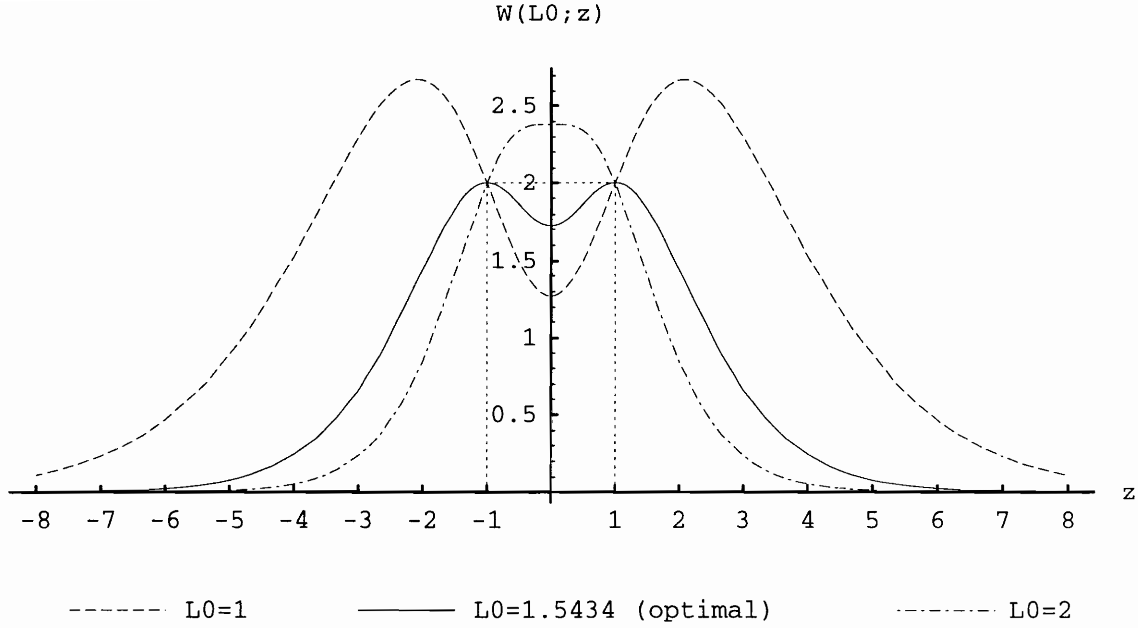


Figure 3.5.1 Scaled prediction variance for two-level designs, one-variable model

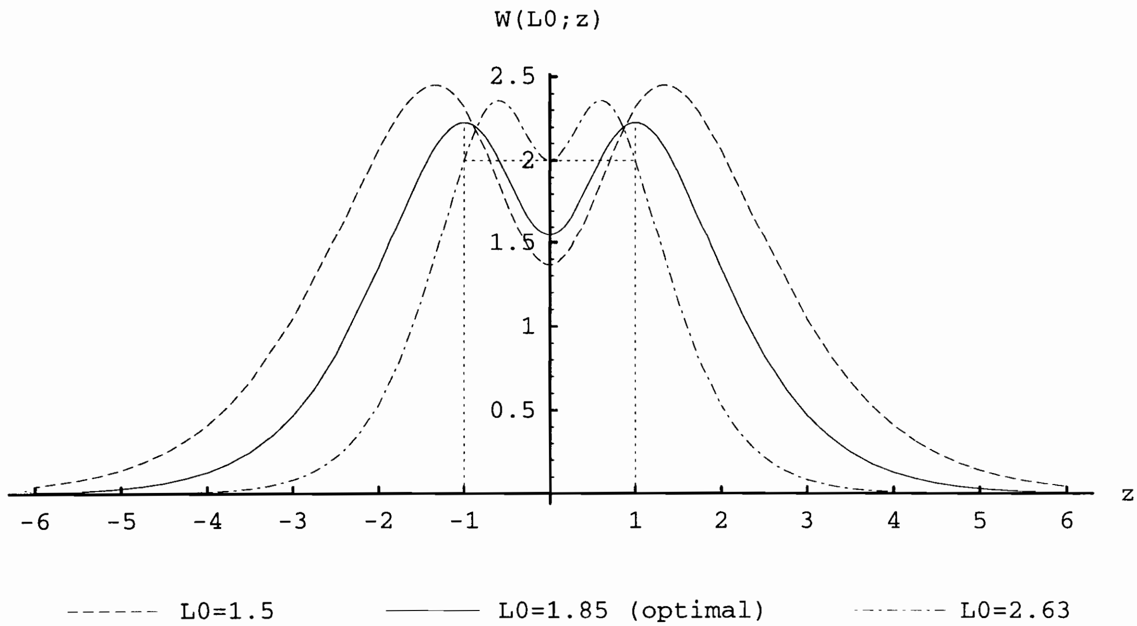


Figure 3.6.1 Scaled prediction variance for three-level designs, one-variable model

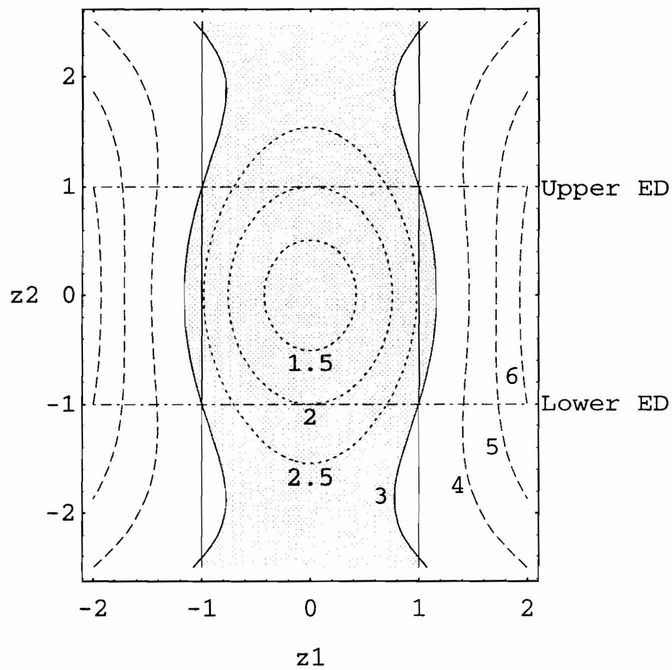


Figure 3.7.2 Scaled prediction variance for parallel-line design with L0 = 1

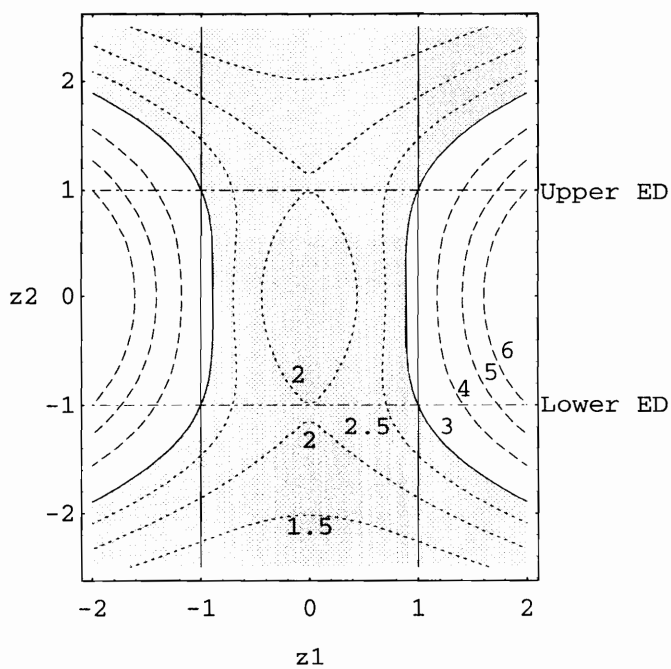


Figure 3.7.3 Scaled prediction variance for parallel-line design with L0 = 1.5

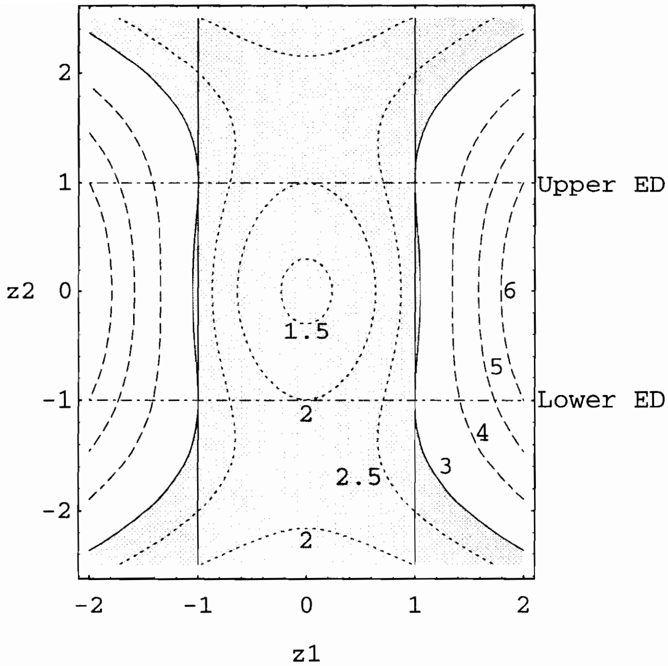


Figure 3.7.4 Scaled prediction variance for D optimal parallel-line design (L0 = 1.2229)

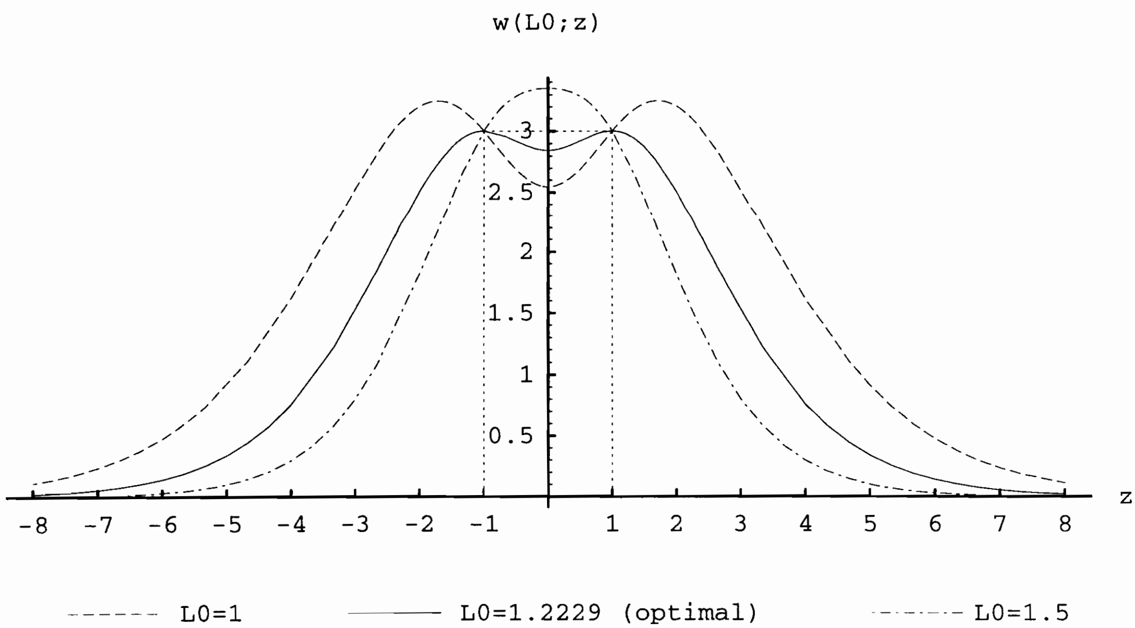


Figure 3.7.5 Scaled prediction variance for parallel-line designs along region edges

CHAPTER 4

DESIGN EFFICIENCY IN THE PRESENCE OF INTERACTION

§4.1 Overview

Designs assuming the no-interaction model are built on straight line ED's and have relatively simple structures while those assuming the interaction model have more complicated structures due to the hyperbolic ED curves. In addition, design construction is done without using the interaction coefficient for designs based on straight line ED's but does require knowledge of the interaction parameter for designs involving hyperbola ED's. Therefore design implementation seems easier and more practical for designs on straight line ED's than for those on hyperbola ED's.

However, since an interaction effect is of potential concern in many experimental studies, a simple design on straight line ED's can often be intended for situations with interaction. Such consideration is reflected in the design criterion adopted by Brunden *et al.* which comes from fitting an interaction model while assuming $\beta_{12} = 0$ in the true relationship. How effective such an approach is in bringing a good design to an interaction situation can be evaluated through relative efficiencies between alternative designs. In general, it is of interest to investigate all designs based on straight line ED's in the presence of interaction. This will be done by finding their efficiencies relative to the hyperbola-based design. The D efficiency in this context can be expressed as

$$\text{D-EFF} = \left(\frac{|I(\beta)| \text{ from design A, given } \beta_{12} \neq 0}{|I(\beta)| \text{ from design B, given } \beta_{12} \neq 0} \right)^{\frac{1}{p}} \tag{4.1.1}$$

where p is the number of parameters, or p = 4, B currently represents the hyperbola-based design which is the common basis of comparison, and A is a design under investigation such as the ray, factorial, or parallel-line designs.

Brunden *et al.* (1988) computed a different type of efficiency for their ray and factorial designs. Their efficiency is defined as

$$\frac{|I(\beta)| \text{ from design A, given } \beta_{12} \neq 0}{|I(\beta)| \text{ from design A, given } \beta_{12} = 0}. \tag{4.1.2}$$

This efficiency expression does not compare two designs but rather measures the change in $|I(\beta)|$ as β_{12} varies from zero for the same design. A large value of this efficiency says nothing as to whether the design is a successful candidate in the respective situation.

To evaluate the efficiency defined in (4.1.1), let X* be the model matrix from fitting an interaction model due to a simple design. Let P_i^* denote the actual response probability given $\beta_{12} \neq 0$ at the ith point of the simple design. It is shown that the efficiency can be expressed as

$$\text{D-EFF} = \frac{1}{4} \cdot \frac{\left(\beta_{12}^4 |X^{*\prime} X^*|\right)^{\frac{1}{4}}}{L} \cdot \frac{\left(\prod_{i=1}^{4} P_i^*(1-P_i^*)\right)^{\frac{1}{4}}}{P(1-P)} \tag{4.1.3}$$

where P is the upper optimal probability for the hyperbola-based design, or P = 0.823959, and L = logit(P).

§4.2 Efficiency of Ray Design

For a ray design, it is shown that

$$\prod_{i=1}^{4} P_i^*(1-P_i^*) = \left(\prod_{s=-1,1} \prod_{j=1}^{2} \left(1+e^{s\left(L_j+\frac{gk(L_j-\beta_0)^2}{(1+gk)^2\Delta_\beta}\right)}\right)\left(1+e^{s\left(L_j+\frac{(k/g)(L_j-\beta_0)^2}{(1+(k/g))^2\Delta_\beta}\right)}\right) \right)^{-1} \qquad (4.2.1)$$

where $k = \frac{\beta_1}{\beta_2}$, g is a function of k given in (2.3.3), L_j is an optimal logit for the ray design, and Δ_β is the difference in logit between the origin and the hyperbola enter, which is given by

$$\Delta_\beta = \beta_0 - \beta_0^* = \frac{\beta_1\beta_2}{\beta_{12}}. \qquad (4.2.2)$$

The D efficiency of the ray design relative to the hyperbola-based design is given by

$$\text{D-EFF}=F(k, \beta_0, \Delta_\beta)=\frac{f(k)^{\frac{1}{4}}(L_2-L_1)\sqrt{(L_1-\beta_0)(L_2-\beta_0)}}{4L|\Delta_\beta|}\,\frac{\left(\prod_{i=1}^{4} P_i^*(1-P_i^*)\right)^{\frac{1}{4}}}{P(1-P)}, \qquad (4.2.3)$$

where L_1 and L_2 are the optimal logits of the ray design, which vary with β_0, and f(k) is a function jointly given by (2.3.2) and (2.3.3).

Essentially, the efficiency is a function of the parameters k, β_0, and Δ_β. Due to the symmetry in k and the assumption $\beta_0 \le 0$, only cases of k<1 and $\beta_0 \le 0$ will be evaluated. Recall that when k=1, the ray design will cause $|I(\beta)|=0$ and hence zero efficiency. The efficiency is tabulated in table 4.2.1 and also plotted in figure 4.2.1.

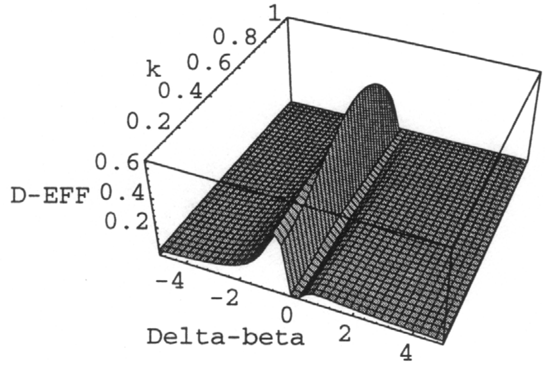
The assumptions in a ray design imply that $\Delta_\beta > 0$ represents synergism whereas $\Delta_\beta < 0$ corresponds to antagonism. Both the table and the graphs indicate that the ray design is not efficient at all with synergism as the efficiency gets close to zero most of the time. The efficiency is much better with antagonism. Nevertheless, as seen from the plots, unfailing performance only occurs over limited regions of the parameter space. For a given β_0, there is a maximum efficiency at a certain condition of k and Δ_β. These results appear in table 4.2.2.

Table 4.2.1 D efficiency of ray design

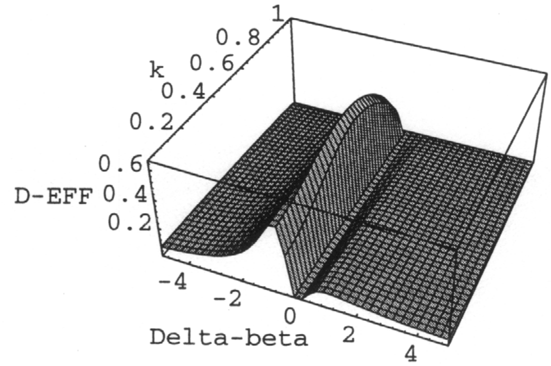
β_0	k	Δ_β -5	-2	-1	-0.5	0.5	1	2	5
0	0.1	.04	.11	.28	.43	.06	.07	.05	.03
	0.3	.03	.10	.27	.52	.05	.06	.04	.02
($ED_{71.1}$,	0.5	.03	.08	.24	.58	.04	.05	.04	.02
$ED_{97.8}$)	0.7	.02	.06	.18	.53	.03	.03	.03	.01
	0.9	.01	.03	.10	.32	.01	.02	.01	.01
-1	0.1	.07	.20	.44	.42	.08	.10	.09	.05
	0.3	.06	.19	.45	.53	.06	.09	.08	.04
($ED_{56.1}$,	0.5	.05	.16	.42	.64	.05	.07	.06	.03
$ED_{96.0}$)	0.7	.04	.12	.33	.60	.03	.05	.04	.02
	0.9	.02	.07	.19	.36	.02	.03	.02	.01
-2	0.1	.11	.31	.47	.27	.07	.12	.12	.07
	0.3	.10	.30	.52	.37	.05	.11	.10	.06
($ED_{44.5}$,	0.5	.08	.26	.53	.44	.04	.08	.08	.05
$ED_{94.1}$)	0.7	.06	.20	.45	.38	.03	.06	.06	.04
	0.9	.03	.11	.26	.22	.02	.03	.03	.02
-5	0.1	.23	.27	.09	.00	.00	.04	.12	.12
	0.3	.21	.30	.12	.01	.00	.04	.11	.11
($ED_{29.3}$,	0.5	.19	.31	.13	.00	.00	.03	.09	.09
$ED_{89.8}$)	0.7	.14	.27	.10	.00	.00	.02	.06	.06
	0.9	.08	.15	.05	.00	.00	.01	.04	.04

Table 4.2.2 Maximum D efficiencies of ray design

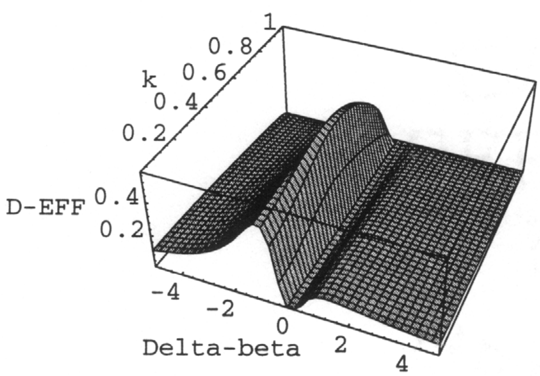
β_0	Max D-EFF	k	Δ_β
0	0.644441	0.595058	-0.382167
-1	0.650155	0.547976	-0.529677
-2	0.576373	0.505713	-0.773346
-5	0.323374	0.401569	-2.33088



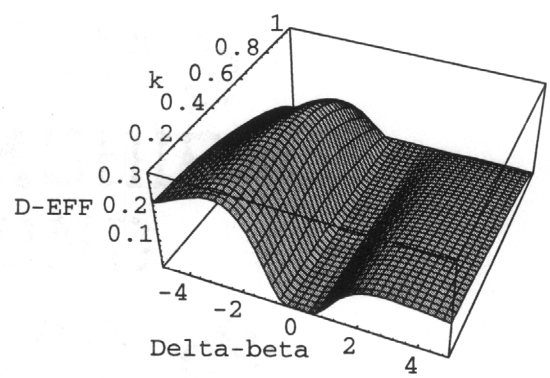
(1) βo = 0



(2) βo = −1



(3) βo = −2



(4) βo = −5

Figure 4.2.1 D efficiency of ray design

As seen in §2.6, the hyperbola center falls in the third quadrant under synergism but in the first under antagonism. The better efficiencies with antagonism then suggests that the simple design, currently the ray design, performs better when it lies closer to the hyperbola center and therefore closer to the hyperbola-based design as well. Intuitively, geographical closeness presumably should yield less inferior performance. As a matter of fact, this is a common feature revealed later in all other simple designs.

§4.3 Efficiency of Factorial Design

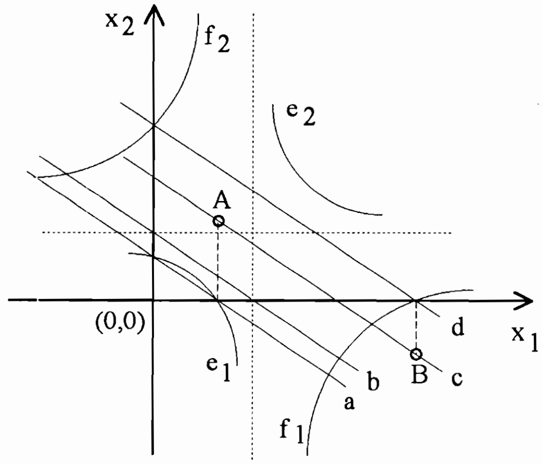
For the 2×2 factorial design, it is shown that

$$\prod_{i=1}^{4} P_i^*(1-P_i^*) = \left(\prod_{s_1=-1,1} \prod_{s_2=-1,1} (1+e^{\frac{s_1}{\Delta_\beta}(L_x+\frac{s_2L_0}{2})(\beta_0+L_x+\frac{s_2L_0}{2})}) \cdot (1+e^{s_1s_2L_0+\frac{s_1}{\Delta_\beta}(L_x-\frac{s_2L_0}{2})(\beta_0+L_x+\frac{s_2L_0}{2})}) \right)^{-1}, \quad (4.3.1)$$

where L_0 is the upper optimal logit for the factorial design, or $L_0 = 4.13068$, Δ_β is given by (4.2.2), and $L_x = \beta_1 x_{10}$ or $\beta_2 x_{20}$ where (x_{10}, x_{20}) is the design center as shown in figure 2.4.1. The efficiency of the factorial design is given by

$$\text{D-EFF} = F(\beta_0, \Delta_\beta, L_x) = \frac{L_0^2}{4L|\Delta_\beta|} \cdot \frac{\left(\prod_{i=1}^{4} P_i^*(1-P_i^*)\right)^{\frac{1}{4}}}{P(1-P)} . \quad (4.3.2)$$

The efficiency is a function of the parameters β_0 and Δ_β as well as the design location index L_x. The logit type quantity L_x is a scale free index representing the user selected position for the design center. For a given set of parameters, L_x uniquely determines the design location as illustrated in figure 4.3.1.



Legend	a:	ED for logit L_A,	assuming no interaction
	b:	ED for logit β_0^*,	assuming no interaction
	c:	ED_{50},	assuming no interaction
	d:	ED for logit L_B,	assuming no interaction
	e_1, e_2:	ED for logit L_A,	assuming interaction
	f_1, f_2:	ED for logit L_B,	assuming interaction
	A:	design center given by $L_x = L_A - \beta_0$ (assuming $L_x = \beta_1 x_{10}$)	
	B:	design center given by $L_x = L_B - \beta_0$ (assuming $L_x = \beta_1 x_{10}$)	

Figure 4.3.1 Correspondence between L_x and design center on ED_{50}

Due to the symmetries

(1) $F(\beta_0, \Delta\beta, L_x) = F(-\beta_0, -\Delta\beta, -L_x)$

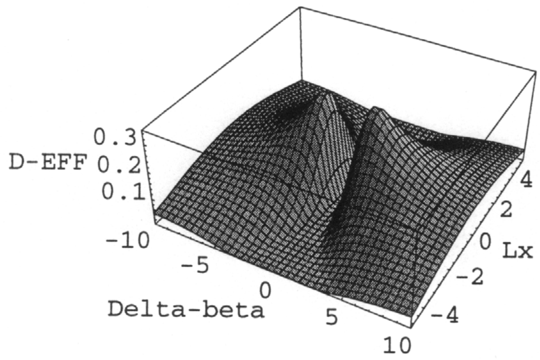
and (2) $F(\beta_0, \Delta\beta, L_x) = F(\beta_0, \Delta\beta, -\beta_0 - L_x)$,

only cases of $\beta_0 \le 0$ and $L_x \ge -\frac{\beta_0}{2}$ will be tabulated. The condition $L_x = -\frac{\beta_0}{2}$ corresponds to a design center at the midpoint of the ED_{50} line segment intersected by the x_1 and x_2 axes. The efficiency is listed in table 4.3.1 and also plotted in figure 4.3.2.

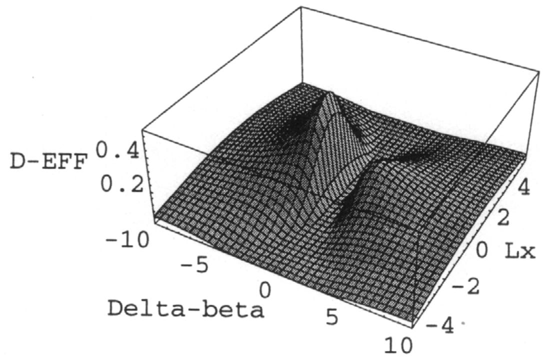
Table 4.3.1 D efficiency of factorial design

β_0	L_x	Δ_β -10	-5	-2	-1	1	2	5	10
0	0	.12	.22	.35	.19	.19	.35	.22	.12
	1	.11	.19	.25	.13	.13	.25	.19	.11
	2	.10	.14	.11	.03	.03	.11	.14	.10
	3	.09	.09	.03	.00	.00	.03	.09	.09
	4	.06	.05	.00	.00	.00	.00	.05	.06
-1	0.5	.13	.27	.55	.25	.10	.23	.18	.11
	1.5	.12	.24	.40	.20	.06	.16	.16	.10
	2.5	.11	.18	.18	.07	.01	.07	.12	.09
	3.5	.09	.11	.05	.00	.00	.02	.08	.08
	4.5	.07	.06	.00	.00	.00	.00	.04	.06
-2	1	.14	.33	.75	.26	.06	.16	.15	.10
	2	.14	.30	.57	.19	.03	.10	.14	.09
	3	.13	.22	.29	.11	.00	.04	.10	.08
	4	.11	.14	.08	.00	.00	.01	.07	.07
	5	.08	.07	.00	.00	.00	.00	.04	.05
-5	2.5	.19	.46	.30	.03	.00	.04	.08	.07
	3.5	.19	.45	.27	.02	.00	.03	.08	.07
	4.5	.18	.40	.22	.01	.00	.01	.06	.07
	5.5	.15	.30	.17	.00	.00	.00	.04	.06
	6.5	.12	.17	.03	.00	.00	.00	.03	.05

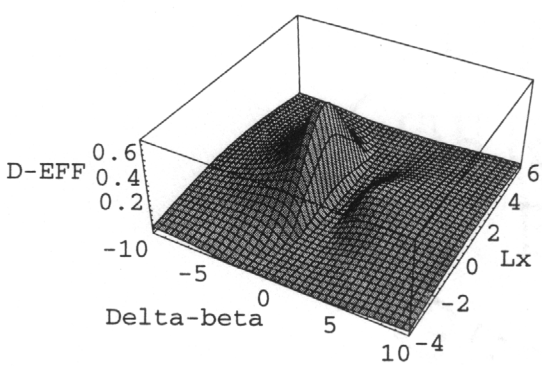
Since β_1 and β_2 are not restricted to have identical signs as in the ray design, the condition $\Delta_\beta>0$ now describes a broader situation than synergism. The correct interpretation is that $\Delta_\beta>0$ corresponds to the situation where the ED_{50} and the hyperbola center are separated by the origin. Furthermore, $\Delta_\beta>0$ implies that if $L_x = -\frac{\beta_0}{2}$, the center of the factorial design and the hyperbola center would fall in opposite quadrants, either



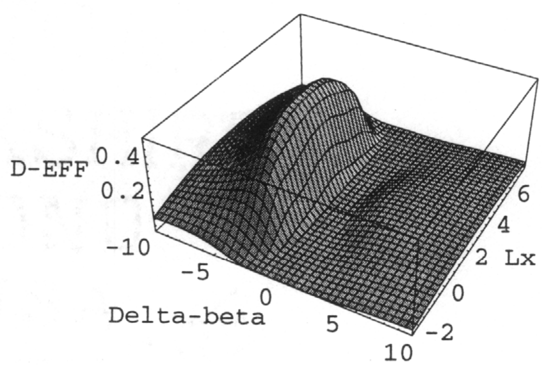
(1) $\beta o = 0$



(2) $\beta o = -1$



(3) $\beta o = -2$



(4) $\beta o = -5$

Figure 4.3.2 D efficiency of factorial design first/third or second/fourth. On the contrary, $\Delta_\beta<0$ implies that if $L_x = -\frac{\beta_0}{2}$, the design center and the hyperbola center would lie in the same quadrant.

The efficiency is worse when $\Delta_\beta>0$ than when $\Delta_\beta<0$ and especially so if the origin is relatively far from ED_{50}, or β_0 far from 0. Supported by the above paragraph, this suggests that in general, the factorial design seems less inferior when it lies closer to the hyperbola-based design but often deteriorates if away from it in the opposite direction.

For the factorial design, the best position of the design center is always given by the condition $L_x = -\frac{\beta_0}{2}$, which leads to the highest possible efficiency given the parameters. The maximum efficiencies given β_0 or or with no restriction are summarized in table 4.3.2.

Table 4.3.2 Maximum D efficiencies of factorial design

β_0	Max D-EFF	L_x	Δ_β
0	0.348610	0	-1.95936
-1	0.552402	0.5	-1.81418
-2	0.751089	1	-2.00687
-5	0.538607	2.5	-3.67626
	Global Max:		
-2.85649	0.816074	1.42825	-2.27609

§4.4 Efficiency of Parallel-Line Design

As seen in §2.5, when $k = -\frac{\beta_1}{\beta_2} = \pm 1$, a rectangular parallel-line design will result in a singular information matrix for fitting the interaction model. Similar to §2.5, the assumption here is to make the non-ED edges of the design horizontal if $|k| = \pm 1$ and leave the design rectangular otherwise. The D criterion of a parallel-line design is directly proportional to the squared length of an ED edge. The comparison of the parallel-line design against the hyperbola-based design will then have to be constrained on the two designs having equal areas.

For the above described parallel-line design, it is shown that

$$\prod_{i=1}^{4} P_i^*(1-P_i^*) = \left(\prod_{s_1=-1,1} \prod_{s_2=-1,1} \prod_{s_3=-1,1} \left(1+e^{s_1 s_3 L_0 + \frac{s_3}{\Delta_\beta}\left(L_x + \frac{s_1 L_0}{2}(1+f(k)) + s_2 \Delta_\beta \frac{L}{L_0}\right)\left(-\beta_0 - L_x + \frac{s_1 L_0}{2}(1-f(k)) - s_2 \Delta_\beta \frac{L}{L_0}\right)}\right) \right)^{-1} \tag{4.4.1}$$

where $k = -\dfrac{\beta_1}{\beta_2}$, L_0 is the upper optimal logit for the parallel-line design, or $L_0 = 1.22291$, Δ_β is as defined in (4.2.2), $L_x = \beta_1 x_{10}$ where x_{10} is the x_1 level at the design center as shown in figure 2.1.1, and the function f(k) is given by (2.5.2). The D efficiency relative to the hyperbola-based design is given by

$$\text{D-EFF} = F(k, \beta_0, \Delta_\beta, L_x) = \frac{\left(\prod_{i=1}^{4} P_i^*(1-P_i^*)\right)^{\frac{1}{4}}}{P(1-P)} \sqrt{|f(k)|}\ . \tag{4.4.2}$$

Due to the symmetries

(1) $F(k, \beta_0, \Delta_\beta, L_x) = F(k, -\beta_0, -\Delta_\beta, -L_x)$,

(2) $F(k, \beta_0, \Delta_\beta, L_x) = F(-k, \beta_0, \Delta_\beta, L_x)$,

and (3) for $k \neq \pm 1$, $F(k, \beta_0, \Delta_\beta, L_x) = F(\frac{1}{k}, \beta_0, \Delta_\beta, -\beta_0 - L_x)$, the efficiencies will be investigated for the restricted parameter space $|k| \in (0,1] \cap \beta_0 \in (-\infty, 0]$. The user selected design location is reflected by the scale free index L_x, which has been discussed with more details in §4.3. Considering the results in §4.3, L_x will be examined at values symmetric about $-\frac{\beta_0}{2}$. The efficiency is listed in table 4.4.1. For |k| = 0.5, the efficiency is plotted in figure 4.4.1. For other tabulated values of |k|, the graphs would look quite similar. Situations producing maximum efficiencies are summarized in table 4.4.2.

Table 4.4.1 D efficiency of parallel-line design

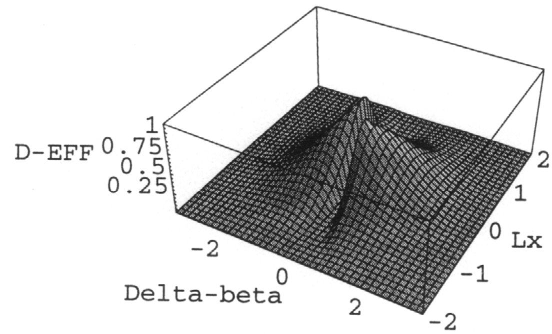
$\|k\|$	β_0	L_x	Δ_β -3	-2	-1	-0.5	0.5	1	2	3
0.2	0	-2	.01	.02	.01	.00	.00	.02	.02	.01
		-1	.04	.12	.21	.12	.23	.20	.09	.03
		0	.05	.21	.58	.76	.76	.58	.21	.05
		1	.03	.09	.20	.23	.12	.21	.12	.04
		2	.01	.02	.02	.00	.00	.01	.02	.01
	-1	-1.5	.01	.02	.01	.00	.00	.02	.01	.00
		-0.5	.04	.14	.27	.24	.23	.18	.09	.03
		0.5	.06	.25	.75	1.03	.41	.49	.22	.06
		1.5	.03	.11	.26	.24	.05	.18	.13	.04
		2.5	.01	.02	.02	.00	.00	.00	.02	.01
	-2	-1	.01	.03	.02	.01	.01	.02	.02	.01
		0	.05	.18	.35	.37	.26	.20	.09	.04
		1	.07	.34	.99	.63	.14	.40	.26	.07
		2	.04	.15	.35	.14	.02	.13	.15	.05
		3	.01	.02	.02	.00	.00	.01	.02	.01
	-5	0.5	.03	.04	.04	.03	.03	.04	.03	.02
		1.5	.18	.39	.12	.00	.00	.09	.20	.08
		2.5	.35	1.02	.05	.00	.00	.03	.33	.22
		3.5	.17	.38	.03	.00	.00	.01	.15	.16
		4.5	.03	.04	.02	.00	.00	.01	.03	.02

Table 4.4.1 (continued)

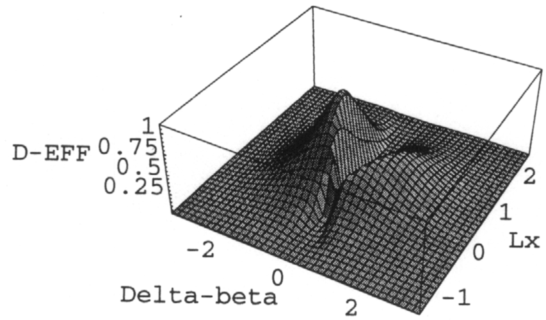
|k|	β_0	L_x	Δ_β -3	-2	-1	-0.5	0.5	1	2	3
0.5	0	-2	.01	.01	.01	.00	.00	.01	.01	.01
		-1	.03	.11	.19	.15	.22	.18	.08	.03
		0	.05	.20	.60	.79	.79	.60	.20	.05
		1	.03	.08	.18	.22	.15	.19	.11	.03
		2	.01	.01	.01	.00	.00	.01	.01	.01
	-1	-1.5	.01	.02	.01	.00	.00	.01	.01	.01
		-0.5	.03	.12	.24	.26	.20	.17	.08	.03
		0.5	.05	.23	.80	1.08	.37	.49	.21	.05
		1.5	.03	.10	.24	.26	.08	.17	.11	.03
		2.5	.01	.02	.02	.00	.00	.01	.01	.01
	-2	-1	.01	.02	.02	.01	.01	.02	.01	.01
		0	.04	.16	.31	.34	.21	.18	.09	.03
		1	.06	.32	1.06	.51	.11	.36	.25	.06
		2	.04	.13	.31	.19	.03	.14	.13	.04
		3	.01	.02	.02	.00	.00	.01	.02	.01
	-5	0.5	.03	.03	.03	.02	.01	.03	.03	.01
		1.5	.16	.34	.09	.00	.00	.05	.18	.08
		2.5	.32	1.06	.04	.00	.00	.02	.29	.21
		3.5	.15	.34	.04	.00	.00	.01	.15	.12
		4.5	.02	.03	.02	.00	.00	.01	.03	.02
0.8	0	-2	.01	.01	.01	.00	.00	.01	.01	.00
		-1	.02	.06	.12	.12	.14	.12	.06	.02
		0	.03	.13	.43	.56	.56	.43	.13	.03
		1	.02	.06	.12	.14	.12	.12	.06	.02
		2	.00	.01	.01	.00	.00	.01	.01	.01
	-1	-1.5	.01	.01	.01	.00	.00	.01	.01	.00
		-0.5	.02	.07	.15	.18	.11	.11	.06	.02
		0.5	.03	.15	.58	.75	.23	.34	.14	.03
		1.5	.02	.07	.15	.18	.08	.11	.06	.02
		2.5	.01	.01	.01	.00	.00	.01	.01	.01

Table 4.4.1 (continued)

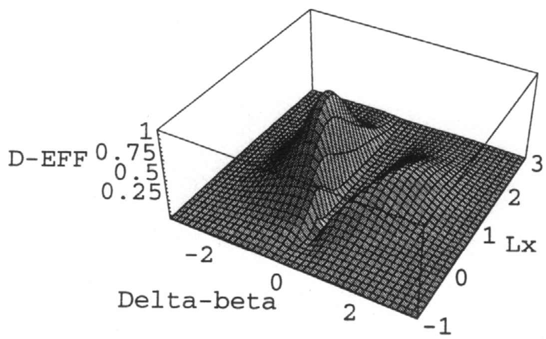
|k|	β_0	L_x	Δ_β -3	-2	-1	-0.5	0.5	1	2	3
0.8	-2	-1	.01	.01	.01	.01	.00	.01	.01	.01
		0	.03	.10	.20	.21	.09	.11	.06	.02
		1	.04	.22	.77	.30	.07	.23	.17	.04
		2	.03	.09	.20	.17	.04	.10	.07	.02
		3	.01	.01	.01	.00	.00	.01	.01	.01
	-5	0.5	.02	.02	.02	.00	.00	.02	.02	.01
		1.5	.10	.22	.04	.00	.00	.02	.12	.06
		2.5	.21	.76	.02	.00	.00	.01	.18	.14
		3.5	.10	.22	.03	.00	.00	.01	.11	.07
		4.5	.02	.02	.02	.00	.00	.01	.02	.01
1	0	-2	.01	.02	.02	.00	.00	.01	.02	.01
		-1	.03	.09	.20	.23	.11	.21	.13	.04
		0	.05	.21	.56	.73	.73	.56	.21	.05
		1	.04	.13	.21	.11	.23	.20	.09	.03
		2	.01	.02	.01	.00	.00	.02	.02	.01
	-1	-1.5	.01	.02	.02	.00	.00	.01	.02	.01
		-0.5	.03	.11	.16	.23	.05	.17	.13	.04
		0.5	.06	.25	.73	.99	.41	.48	.22	.06
		1.5	.04	.14	.27	.23	.23	.18	.08	.03
		2.5	.01	.02	.01	.00	.00	.02	.02	.01
	-2	-1	.01	.02	.02	.00	.00	.01	.02	.01
		0	.04	.15	.25	.13	.01	.13	.16	.05
		1	.08	.34	.95	.65	.14	.40	.25	.07
		2	.05	.18	.35	.37	.27	.20	.09	.04
		3	.01	.03	.02	.00	.01	.02	.02	.01
	-5	0.5	.03	.04	.02	.00	.00	.00	.03	.03
		1.5	.17	.38	.03	.00	.00	.01	.15	.16
		2.5	.35	.99	.06	.00	.00	.03	.33	.22
		3.5	.19	.39	.12	.00	.00	.10	.20	.08
		4.5	.03	.04	.04	.03	.03	.04	.03	.02



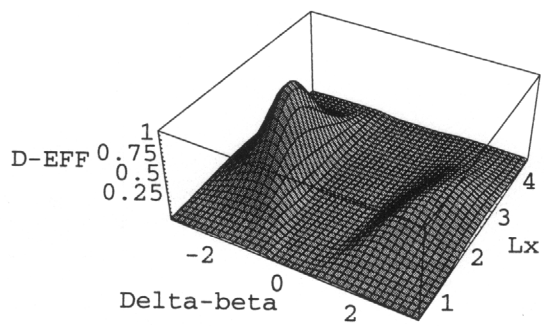
(1) $\beta o = 0$



(2) $\beta o = -1$



(3) $\beta o = -2$



(4) $\beta o = -5$

Figure 4.4.1 D efficiency of parallel-line design, $|k| = 0.5$

Table 4.4.2 Maximum D efficiencies of parallel-line design

$\lvert k \rvert$	β_0	Max D-EFF	L_x	Δ_β
0.2	0	1.01655	0.122949	-0.121582
	-1	1.03369	0.476906	-0.455597
	-2	1.02722	0.965984	-0.842871
	-5	1.02092	2.47883	-2.00833
		Max given $\lvert k \rvert$:		
	-0.610870	1.03542	0.305435	-0.305435
0.5	0	0.953837	0.259877	-0.236806
	-1	1.08343	0.515547	-0.539878
	-2	1.08083	0.985573	-0.900313
	-5	1.06786	2.48735	-2.03742
		Max given $\lvert k \rvert$:		
	-1.27061	1.08527	0.635305	-0.635305
0.8	0	0.569423	0.155526	-0.409865
	-1	0.777025	0.514442	-0.587630
	-2	0.783126	0.996431	-0.934492
	-5	0.770943	2.49563	-2.05558
		Max given $\lvert k \rvert$:		
	-1.54952	0.784779	0.774762	-0.774762
1	0	1*	0*	0*
	-1	0.995652	0.539806	-0.427234
	-2	0.990787	1.04092	-0.824623
	-5	0.986517	2.52350	-1.99947
		Max given $\lvert k \rvert$:		
	0*	1*	0*	0*
		Global Max:		
0.439529	-1.17012	1.09196	0.585060	-0.585060

* : Since Δ_β by definition is not allowed to be 0, the values with asterisks are meaningful only in the limiting sense, which means that for $\lvert k \rvert = 1$, the efficiency approaches 1 as β_0, L_x, and Δ_β all tend to zero in a manner such that $L_x = -\beta_0/2$ and $\Delta_\beta = \beta_0/2$.

As seen from table 4.4.1 and figure 4.4.1, the same feature found in the ray and factorial designs is once again present here. The efficiency is usually poor when $\Delta_\beta > 0$ but considerably better when $\Delta_\beta < 0$. The difference becomes more obvious as ED_{50} shifts away from the origin, or β_0 deviates from 0. Similar to the interpretation given in the previous sections, it is felt that the parallel-line design is likely to be less inefficient when located near the hyperbola-based design.

Table 4.4.2 shows that the best L_x, indicating the design location, is always close to $-\frac{\beta_0}{2}$ for given $|k|$ and β_0. For the less restricted maximum given $|k|$ only as well as the unrestricted global maximum, the best L_x is exactly at $-\frac{\beta_0}{2}$. Whenever $L_x = -\frac{\beta_0}{2}$ accompanies a maximum efficiency, the condition $\Delta_\beta = \frac{\beta_0}{2}$ is also present.

§4.5 Efficiency of Alternative Parallel-Line Design

The study is conducted in exactly the same fashion as seen for the regular parallel-line design. The only adjustment needed is to use the appropriate optimal logit for the alternative parallel-line design, which is 1.54341 instead of 1.22291 for the regular parallel-line design. Numerical results are given in tables 4.5.1 and 4.5.2. The efficiency is plotted in figure 4.5.1 for selected situations. No remarkable difference is found for the alternative parallel-line design and the same comments given in §4.4 apply here.

Based on §4.2 through §4.5, the simple type of designs including the ray, factorial, and parallel-line designs do not seem to offer reliable performance for the purpose of fitting an interaction model when the interaction is indeed present. The efficiencies are quite sensitive to the parameter structure and where needed, to the user selected design location as well. In all four designs, the efficiency can be rather low when $\Delta_\beta>0$. The interpretation is that a simple design is especially unsuccessful when it lies far from the hyperbola-based design. Best efficiencies roughly around 0.7 to 1.0 occur only in limited situations. As far as the efficiency is concerned, the hyperbola-based design still remains a better approach to handle an interaction situation.

Table 4.5.1 D efficiency of alternative parallel-line design

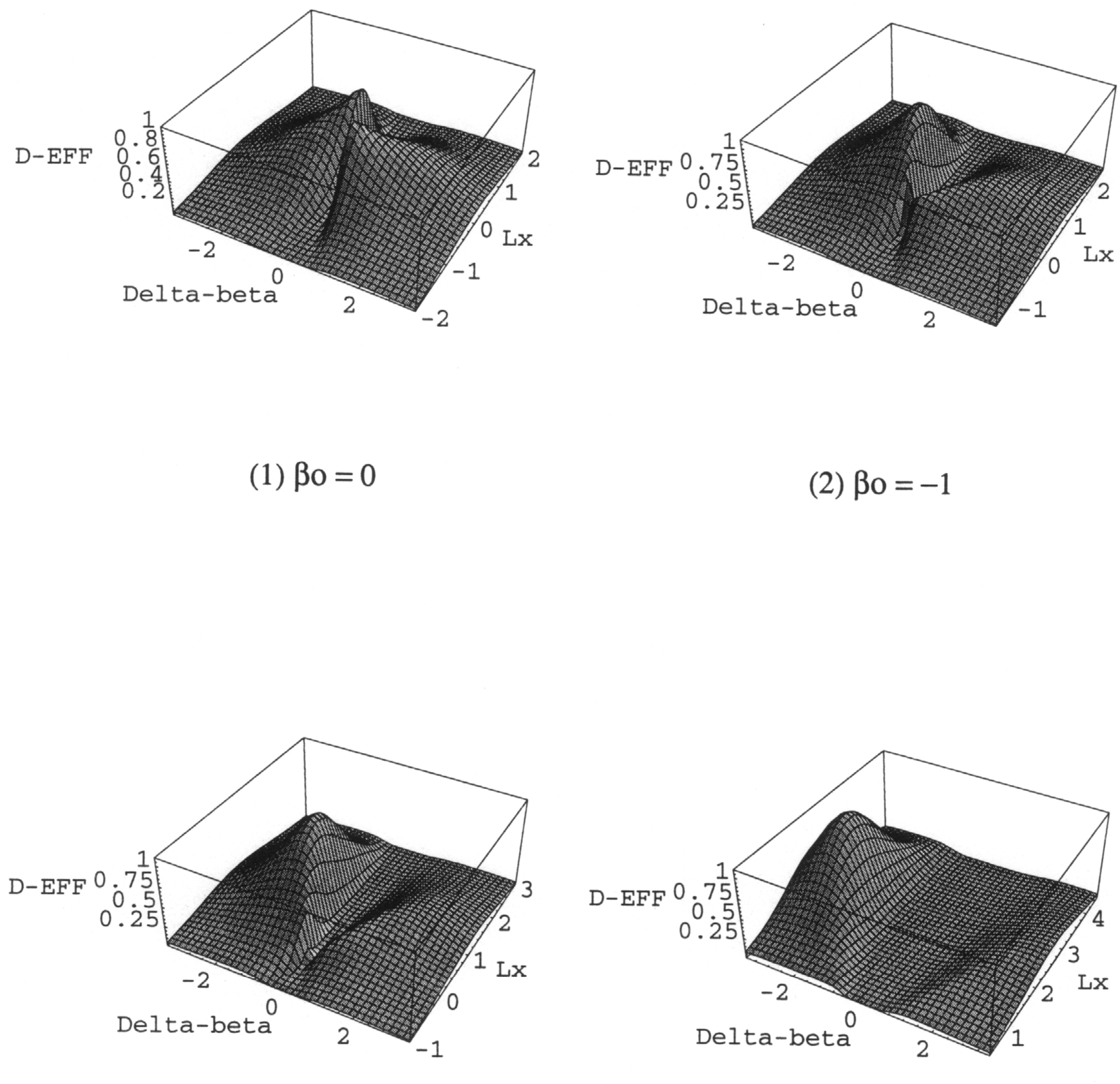
$\|k\|$	β_0	L_x	Δ_β -3	-2	-1	-0.5	0.5	1	2	3
0.2	0	-2	.04	.05	.02	.00	.00	.03	.05	.03
		-1	.17	.27	.22	.07	.34	.34	.20	.12
		0	.23	.42	.63	.70	.70	.63	.42	.23
		1	.12	.20	.34	.34	.07	.22	.27	.17
		2	.03	.05	.03	.00	.00	.02	.05	.04
	-1	-1.5	.05	.06	.03	.00	.00	.04	.04	.03
		-0.5	.19	.31	.35	.24	.34	.29	.19	.11
		0.5	.26	.51	.88	1.03	.25	.43	.38	.22
		1.5	.13	.25	.44	.24	.02	.13	.24	.17
		2.5	.04	.06	.04	.00	.00	.01	.04	.04
	-2	-1	.06	.07	.05	.01	.01	.05	.04	.03
		0	.23	.39	.48	.48	.34	.29	.19	.11
		1	.32	.68	1.03	.48	.06	.25	.36	.24
		2	.17	.34	.48	.07	.00	.07	.21	.19
		3	.05	.07	.05	.00	.00	.01	.04	.05
	-5	0.5	.10	11	.10	.05	.03	.10	.08	.05
		1.5	.50	.50	.08	.00	.00	.05	.26	.18
		2.5	.91	.82	.03	.00	.00	.01	.17	.29
		3.5	.49	.40	.02	.00	.00	.00	.07	.18
		4.5	.10	.11	.02	.00	.00	.00	.03	.06

Table 4.5.1 (continued)

|k|	β_0	L_x	Δ_β -3	-2	-1	-0.5	0.5	1	2	3
0.5	0	-2	.04	.04	.02	.00	.00	.03	.04	.03
		-1	.15	.24	.25	.13	.37	.32	.20	.11
		0	.22	.43	.65	.68	.68	.65	.43	.22
		1	.11	.20	.32	.37	.13	.25	.24	.15
		2	.03	.04	.03	.00	.00	.02	.04	.04
	-1	-1.5	.04	.05	.03	.00	.00	.03	.04	.03
		-0.5	.16	.28	.36	.35	.30	.27	.18	.11
		0.5	.25	.52	.94	.97	.21	.40	.38	.22
		1.5	.13	.24	.42	.35	.04	.16	.22	.15
		2.5	.03	.05	.04	.00	.00	.02	.04	.04
	-2	-1	.05	.06	.05	.01	.01	.04	.04	.03
		0	.20	.36	.47	.50	.21	.25	.18	.11
		1	.31	.70	1.05	.33	.05	.21	.34	.23
		2	.16	.32	.47	.13	.01	.08	.20	.16
		3	.04	.06	.05	.00	.00	.01	.04	.04
	-5	0.5	.08	.09	.09	.02	.01	.06	.06	.05
		1.5	.46	.47	.05	.00	.00	.02	.19	.17
		2.5	.95	.75	.02	.00	.00	.01	.14	.26
		3.5	.44	.40	.02	.00	.00	.00	.08	.17
		4.5	.08	.09	.02	.00	.00	.00	.04	.05
0.8	0	-2	.02	.03	.02	.00	.00	.02	.03	.02
		-1	.09	.15	.19	.16	.23	.21	.14	.08
		0	.15	.30	.45	.43	.43	.45	.30	.15
		1	.08	.14	.21	.23	.16	.19	.15	.09
		2	.02	.03	.02	.00	.00	.02	.03	.02
	-1	-1.5	.02	.03	.02	.00	.00	.02	.02	.02
		-0.5	.10	.18	.26	.30	.15	.17	.13	.08
		0.5	.17	.37	.68	.58	.13	.26	.26	.15
		1.5	.09	.17	.27	.30	.06	.14	.14	.09
		2.5	.02	.03	.03	.00	.00	.01	.02	.02

Table 4.5.1 (continued)

|k|	β_0	L_x	Δ_β -3	-2	-1	-0.5	0.5	1	2	3
0.8	-2	-1	.03	.04	.04	.01	.00	.02	.02	.02
		0	.12	.23	.33	.30	.05	.13	.12	.08
		1	.21	.50	.73	.17	.03	.13	.23	.16
		2	.11	.22	.33	.17	.02	.08	.13	.09
		3	.03	.04	.04	.01	.00	.01	.02	.02
	-5	0.5	.05	.06	.05	.00	.00	.02	.04	.03
		1.5	.30	.31	.02	.00	.00	.01	.09	.12
		2.5	.68	.49	.01	.00	.00	.01	.09	.16
		3.5	.29	.29	.02	.00	.00	.00	.06	.11
		4.5	.05	.06	.03	.00	.00	.01	.03	.03
1	0	-2	.03	.05	.03	.00	.00	.01	.05	.05
		-1	.11	.20	.34	.32	.06	.21	.27	.18
		0	.23	.41	.61	.68	.68	.61	.41	.23
		1	.18	.27	.21	.06	.32	.34	.20	.11
		2	.05	.05	.01	.00	.00	.03	.05	.03
	-1	-1.5	.04	.06	.04	.00	.00	.01	.04	.05
		-0.5	.13	.25	.44	.21	.01	.12	.24	.18
		0.5	.25	.50	.84	1	.26	.43	.37	.22
		1.5	.19	.31	.34	.21	.34	.29	.18	.11
		2.5	.05	.06	.02	.00	.00	.04	.04	.03
	-2	-1	.05	.07	.05	.00	.00	.01	.04	.05
		0	.17	.34	.47	.06	.00	.06	.21	.19
		1	.32	.66	1	.52	.06	.26	.35	.24
		2	.23	.39	.47	.46	.35	.29	.19	.11
		3	.06	.07	.05	.01	.01	.05	.05	.03
	-5	0.5	.10	.11	.01	.00	.00	.00	.03	.06
		1.5	.48	.39	.02	.00	.00	.00	.07	.18
		2.5	.88	.81	.03	.00	.00	.01	.17	.29
		3.5	.50	.49	.09	.00	.00	.06	.27	.18
		4.5	.11	.11	.10	.05	.05	.10	.08	.05



(1) $\beta o = 0$

(2) $\beta o = -1$

(3) $\beta o = -2$

(4) $\beta o = -5$

Figure 4.5.1 D efficiency of alternative parallel-line design, $|k| = 0.5$

Table 4.5.2 Maximum D efficiencies of alternative parallel-line design

$\|k\|$	β_0	Max D-EFF	L_x	Δ_β
0.2	0	1.01122	0.176667	-0.172197
	-1	1.03364	0.534030	-0.549454
	-2	1.03509	1.01119	-1.03153
	-5	1.03540	2.50131	-2.51615
		Max given $\|k\|$:		
	$-\infty$	1.03542*	$-\beta_0/2$	$\beta_0/2$
0.5	0	0.919936	0.392824	-0.314002
	-1	1.05136	0.610963	-0.683746
	-2	1.07768	1.03433	-1.12989
	-5	1.08487	2.50373	-2.56918
		Max given $\|k\|$:		
	$-\infty$	1.08527*	$-\beta_0/2$	$\beta_0/2$
0.8	0	0.500254	0.664515	-0.191627
	-1	0.721647	0.583482	-0.762561
	-2	0.770657	1.02082	-1.18972
	-5	0.784024	2.50206	-2.60232
		Max given $\|k\|$:		
	$-\infty$	0.784779*	$-\beta_0/2$	$\beta_0/2$
1	0	1**	0**	0**
	-1	1	0.5	-0.5
	-2	1	1	-1
	-5	1	2.5	-2.5
		Max given $\|k\|$:		
	$\beta_0 \neq 0$	1	$-\beta_0/2$	$\beta_0/2$
		Global Max:		
0.439529*	$-\infty$	1.09196*	$-\beta_0/2$	$\beta_0/2$

* : These maximums are in fact the limits of the efficiency as β_0 tends to infinity while L_x assumes the value $-\beta_0/2$ and Δ_β assumes $\beta_0/2$.

** : Since Δ_β by definition is not allowed to be 0, the values are meaningful only in the limiting sense, which means that for $|k| = 1$ and $\beta_0 = 0$, the efficiency approaches 1 as L_x and Δ_β tend to zero.

§4.6 Ray Design vs. Modified Ray Design

As the ray design often results in poor performance with synergism, the modified ray design is proposed primarily as an alternative to better cope with synergistic conditions. More insight and motivation were discussed in §2.6. Taking into account the effect of interaction, the modified ray design relies on the hyperbola ED curves rather than the ED lines that were used in the ray design as a result of assuming no interaction. Let L_1' and L_2' denote the logits of the ray design. Let L_1 and L_2 denote the logits and g denote a ray slope of the modified ray design. By the definition in (4.1.1), the D efficiency of the ray design relative to the modified ray design can be expressed as

$$\text{D-EFF} = F(k, \beta_0, \Delta_\beta) = 4k\,f(k)^{\frac{1}{4}} \frac{(L_2' - L_1')[(L_1' - \beta_0)(L_2' - \beta_0)]^{\frac{1}{2}}}{\Delta_\beta^2}$$

$$\cdot \left([S(g, L_2) - S(g, L_1)][S(\frac{1}{g}, L_2) - S(\frac{1}{g}, L_1)] \right.$$

$$\left. \cdot [(\frac{1}{g^2} - 1)S(g, L_1)S(g, L_2) + (g^2 - 1)S(\frac{1}{g}, L_1)S(\frac{1}{g}, L_2)] \right)^{-\frac{1}{2}}$$

$$\cdot \frac{\left(\prod_{i=1}^{2} \prod_{j=-1,1} (1 + e^{jL_i}) \right)^{\frac{1}{2}}}{\left(\prod_{i=1}^{2} \prod_{j=-1,1} (1 + e^{j(L_i' + \frac{gk(L_i' - \beta_0)^2}{(1+gk)^2 \Delta_\beta})})(1 + e^{j(L_i' + \frac{(k/g)(L_i' - \beta_0)^2}{(1+(k/g))^2 \Delta_\beta})}) \right)^{\frac{1}{4}}}$$

(4.6.1)

where $k = \frac{\beta_1}{\beta_2}$, g is a function of k given in (2.3.3), the function f(k) is jointly given by (2.3.2) and (2.3.3), Δ_β is defined in (4.2.2), and the function S(g, L) is defined in (2.6.1). The efficiency is evaluated in table 4.6.1.

Table 4.6.1 D efficiency of ray design relative to modified ray design

		Δ_β			
β_0	k	0.5	1	2	5
0	.1	.4086	.7404	.9295	1.006
	.3	.3927	.7315	.9274	1.007
	.5	.3775	.7220	.9245	1.007
	.7	.3691	.7165	.9227	1.007
	.9	.3659	.7143	.9220	1.007
-1	.1	.3108	.6657	.8972	1.001
	.3	.2966	.6563	.8948	1.002
	.5	.2828	.6460	.8915	1.002
	.7	.2751	.6399	.8894	1.002
	.9	.2722	.6375	.8885	1.002
-2	.1	.1875	.5470	.8378	.9897
	.3	.1779	.5382	.8355	.9908
	.5	.1681	.5277	.8318	.9911
	.7	.1627	.5214	.8294	.9912
	.9	.1607	.5189	.8284	.9912
-5	.1	.0032	.1024	.4608	.8720
	.3	.0031	.1050	.4684	.8781
	.5	.0028	.1053	.4718	.8820
	.7	.0028	.1051	.4728	.8838
	.9	.0025	.1051	.4731	.8845

The efficiency is quite insensitive to the ratio k but is greatly affected by β_0 and Δ_β. Decreasing efficiency occurs as β_0 goes down, or the response probability at zero doses becomes lower. In the meantime, the efficiency drops very quickly as Δ_β decreases.

In fact, when β_1 and β_2 are considered to remain constant, a smaller value of $\Delta_\beta = \frac{\beta_1\beta_2}{\beta_{12}}$, or a shorter logit distance from the origin to the hyperbola center, would indicate stronger synergism. This means that as the interaction in the form of synergism gets more severe, the modified ray design would become more advantageous over the regular ray design.

By comparing the efficiencies in table 4.2.1 to those in table 4.6.1, one can obtain the efficiency of the modified ray design relative to the hyperbola-based design, which is given in table 4.6.2

Table 4.6.2 D efficiency of modified ray design relative to hyperbola-based design
(efficiency of ray design relative to hyperbola-based design in parentheses)

		Δ_β			
β_0	k	0.5	1	2	5
0	.1	**.14** (.06)	**.09** (.07)	**.05** (.05)	**.03** (.03)
	.3	**.12** (.05)	**.08** (.06)	**.05** (.04)	**.02** (.02)
	.5	**.10** (.04)	**.06** (.05)	**.04** (.04)	**.02** (.02)
	.7	**.07** (.03)	**.04** (.03)	**.03** (.03)	**.01** (.01)
	.9	**.04** (.01)	**.03** (.02)	**.02** (.01)	**.01** (.01)
-1	.1	**.24** (.08)	**.16** (.10)	**.10** (.09)	**.05** (.05)
	.3	**.21** (.06)	**.14** (.09)	**.08** (.08)	**.04** (.04)
	.5	**.17** (.05)	**.11** (.07)	**.07** (.06)	**.03** (.03)
	.7	**.13** (.03)	**.08** (.05)	**.05** (.04)	**.02** (.02)
	.9	**.07** (.02)	**.04** (.03)	**.03** (.02)	**.01** (.01)
-2	.1	**.35** (.07)	**.23** (.12)	**.14** (.12)	**.07** (.07)
	.3	**.30** (.05)	**.20** (.11)	**.12** (.10)	**.06** (.06)
	.5	**.24** (.04)	**.16** (.08)	**.10** (.08)	**.05** (.05)
	.7	**.18** (.03)	**.12** (.06)	**.07** (.06)	**.04** (.04)
	.9	**.10** (.02)	**.06** (.03)	**.04** (.03)	**.02** (.02)
-5	.1	**.60** (.00)	**.40** (.04)	**.26** (.12)	**.14** (.12)
	.3	**.52** (.00)	**.35** (.04)	**.23** (.11)	**.12** (.11)
	.5	**.42** (.00)	**.28** (.03)	**.19** (.09)	**.10** (.09)
	.7	**.31** (.00)	**.21** (.02)	**.14** (.06)	**.07** (.06)
	.9	**.17** (.00)	**.11** (.01)	**.08** (.04)	**.04** (.04)

For a quick reference on how much improvement the modified ray design possesses, the efficiency of the ray design relative to the hyperbola-based design is also listed next in parentheses in table 4.6.2. Apparently, the most difficult situation for the ray design is strong synergism accompanied by a fairly low zero-dose response probability. Quite to the contrary, the very same type of scenario is actually the most optimistic condition for the modified ray design. It seems that the worse the ray design performs, the greater the improvement of the modified ray design tends to be. Even though the modified ray design may still seem not so efficient compared to the hyperbola-based design, one should keep in mind that the table numbers do not really present the practical side of the story, in which the hyperbola-based design is often ruled out even from the candidate list because of negative doses or other practical concerns discussed in §2.6.

In general, the ray design appears less efficient than the modified ray design most of the time, especially under strong synergism with a close to zero response probability at zero doses. From the efficiency stand point, the modified ray design may indeed be a meaningful alternative to the regular ray design in a synergistic situation.

CHAPTER 5

ROBUSTNESS TO PARAMETER MISSPECIFICATION

The optimal designs developed in the previous chapters are expressed through ED_{100P}'s along with certain geometric constraints due to the two-variable design space. To determine the actual design points, the values of all model parameters must be known. Parameter guesses in a real experiment can rarely be perfect, especially for the interaction case where the knowledge about the interaction coefficient is often scarce. If the parameters are misspecified, the implemented design will not be the desired optimal design. This chapter examines the effect of parameter misspecification on several D optimal designs. The robustness property of the design under parameter misspecification is evaluated through the D efficiency defined as

$$\text{D-EFF} = \left(\frac{|I(\beta)| \text{ for the design obtained from misspecified parameters}}{|I(\beta)| \text{ for the design obtained from true parameters}} \right)^{\frac{1}{p}}$$

where p is the number of parameters. Due to the additional geometric constraints, parameter misspecification can distort a design in a more complicated way than it does in the one-variable logistic case where a design is given purely in terms of ED's.

§5.1 Parallel-Line Design

The D optimal parallel-line design is given by $ED_{22.7}$ and $ED_{77.3}$ with equal sample allocations. The construction of such a design is discussed in §2.1. Assume that the experimenter computes the design points using the design matrix given in (2.1.7). Let

β_0, β_1, and β_2 denote the true parameters and β_0', β_1', and β_2' the specified values. The degree of misspecification can be quantified through the ratios

$$m_0 = \frac{\beta_0'}{\beta_0}, \qquad m_1 = \frac{\beta_1'}{\beta_1}, \qquad \text{and} \qquad m_2 = \frac{\beta_2'}{\beta_2}, \tag{5.1.1}$$

where $\beta_j \neq 0$, j = 0, 1, 2. The design points computed with the misspecified parameters no longer fall on $ED_{22.7}$ and $ED_{77.3}$ but rather have their actual logit levels given by

$$L_1 = \beta_0(1 - \frac{m_0}{m_2}) + (L_x - L_D)(1 - \frac{m_1}{m_2}) - L\frac{m_1k^2 + m_2}{m_1^2k^2 + m_2^2},$$

$$L_2 = \beta_0(1 - \frac{m_0}{m_2}) + (L_x - L_D)(1 - \frac{m_1}{m_2}) + L\frac{m_1k^2 + m_2}{m_1^2k^2 + m_2^2},$$

$$L_3 = \beta_0(1 - \frac{m_0}{m_2}) + (L_x + L_D)(1 - \frac{m_1}{m_2}) - L\frac{m_1k^2 + m_2}{m_1^2k^2 + m_2^2},$$

$$L_4 = \beta_0(1 - \frac{m_0}{m_2}) + (L_x + L_D)(1 - \frac{m_1}{m_2}) + L\frac{m_1k^2 + m_2}{m_1^2k^2 + m_2^2}, \tag{5.1.2}$$

where $L_x = \beta_1 x_{10}$, $L_D = \beta_1 D_1$, k is the slope of the true ED's given by $-\frac{\beta_1}{\beta_2}$, and L is the optimal logit 1.22291 for the parallel-line design. When $\beta_0 = 0$, the term $\beta_0(1 - \frac{m_0}{m_2})$ in (5.1.2) should be replaced by $-\frac{\beta_0'}{m_2}$. The scale free index L_x stands for the user selected design location and was explained in detail in §4.3. The scale free distance $L_D = \beta_1 D_1$ describes the length of the design along the ED edges, where the distance D_1 is initially defined in figure 2.1.1. The scale of the design is in fact characterized by the two logit type distances L and L_D, which address two orthogonal directions: the direction with the steepest variation in the response probability and the one with a constant response probability.

The D criterion of the constructed design is then compared to that of the optimal design. This relative efficiency measures the design robustness and is given by

$$\text{D-EFF} = F(k, \beta_0, L_x, L_D, m_0, m_1, m_2) = (m_2)^{-\frac{2}{3}} \frac{2+e^{-L}+e^{L}}{\left(\dfrac{\prod_{i=1}^{4}(2+e^{-L_i}+e^{L_i})}{\frac{1}{4}\sum_{i=1}^{4}(2+e^{-L_i}+e^{L_i})}\right)^{\frac{1}{3}}}. \tag{5.1.3}$$

An example of design misplacement appears in figure 5.1.1. The optimal design in solid lines, when implemented with the wrong parameters, becomes the incorrect design in dashed lines. Table 5.1.1 gives the design points and their actual logits for both the correct and incorrect designs. In this example, the D efficiency of the constructed design relative to the optimal design is 45.95%.

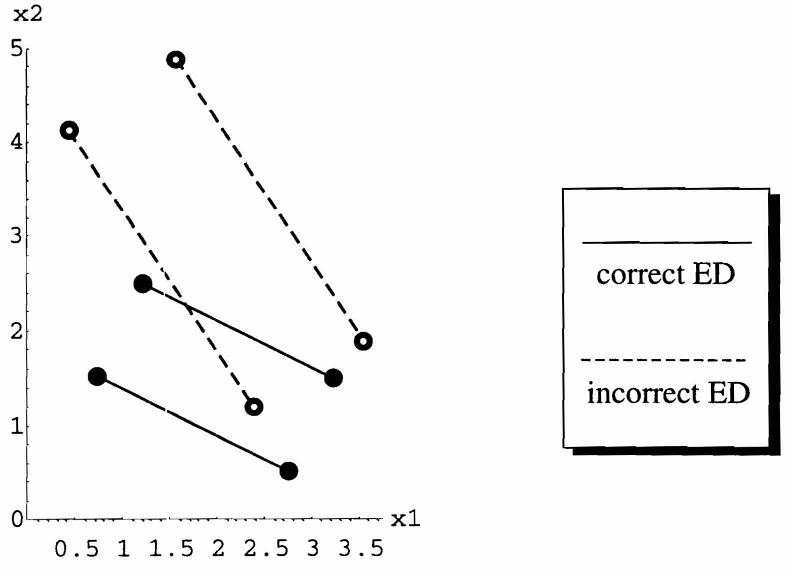


Figure 5.1.1 Design shift due to parameter misspecification, parallel-line design

Table 5.1.1 Design shift due to parameter misspecification, parallel-line design ($x_{10} = 2, D_1 = 1$)

Correct design ($\beta_0 = -5, \beta_1 = 1, \beta_2 = 2$)		Misplaced design ($\beta'_0 = -6, \beta'_1 = 1.5, \beta'_2 = 1$)	
(x_1, x_2)	Logit	(x_1, x_2)	Logit
(0.76, 1.51)	-1.22	(0.44, 4.12)	3.68
(1.24, 2.49)	1.22	(1.56, 4.88)	6.32
(2.76, 0.51)	-1.22	(2.44, 1.12)	-0.32
(3.24, 1.49)	1.22	(3.56, 1.88)	2.32

The robustness property in a simplified situation assuming no misspecification in the slope k, or $m_1 = m_2 = m$, is first inspected in table 5.1.2 so that the effect of m_0 can be looked at more carefully. The efficiency function is free from k, L_x, and L_D due to the constraint $m_1 = m_2$. It seems that the design is slightly more robust to overestimation than to underestimation of the linear effects, especially when $|\beta_0|$ is small. The worst scenario occurs with an overestimated β_0 accompanied by underestimated linear effects, especially when $|\beta_0|$ is already large. Furthermore, the efficiency function will be free from β_0 whenever the misspecification is proportional across all the parameters, or $m_0 = m_1 = m_2 = m$, and is a function of m only. A general impression here is that the design appears somewhat more robust to overestimation than to underestimation of the parameters.

The effect of misspecification is further investigated subject to no constraint. Selected results are given in tables 5.1.3 and 5.1.4. Many other tables are obtained but not shown here as they reflect similar results. It seems that the design is less robust when it is constructed longer in the direction of constant probabilities, characterized by a larger L_D. Recall that the D criterion of the optimal design based on the true parameters always improves as L_D increases. A design with longer ED edges is hence more desirable under good parameter knowledge but will be somewhat penalized when the parameter knowledge is poor.

Table 5.1.2 D efficiency under parameter misspecification, parallel-line design, $m_1 = m_2 = m$

		m							
$\|\beta_0\|$	m_0	.25	.5	.75	1	1.25	1.5	1.75	2
0	0 ($\|\beta_0'\|$)	.11	.66	.94	1	.97	.92	.87	.82
	.5 ($\|\beta_0'\|$)	.16	.68	.92	.97	.95	.90	.86	.81
	1 ($\|\beta_0'\|$)	.21	.63	.82	.88	.88	.85	.81	.78
	1.5 ($\|\beta_0'\|$)	.10	.44	.64	.73	.77	.77	.75	.73
	2 ($\|\beta_0'\|$)	.02	.23	.44	.57	.64	.66	.67	.66
1	.25	.11	.67	.92	.93	.88	.81	.75	.70
	.5	.12	.66	.94	.97	.92	.85	.78	.73
	.75	.16	.67	.94	.99	.95	.88	.81	.75
	1	.19	.68	.94	1	.97	.90	.84	.78
	1.25	.21	.67	.92	.99	.97	.92	.86	.79
	1.5	.17	.63	.88	.97	.97	.92	.87	.81
	1.75	.10	.55	.82	.93	.95	.92	.87	.81
	2	.05	.44	.74	.88	.92	.90	.87	.82
2	.25	.11	.68	.82	.73	.64	.56	.49	.44
	.5	.16	.66	.92	.88	.77	.66	.58	.52
	.75	.21	.68	.94	.97	.88	.77	.67	.59
	1	.10	.63	.92	1	.95	.85	.75	.66
	1.25	.02	.44	.82	.97	.97	.90	.81	.73
	1.5	.00	.23	.64	.88	.95	.92	.86	.78
	1.75	.00	.10	.44	.73	.88	.90	.87	.81
	2	.00	.04	.27	.57	.77	.85	.86	.82
5	.25	.11	.55	.27	.15	.10	.07	.06	.05
	.5	.17	.66	.74	.41	.24	.15	.11	.08
	.75	.00	.55	.94	.81	.50	.31	.21	.15
	1	.00	.10	.74	1	.82	.56	.37	.26
	1.25	.00	.01	.27	.81	.97	.81	.58	.41
	1.5	.00	.00	.06	.41	.82	.92	.78	.59
	1.75	.00	.00	.01	.15	.50	.81	.87	.75
	2	.00	.00	.00	.05	.24	.56	.78	.82

Table 5.1.3 D efficiency under parameter misspecification, parallel-line design, $\beta_0 = -1$, $L_x = 0.5$, $L_D = 1$

			m_2							
|k|	m_0	m_1	.25	.5	.75	1	1.25	1.5	1.75	2
.5	.5	.25	.12	.60	.87	.90	.86	.80	.74	.68
		.5	.40	.66	.89	.93	.89	.82	.76	.70
		.75	.91	.84	.94	.95	.90	.84	.78	.73
		1	.74	1.01	.99	.97	.92	.85	.79	.73
		1.25	.35	1.02	1.02	.97	.92	.86	.79	.74
		1.5	.15	.86	.99	.96	.91	.85	.79	.74
		1.75	.06	.65	.90	.92	.88	.84	.78	.74
		2	.03	.47	.78	.86	.85	.81	.77	.73
	1	.25	.19	.62	.85	.89	.86	.80	.75	.69
		.5	.48	.68	.88	.93	.90	.84	.78	.72
		.75	.79	.84	.94	.96	.93	.87	.80	.74
		1	.72	1.01	1.01	1	.95	.89	.82	.76
		1.25	.35	1.04	1.07	1.03	.97	.90	.83	.77
		1.5	.12	.89	1.06	1.03	.97	.90	.84	.78
		1.75	.04	.66	.98	1.00	.96	.90	.84	.78
		2	.02	.44	.85	.95	.93	.88	.83	.78
	1.5	.25	.17	.56	.77	.83	.82	.78	.74	.69
		.5	.27	.63	.81	.88	.87	.82	.77	.72
		.75	.37	.76	.88	.92	.90	.86	.80	.75
		1	.44	.89	.96	.97	.94	.89	.83	.77
		1.25	.37	.95	1.02	1.01	.97	.91	.85	.79
		1.5	.18	.86	1.03	1.03	.98	.92	.86	.80
		1.75	.06	.66	.98	1.02	.99	.93	.87	.81
		2	.02	.44	.86	.98	.97	.92	.86	.81
1	.5	.25	.12	.44	.74	.83	.82	.78	.72	.68
		.5	.75	.66	.79	.84	.83	.79	.74	.69
		.75	1.22	1.03	.94	.90	.86	.81	.75	.70
		1	.73	1.22	1.07	.97	.89	.83	.77	.71
		1.25	.32	1.13	1.12	1.01	.92	.84	.78	.72
		1.5	.14	.90	1.06	1.01	.92	.85	.79	.73
		1.75	.06	.66	.95	.96	.91	.84	.78	.73
		2	.03	.47	.81	.89	.87	.83	.77	.73
	1	.25	.19	.48	.73	.82	.82	.78	.73	.69
		.5	.71	.68	.78	.84	.84	.80	.75	.70
		.75	.98	1.01	.94	.91	.88	.83	.77	.72
		1	.75	1.22	1.10	1	.92	.86	.80	.74
		1.25	.32	1.17	1.18	1.07	.97	.88	.82	.75
		1.5	.11	.93	1.16	1.09	.99	.90	.83	.77
		1.75	.04	.65	1.04	1.06	.99	.91	.84	.77
		2	.01	.43	.88	.99	.96	.90	.84	.78
	1.5	.25	.17	.47	.69	.78	.79	.76	.72	.68
		.5	.27	.63	.74	.81	.81	.79	.75	.70
		.75	.39	.85	.88	.88	.86	.82	.77	.73
		1	.48	1.01	1.02	.97	.91	.86	.80	.75
		1.25	.38	1.04	1.12	1.05	.97	.89	.83	.77
		1.5	.17	.90	1.12	1.09	1.01	.92	.85	.79
		1.75	.06	.66	1.04	1.08	1.02	.94	.87	.80
		2	.02	.43	.89	1.03	1.01	.94	.87	.81

Table 5.1.4 D efficiency under parameter misspecification, parallel-line design, $\beta_0 = -2$, $L_x = 1$, $|k| = 0.5$

			m_2							
L_D	m_0	m_1	.25	.5	.75	1	1.25	1.5	1.75	2
1	.5	.25	.16	.60	.87	.90	.85	.78	.71	.65
		.5	.42	.66	.88	.91	.85	.78	.71	.66
		.75	.91	.84	.92	.90	.84	.77	.71	.65
		1	.74	.97	.93	.88	.81	.75	.69	.64
		1.25	.39	.89	.88	.82	.77	.71	.66	.61
		1.5	.19	.70	.77	.75	.71	.66	.62	.59
		1.75	.10	.52	.65	.66	.64	.61	.58	.55
		2	.05	.38	.53	.57	.57	.56	.54	.52
	1	.25	.10	.52	.77	.85	.85	.80	.75	.69
		.5	.20	.63	.83	.91	.89	.84	.77	.71
		.75	.37	.80	.92	.96	.93	.86	.79	.73
		1	.53	.97	1.01	1	.95	.87	.80	.73
		1.25	.39	1.04	1.07	1.02	.95	.87	.79	.73
		1.5	.13	.89	1.05	1.00	.93	.85	.78	.71
		1.75	.04	.66	.95	.95	.88	.81	.75	.69
		2	.02	.46	.81	.86	.82	.77	.71	.66
	1.5	.25	.00	.16	.44	.63	.71	.73	.71	.68
		.5	.01	.23	.54	.71	.78	.78	.75	.71
		.75	.01	.33	.64	.79	.84	.83	.79	.74
		1	.04	.47	.76	.88	.90	.87	.82	.77
		1.25	.10	.62	.88	.95	.95	.91	.85	.78
		1.5	.16	.70	.95	1.00	.98	.92	.86	.79
		1.75	.10	.64	.95	1.01	.99	.92	.86	.79
		2	.03	.46	.86	.98	.96	.91	.84	.78
2	.5	.25	.16	.61	.78	.72	.63	.55	.48	.43
		.5	.48	.66	.87	.83	.73	.63	.55	.49
		.75	.37	.84	.92	.88	.79	.69	.61	.54
		1	.06	.77	.90	.88	.80	.71	.64	.57
		1.25	.01	.47	.78	.81	.77	.70	.64	.58
		1.5	.00	.24	.60	.70	.70	.66	.62	.57
		1.75	.00	.11	.42	.58	.62	.61	.58	.55
		2	.00	.05	.28	.46	.53	.54	.54	.52
	1	.25	.10	.51	.70	.70	.63	.55	.49	.44
		.5	.19	.63	.82	.83	.75	.66	.58	.51
		.75	.24	.78	.92	.94	.86	.75	.66	.58
		1	.10	.77	.98	1	.93	.82	.72	.64
		1.25	.01	.46	.91	.99	.95	.86	.76	.67
		1.5	.00	.19	.69	.90	.91	.85	.77	.69
		1.75	.00	.07	.45	.74	.82	.80	.75	.69
		2	.00	.03	.27	.57	.71	.73	.71	.66
	1.5	.25	.00	.17	.43	.56	.57	.53	.48	.44
		.5	.01	.23	.53	.67	.68	.63	.57	.51
		.75	.02	.34	.64	.78	.79	.73	.66	.59
		1	.05	.44	.75	.88	.88	.82	.74	.66
		1.25	.03	.41	.78	.93	.95	.89	.81	.72
		1.5	.00	.24	.67	.90	.96	.92	.85	.76
		1.75	.00	.09	.45	.77	.90	.91	.86	.78
		2	.00	.03	.25	.59	.79	.85	.83	.78

§5.2 Hyperbola-Based Design

The D optimal hyperbola-based design is given by $ED_{17.6}$ and $ED_{82.4}$ with equal sample allocations. The design implementation was discussed in §2.2. Assume that the design is constructed from the design matrix given in (2.2.7) with the guessed parameters $\beta_0^{*\prime}$, β_1', β_2', and β_{12}', where $\beta_0^{*\prime}$ is the guessed logit on the centered axes and is given by $\beta_0^{*\prime} = \beta_0' - \frac{\beta_1'\beta_2'}{\beta_{12}'}$. The misspecification can be characterized by the ratios

$$m_0^* = \frac{\beta_0^{*\prime}}{\beta_0^*}, \quad m_1 = \frac{\beta_1'}{\beta_1}, \quad m_2 = \frac{\beta_2'}{\beta_2}, \quad \text{and} \quad m_{12} = \frac{\beta_{12}'}{\beta_{12}}, \tag{5.2.1}$$

where $\beta_0^* = \beta_0 - \frac{\beta_1\beta_2}{\beta_{12}}$ and $\beta_0^* \neq 0$. The actual logits of the constructed design will not correspond to $ED_{17.6}$ and $ED_{82.4}$ but are given by

$$L_1 = \beta_0^* + [\,\Delta_\beta(1 - \frac{m_2}{m_{12}}) - L_D\,]\,[\,(1 - \frac{m_1}{m_{12}}) - \frac{L - m_0^*\beta_0^*}{m_{12}L_D}\,],$$
$$L_2 = \beta_0^* + [\,\Delta_\beta(1 - \frac{m_2}{m_{12}}) - L_D\,]\,[\,(1 - \frac{m_1}{m_{12}}) + \frac{L + m_0^*\beta_0^*}{m_{12}L_D}\,],$$
$$L_3 = \beta_0^* + [\,\Delta_\beta(1 - \frac{m_2}{m_{12}}) + L_D\,]\,[\,(1 - \frac{m_1}{m_{12}}) - \frac{L + m_0^*\beta_0^*}{m_{12}L_D}\,],$$
$$L_4 = \beta_0^* + [\,\Delta_\beta(1 - \frac{m_2}{m_{12}}) + L_D\,]\,[\,(1 - \frac{m_1}{m_{12}}) + \frac{L - m_0^*\beta_0^*}{m_{12}L_D}\,], \tag{5.2.2}$$

where L is the optimal logit 1.54341 for the hyperbola-based design, $\Delta_\beta = \beta_0 - \beta_0^* = \frac{\beta_1\beta_2}{\beta_{12}}$, and $L_D = \beta_1 z_1$, where z_1 is the equidistant deviation in x_1 selected by the user and was better explained in §2.2. The quantity L_D is a logit type measure of the centered design level. As initially defined in (4.2.2), the quantity Δ_β is the difference in logit between the origin and the hyperbola center. When $\beta_0^* = 0$, the term $m_0^*\beta_0^*$ in (5.2.2) is replaced by $\beta_0^{*\prime}$.

The robustness is measured by the D efficiency of the design obtained from the guessed parameters relative to the one from the true parameters. The D efficiency is given by

$$\text{D-EFF} = F(\beta_0^*, \Delta_\beta, L_D, m_0^*, m_1, m_2, m_{12}) = \frac{2+e^{-L}+e^{L}}{|m_{12}|\left(\prod_{i=1}^{4}(2+e^{-L_i}+e^{L_i})\right)^{\frac{1}{4}}}. \quad (5.2.3)$$

An example of design misplacement due to wrong parameter knowledge is shown in figure 5.2.1. The solid curves are associated with the optimal design and the dashed ones with the misplaced design. Table 5.2.1 lists the design points along with their actual logits for both setups. In this example, the D efficiency of the constructed design relative to the optimal design is 51.44%.

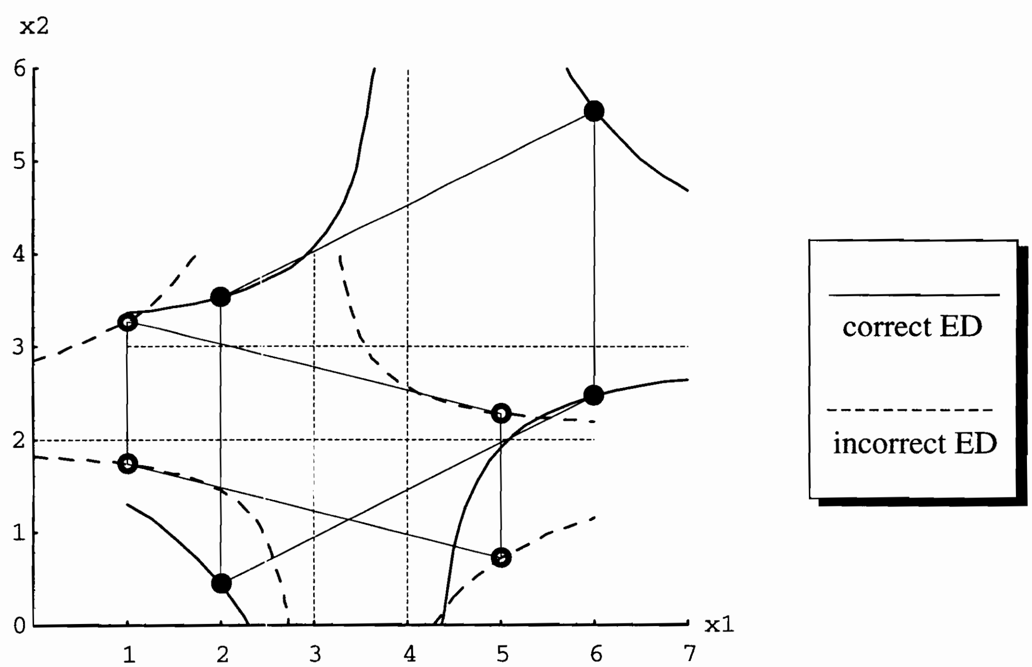


Figure 5.2.1 Design shift due to parameter misspecification, hyperbola-based design

Table 5.2.1 Design shift due to parameter misspecification, hyperbola-based design ($z_1 = 2$)

Correct design ($\beta_0 = -5, \beta_1 = 1.5, \beta_2 = 2, \beta_{12} = -0.5$)		Misplaced design ($\beta_0' = -7, \beta_1' = 2, \beta_2' = 3, \beta_{12}' = -1$)	
(x_1, x_2)	Logit	(x_1, x_2)	Logit
(2, 3.54)	1.54	(1, 3.27)	1.41
(2, 0.46)	-1.54	(1, 1.73)	-0.91
(6, 5.54)	-1.54	(5, 2.27)	1.36
(6, 2.46)	1.54	(5, 0.73)	2.14

The robustness property in a simplified situation assuming $m_1 = m_2 = m_{12} = m$ is first illustrated in table 5.2.2 where the effect of m_0^* can be observed more carefully. The efficiency function is free from Δ_β and L_D due to the constraint. The design seems more robust to overestimation in β_0^* when the other effects are also overestimated. However, the design is likely to be badly affected by an overestimated β_0^* when the other effects are underestimated, especially if $|\beta_0^*|$ is already large. The efficiency function will be further free from β_0^* when the misspecification is proportional across all the parameters, or $m_0^* = m_1 = m_2 = m_{12} = m$, and becomes a function of m only. A general impression is that the design appears somewhat more robust to overestimation than to underestimation of the parameters.

The effect of misspecification is further studied subject to less constraint. Selected results are given in tables 5.2.3 and 5.2.4. Many other tables were obtained but not shown here as they reflect the same results. To be more consistent with other illustrations in this dissertation, tables 5.2.3 and 5.2.4 are indexed by β_0 and Δ_β whereas they could have been indexed by β_0^* and Δ_β which appear in the efficiency function in (5.2.3). A tendency similar to the previous observation is revealed here. The design appears to be most adversely affected when the central logit β_0^* is overestimated while the interaction effect is underestimated, especially if the linear effects are both overestimated. Apart from this vulnerable point, the efficiency often seems quite reasonable.

Table 5.2.2 D efficiency under parameter misspecification, hyperbola-based design, $m_1 = m_2 = m_{12} = m$

		m							
$\lvert\beta_0^*\rvert$	m_0^*	.25	.5	.75	1	1.25	1.5	1.75	2
0	0 ($\lvert\beta_0^{*\prime}\rvert$)	.06	.58	.92	1	.96	.89	.82	.75
	.5 ($\lvert\beta_0^{*\prime}\rvert$)	.06	.55	.88	.96	.94	.87	.80	.74
	1 ($\lvert\beta_0^{*\prime}\rvert$)	.05	.47	.77	.86	.86	.82	.76	.71
	1.5 ($\lvert\beta_0^{*\prime}\rvert$)	.03	.33	.59	.72	.75	.74	.70	.66
	2 ($\lvert\beta_0^{*\prime}\rvert$)	.01	.18	.41	.56	.62	.64	.63	.60
1	.25	.06	.57	.88	.92	.86	.78	.70	.63
	.5	.06	.58	.91	.96	.90	.82	.73	.66
	.75	.06	.57	.92	.99	.94	.85	.76	.69
	1	.06	.55	.91	1	.96	.87	.79	.71
	1.25	.05	.52	.88	.99	.96	.89	.80	.72
	1.5	.04	.47	.83	.96	.96	.89	.81	.74
	1.75	.03	.40	.77	.92	.94	.89	.82	.74
	2	.02	.33	.68	.86	.90	.87	.81	.75
2	.25	.06	.55	.77	.72	.62	.53	.46	.40
	.5	.06	.58	.88	.86	.75	.64	.54	.47
	.75	.05	.55	.92	.96	.86	.74	.63	.54
	1	.03	.47	.88	1	.94	.82	.70	.60
	1.25	.01	.33	.77	.96	.96	.87	.76	.66
	1.5	.00	.18	.59	.86	.94	.89	.80	.71
	1.75	.00	.08	.41	.72	.86	.87	.82	.74
	2	.00	.03	.26	.56	.75	.82	.80	.75
5	.25	.06	.40	.26	.15	.09	.07	.05	.04
	.5	.04	.58	.68	.40	.23	.15	.10	.08
	.75	.00	.40	.92	.80	.49	.30	.19	.14
	1	.00	.08	.68	1	.81	.53	.34	.23
	1.25	.00	.01	.26	.80	.96	.78	.54	.37
	1.5	.00	.00	.06	.40	.81	.89	.73	.54
	1.75	.00	.00	.01	.15	.49	.78	.82	.69
	2	.00	.00	.00	.04	.23	.53	.73	.75

Table 5.2.3 D efficiency under parameter misspecification, hyperbola-based design, $m_1 = m_2 = m$, $\beta_0 = -1$, $L_D = 1$

			m_{12}							
Δ_β	m	m_C^*	.25	.5	.75	1	1.25	1.5	1.75	2
-.5	.5	.25	.05	.57	.89	.93	.86	.76	.68	.60
		.5	.05	.58	.90	.94	.87	.78	.69	.62
		.75	.05	.57	.91	.95	.89	.80	.71	.63
		1	.05	.57	.91	.96	.90	.81	.72	.64
		1.25	.05	.56	.90	.96	.91	.82	.73	.65
		1.5	.05	.55	.89	.96	.91	.83	.74	.66
		1.75	.05	.54	.87	.95	.91	.84	.75	.67
		2	.05	.52	.85	.94	.91	.84	.76	.68
	1	.25	.00	.51	.89	.98	.93	.84	.75	.66
		.5	.00	.51	.90	.99	.94	.85	.76	.67
		.75	.00	.51	.91	1.00	.95	.86	.77	.68
		1	.00	.50	.91	1	.95	.87	.78	.69
		1.25	.00	.50	.90	1.00	.96	.87	.78	.70
		1.5	.00	.49	.89	.99	.96	.88	.79	.71
		1.75	.00	.48	.87	.98	.95	.88	.79	.71
		2	.00	.47	.85	.96	.94	.88	.79	.72
	1.5	.25	.00	.29	.77	.93	.93	.86	.78	.70
		.5	.00	.30	.78	.94	.94	.87	.79	.71
		.75	.00	.30	.79	.95	.95	.88	.80	.72
		1	.00	.30	.80	.96	.95	.89	.80	.72
		1.25	.00	.30	.80	.96	.96	.89	.81	.73
		1.5	.00	.30	.80	.96	.96	.89	.81	.73
		1.75	.00	.30	.80	.95	.95	.89	.81	.74
		2	.00	.30	.79	.94	.94	.89	.81	.74
.5	.5	.25	.05	.56	.80	.77	.70	.63	.56	.51
		.5	.05	.58	.86	.84	.75	.67	.59	.53
		.75	.05	.56	.88	.88	.78	.69	.61	.55
		1	.04	.52	.86	.88	.79	.70	.62	.55
		1.25	.02	.44	.79	.85	.78	.70	.62	.56
		1.5	.01	.33	.68	.78	.75	.68	.61	.55
		1.75	.00	.21	.55	.69	.69	.64	.59	.53
		2	.00	.12	.41	.58	.62	.60	.56	.51
	1	.25	.00	.42	.80	.83	.74	.65	.58	.52
		.5	.00	.44	.86	.92	.83	.72	.63	.55
		.75	.00	.43	.88	.98	.89	.77	.67	.59
		1	.00	.37	.86	1	.93	.81	.70	.61
		1.25	.00	.27	.79	.98	.94	.83	.72	.63
		1.5	.00	.17	.68	.92	.92	.83	.73	.64
		1.75	.00	.10	.55	.83	.88	.81	.72	.64
		2	.00	.05	.41	.72	.81	.78	.70	.63
	1.5	.25	.00	.23	.62	.77	.74	.66	.59	.52
		.5	.00	.20	.64	.84	.83	.74	.64	.57
		.75	.00	.13	.62	.88	.89	.80	.70	.61
		1	.00	.07	.57	.88	.93	.85	.74	.65
		1.25	.00	.04	.48	.85	.94	.88	.78	.68
		1.5	.00	.02	.37	.78	.92	.89	.80	.70
		1.75	.00	.01	.28	.69	.88	.88	.80	.71
		2	.00	.00	.20	.58	.81	.85	.80	.71

Table 5.2.4 D efficiency under parameter misspecification, hyperbola-based design, $m_1 = m_2 = m$, $\beta_0 = -2$, $L_D = 1$

			m_{12}							
Δ_β	m	m_0^*	.25	.5	.75	1	1.25	1.5	1.75	2
-1	.5	.25	.03	.57	.83	.81	.71	.61	.52	.45
		.5	.03	.58	.86	.86	.76	.66	.56	.49
		.75	.03	.57	.88	.89	.80	.70	.60	.52
		1	.03	.55	.88	.91	.83	.73	.64	.55
		1.25	.03	.52	.86	.91	.86	.76	.67	.58
		1.5	.03	.47	.82	.91	.87	.78	.69	.61
		1.75	.02	.40	.77	.89	.87	.80	.71	.63
		2	.01	.33	.70	.85	.86	.80	.73	.65
	1	.25	.00	.28	.83	.92	.84	.74	.63	.55
		.5	.00	.28	.86	.96	.89	.78	.67	.58
		.75	.00	.28	.88	.99	.92	.81	.71	.61
		1	.00	.28	.88	1	.94	.84	.73	.64
		1.25	.00	.28	.86	.99	.95	.86	.76	.66
		1.5	.00	.27	.82	.96	.95	.87	.78	.68
		1.75	.00	.26	.77	.92	.93	.87	.79	.70
		2	.00	.24	.70	.86	.90	.86	.79	.71
	1.5	.25	.00	.02	.48	.81	.84	.78	.69	.61
		.5	.00	.02	.50	.86	.89	.82	.73	.64
		.75	.00	.01	.51	.89	.92	.85	.76	.67
		1	.00	.01	.51	.91	.94	.87	.78	.69
		1.25	.00	.01	.51	.91	.95	.89	.80	.71
		1.5	.00	.01	.51	.91	.95	.89	.81	.72
		1.75	.00	.01	.50	.89	.93	.89	.81	.73
		2	.00	.00	.48	.85	.90	.87	.81	.74
1	.5	.25	.01	.52	.51	.41	.36	.32	.29	.26
		.5	.01	.58	.66	.51	.42	.37	.33	.29
		.75	.00	.52	.73	.56	.45	.39	.34	.31
		1	.00	.33	.64	.53	.43	.37	.33	.30
		1.25	.00	.12	.42	.42	.36	.32	.29	.27
		1.5	.00	.03	.22	.29	.27	.26	.24	.23
		1.75	.00	.01	.10	.18	.19	.19	.19	.19
		2	.00	.00	.04	.11	.13	.14	.14	.15
	1	.25	.00	.13	.51	.48	.37	.31	.26	.24
		.5	.00	.12	.66	.72	.54	.42	.34	.29
		.75	.00	.05	.73	.92	.70	.52	.42	.35
		1	.00	.01	.64	1	.82	.61	.47	.39
		1.25	.00	.00	.42	.92	.85	.64	.50	.41
		1.5	.00	.00	.22	.72	.79	.63	.49	.40
		1.75	.00	.00	.10	.48	.64	.56	.45	.38
		2	.00	.00	.04	.28	.47	.46	.39	.34
	1.5	.25	.00	.02	.24	.41	.37	.30	.25	.22
		.5	.00	.00	.22	.51	.54	.43	.34	.29
		.75	.00	.00	.13	.56	.70	.58	.46	.37
		1	.00	.00	.06	.53	.82	.74	.57	.45
		1.25	.00	.00	.02	.42	.85	.85	.67	.52
		1.5	.00	.00	.01	.29	.79	.89	.74	.58
		1.75	.00	.00	.00	.18	.64	.85	.76	.61
		2	.00	.00	.00	.11	.47	.74	.72	.60

§5.3 Ray Design

The ray designs were derived by maximizing the determinant of the information matrix for an interaction model while assuming no interaction in the true relationship. To obtain the robustness efficiency, it would be reasonable to compute the D criterion in the same fashion as it was used to generate the optimal ray design. Namely, $|I(\beta)|$ in the efficiency expression will come from fitting an interaction model while the dose-response relationship involved in $|I(\beta)|$ follows straight line ED's. This will preserve the criterion value of the correctly implemented design based on the true parameters.

The optimal ED's of a ray design depend on β_0 while the optimal ray slope varies with $k = \frac{\beta_1}{\beta_2}$. The optimal design is first found with the specified parameters and can then be implemented using the design matrix

$$\begin{bmatrix} x_{11} & x_{21} \\ x_{12} & x_{22} \\ x_{13} & x_{23} \\ x_{14} & x_{24} \end{bmatrix} = \begin{bmatrix} \dfrac{L_1 - \beta_0}{\beta_1 + \beta_2 / g(k)} & \dfrac{L_1 - \beta_0}{\beta_1 g(k) + \beta_2} \\ \dfrac{L_2 - \beta_0}{\beta_1 + \beta_2 / g(k)} & \dfrac{L_2 - \beta_0}{\beta_1 g(k) + \beta_2} \\ \dfrac{L_1 - \beta_0}{\beta_1 + \beta_2 g(k)} & \dfrac{L_1 - \beta_0}{\beta_1 / g(k) + \beta_2} \\ \dfrac{L_2 - \beta_0}{\beta_1 + \beta_2 g(k)} & \dfrac{L_2 - \beta_0}{\beta_1 / g(k) + \beta_2} \end{bmatrix} \qquad (5.3.1)$$

where L_1 and L_2 are the optimal logits determined by the specific β_0 and $g(k)$ is the optimal ray slope expressed as a function of k given by (2.3.3).

In this section, it is assumed that the user implements the design according to the logits and ray slope solved from the guessed parameters. This design is then compared to the optimal design both solved and implemented with the true parameters. Normally, design robustness may as well be studied by addressing the implementation stage only without combining the issue of wrong solutions. Namely, a given optimal design is subject to pure implementation error and then compared to the same optimal design implemented with the true parameters. The current approach would involve more complexity but sometimes can better represent the reality.

The same ratios m_0, m_1, and m_2 as defined in (5.1.1) are used to characterize the misspecification. Let L_1^* and L_2^* represent the optimal logits resulting from the true intercept β_0 and $L_1^{*\prime}$ and $L_2^{*\prime}$ the optimal logits obtained from the specified intercept β_0'. The actual logits of the design points obtained from the guessed parameters are given by

$$L_1 = \beta_0 + \frac{g'k+1}{m_1 g'k + m_2}(L_1^{*\prime} - m_0\beta_0),$$

$$L_2 = \beta_0 + \frac{g'k+1}{m_1 g'k + m_2}(L_2^{*\prime} - m_0\beta_0),$$

$$L_3 = \beta_0 + \frac{k/g'+1}{m_1 k/g' + m_2}(L_1^{*\prime} - m_0\beta_0),$$

$$L_4 = \beta_0 + \frac{k/g'+1}{m_1 k/g' + m_2}(L_2^{*\prime} - m_0\beta_0), \qquad (5.3.2)$$

where g' is the optimal ray slope resulting from the guessed parameters and is given by the function $g(\frac{m_1}{m_2}k)$. The D efficiency of the design obtained from the specified parameters relative to the one given by the true parameters is expressed as

$$\text{D-EFF} = F(\beta_0, k, m_0, m_1, m_2, L_1^*, L_2^*, L_1^{*\prime}, L_2^{*\prime})$$

$$=\frac{1}{|m_1 m_2|}\left(\frac{f(\frac{m_1}{m_2}k)}{f(k)}\right)^{\frac{1}{4}}\left(\frac{L_2^*{}'-L_1^*{}'}{L_2^*-L_1^*}\right)\left(\frac{(L_1^*{}'-m_0\beta_0)(L_2^*{}'-m_0\beta_0)}{(L_1^*-\beta_0)(L_2^*-\beta_0)}\right)^{\frac{1}{2}}\frac{\left(\prod_{i=1}^{2}(2+e^{-L_i^*}+e^{L_i^*})\right)^{\frac{1}{2}}}{\left(\prod_{i=1}^{4}(2+e^{-L_i}+e^{L_i})\right)^{\frac{1}{4}}}$$

(5.3.3)

where the function f is jointly defined by (2.3.2) and (2.3.3). When $\beta_0 = 0$, the term $m_0\beta_0$ in (5.3.2) and (5.3.3) is replaced by β_0'. Of those parameters involved in the efficiency function F, an independent set of parameters consists of only β_0, k, m_0, m_1, and m_2. Among the other factors, L_1^* and L_2^* are determined by β_0 while $L_1^*{}'$ and $L_2^*{}'$ are determined by β_0 and m_0, though not through any algebraic functions.

Figure 5.3.1 illustrates an example of design shift due to parameter misspecification. Table 5.3.1 contains the relevant information concerning the parameters and the design points. The optimal design resides on the solid lines and the misplaced design on the dashed lines. The true parameters yield the optimal logits $L_1^* = -0.22$ and $L_2^* = 2.76$ as well as the optimal ray slopes 0.15 and 6.62. On the other hand, the misspecified parameters give the optimal logits $L_1^*{}' = -0.53$ and $L_2^*{}' = 2.48$ along with the optimal ray slopes 0.17 and 5.96, which are not optimal with respect to the true parameters. The misplaced design does lie on the rays with slopes 0.17 and 5.96 and meanwhile falls on the parallel lines corresponding to $L_1^*{}'$ and $L_2^*{}'$ defined by the incorrect parameters. However, the actual logits of the design points are defined by the true parameters and are listed in table 5.3.1. From figure 5.3.1, it can be seen that the ray slopes under misspecification are only off by a minimal amount. Distortion of the design is primarily caused by the incorrect ED lines. The D efficiency of the incorrect design relative to the correct one in this particular case is 47.48%.

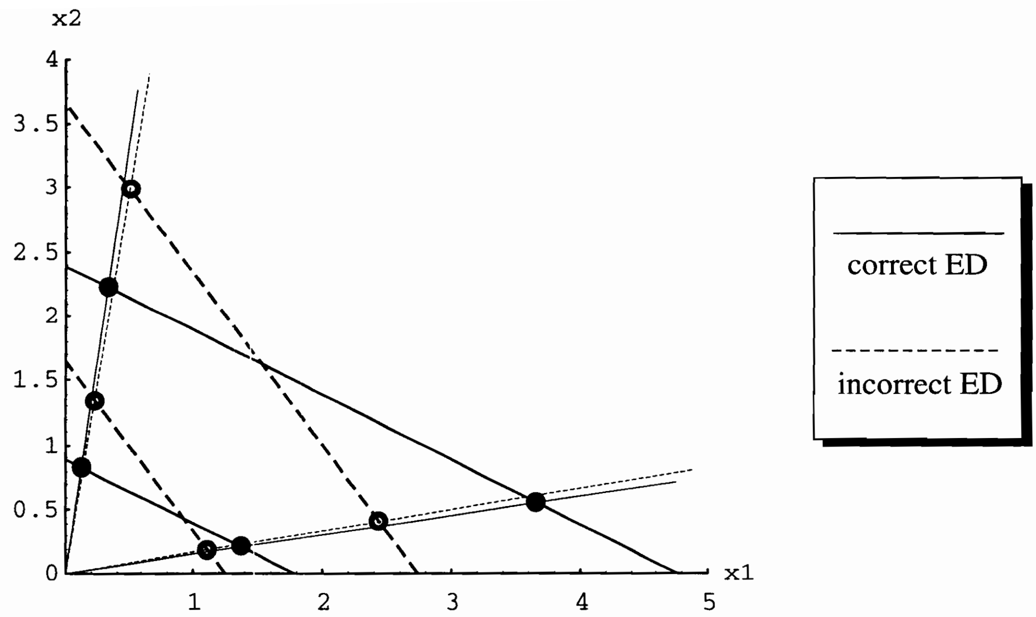


Figure 5.3.1 Design shift due to parameter misspecification, ray design

Table 5.3.1 Design shift due to parameter misspecification, ray design

Correct design ($\beta_0 = -2$, $\beta_1 = 1$, $\beta_2 = 2$)		Misplaced design ($\beta_0' = -3$, $\beta_1' = 2$, $\beta_2' = 1.5$)	
(x_1, x_2)	Logit	(x_1, x_2)	Logit
(0.13, 0.83)	-0.22	(0.23, 1.35)	0.92
(0.33, 2.21)	2.76	(0.50, 2.99)	4.48
(1.37, 0.21)	-0.22	(1.10, 0.18)	-0.53
(3.66, 0.55)	2.76	(2.44, 0.41)	1.25

The robustness property in a simplified situation assuming no misspecification in the slope k, or $m_1 = m_2 = m$, is first displayed in table 5.3.2 so that the effect of m_0 can be looked at more carefully. The efficiency function is free from k due to the constraint $m_1 = m_2$. Relatively speaking, a more optimistic type of misspecification seems to happen when both $|\beta_0|$ and the linear effects are moderately overestimated. The most detrimental circumstance would be to underestimate β_1 and β_2 while overestimating $|\beta_0|$, especially when $|\beta_0|$ is already large.

Table 5.3.2 D efficiency under parameter misspecification, ray design, $m_1 = m_2 = m$

β_0	m_0	m .25	.5	.75	1	1.25	1.5	1.75	2
0	0 (β_0')	.02	.47	.89	1	.94	.83	.72	.62
	-.25 (β_0')	.02	.44	.87	1.00	.96	.85	.74	.64
	-.5 (β_0')	.01	.40	.84	.99	.96	.87	.76	.66
	-.75 (β_0')	.01	.35	.80	.98	.97	.89	.78	.68
	-1 (β_0')	.01	.31	.75	.96	.97	.90	.80	.70
	-1.25(β_0')	.00	.26	.70	.93	.96	.91	.82	.72
	-1.5 (β_0')	.00	.21	.64	.89	.95	.91	.83	.74
	-2 (β_0')	.00	.14	.51	.79	.90	.90	.85	.77
-1	.25	.02	.53	.93	.97	.86	.71	.58	.47
	.5	.02	.50	.92	.99	.88	.74	.61	.49
	.75	.01	.45	.90	1.00	.90	.77	.63	.52
	1	.01	.40	.87	1	.93	.80	.67	.55
	1.25	.01	.35	.83	1.00	.95	.83	.70	.58
	1.5	.00	.29	.78	.98	.96	.85	.73	.61
	1.75	.00	.24	.72	.96	.97	.88	.76	.64
	2	.00	.19	.65	.92	.97	.89	.78	.67
-2	.25	.03	.60	.92	.86	.67	.50	.38	.28
	.5	.01	.53	.94	.92	.75	.58	.43	.33
	.75	.01	.42	.91	.98	.84	.66	.50	.39
	1	.00	.30	.83	1	.91	.74	.58	.45
	1.25	.00	.19	.71	.97	.96	.82	.66	.52
	1.5	.00	.11	.56	.90	.97	.87	.73	.59
	1.75	.00	.06	.40	.78	.93	.90	.79	.66
	2	.00	.03	.28	.64	.86	.90	.82	.71
-5	.25	.05	.58	.51	.23	.11	.05	.03	.02
	.5	.01	.59	.80	.48	.23	.12	.06	.04
	.75	.00	.27	.92	.82	.47	.25	.14	.08
	1	.00	.05	.63	1	.78	.47	.27	.16
	1.25	.00	.01	.24	.81	.96	.73	.46	.28
	1.5	.00	.00	.07	.44	.85	.89	.68	.45
	1.75	.00	.00	.02	.18	.56	.83	.81	.63
	2	.00	.00	.00	.06	.29	.61	.79	.74

Table 5.3.3 D efficiency under parameter misspecification, ray design, $\beta_0 = -1$

k	m_0	m_1	m_2 = .25	.5	.75	1	1.25	1.5	1.75	2
.3	.5	.25	.02	.21	.40	.48	.50	.48	.45	.42
		.5	.05	.50	.89	1.07	1.11	1.08	1.01	.95
		.75	.04	.52	.92	1.10	1.14	1.11	1.04	.97
		1	.07	.46	.82	.99	1.02	.99	.94	.87
		1.25	.13	.35	.70	.85	.88	.86	.81	.76
		1.5	.17	.21	.57	.72	.75	.74	.70	.65
		1.75	.20	.14	.45	.60	.64	.64	.61	.57
		2	.22	.26	.34	.50	.55	.55	.52	.49
	1	.25	.01	.15	.30	.38	.41	.40	.39	.36
		.5	.03	.40	.80	1.01	1.07	1.06	1.01	.95
		.75	.03	.45	.87	1.09	1.16	1.14	1.09	1.02
		1	.05	.41	.80	1	1.06	1.04	1.00	.94
		1.25	.10	.32	.69	.87	.93	.92	.88	.83
		1.5	.13	.20	.58	.75	.80	.80	.77	.72
		1.75	.16	.13	.46	.63	.69	.69	.67	.63
		2	.18	.25	.35	.53	.59	.60	.58	.55
	1.5	.25	.00	.09	.21	.28	.31	.31	.30	.29
		.5	.02	.29	.67	.90	.99	1.00	.97	.92
		.75	.02	.36	.78	1.04	1.14	1.15	1.12	1.06
		1	.03	.34	.74	.98	1.08	1.09	1.06	1.00
		1.25	.07	.28	.66	.88	.96	.97	.94	.90
		1.5	.10	.18	.56	.76	.84	.85	.83	.79
		1.75	.12	.12	.45	.65	.73	.75	.73	.69
		2	.13	.23	.34	.55	.63	.65	.64	.61
.7	.5	.25	.02	.25	.53	.69	.74	.74	.71	.67
		.5	.11	.50	1.08	1.42	1.55	1.57	1.52	1.44
		.75	.31	.26	.92	1.31	1.48	1.52	1.49	1.42
		1	.47	.76	.42	.99	1.20	1.26	1.26	1.22
		1.25	.57	1.00	.62	.58	.88	.99	1.02	1.00
		1.5	.60	1.12	.87	.32	.57	.74	.80	.80
		1.75	.61	1.16	.98	.61	.18	.51	.61	.64
		2	.59	1.16	1.02	.72	.38	.27	.44	.49
	1	.25	.01	.17	.40	.54	.61	.62	.60	.57
		.5	.07	.40	.96	1.32	1.49	1.53	1.51	1.44
		.75	.23	.23	.87	1.29	1.50	1.56	1.55	1.49
		1	.36	.70	.42	1	1.24	1.33	1.34	1.31
		1.25	.45	.94	.62	.60	.93	1.06	1.10	1.09
		1.5	.50	1.08	.88	.34	.61	.80	.87	.88
		1.75	.51	1.14	1.01	.65	.19	.55	.67	.70
		2	.50	1.15	1.06	.77	.42	.30	.48	.55
	1.5	.25	.00	.10	.27	.39	.45	.47	.47	.45
		.5	.04	.29	.79	1.16	1.36	1.44	1.44	1.39
		.75	.15	.18	.78	1.22	1.46	1.56	1.57	1.54
		1	.25	.60	.39	.98	1.26	1.37	1.41	1.39
		1.25	.33	.85	.60	.60	.96	1.12	1.17	1.17
		1.5	.37	.99	.87	.35	.64	.85	.94	.96
		1.75	.39	1.07	1.01	.67	.20	.60	.73	.78
		2	.39	1.09	1.08	.81	.45	.32	.53	.61

Table 5.3.4 D efficiency under parameter misspecification, ray design, β_0 = -2

			m_2							
k	m_0	m_1	.25	.5	.75	1	1.25	1.5	1.75	2
.3	.5	.25	.01	.21	.40	.46	.46	.43	.39	.35
		.5	.05	.53	.93	1.07	1.05	.98	.88	.79
		.75	.04	.55	.94	1.07	1.05	.97	.87	.78
		1	.07	.47	.81	.92	.91	.83	.75	.67
		1.25	.13	.35	.67	.77	.75	.69	.62	.56
		1.5	.16	.21	.53	.63	.62	.58	.52	.46
		1.75	.18	.13	.41	.51	.51	.48	.43	.39
		2	.19	.24	.30	.41	.43	.40	.37	.33
	1	.25	.00	.07	.19	.25	.27	.26	.25	.23
		.5	.01	.30	.71	.93	1.00	.97	.90	.83
		.75	.01	.38	.83	1.08	1.14	1.10	1.02	.93
		1	.03	.36	.78	1	1.05	1.01	.93	.85
		1.25	.06	.29	.68	.87	.91	.87	.80	.73
		1.5	.08	.18	.56	.73	.77	.74	.68	.62
		1.75	.10	.12	.44	.61	.65	.63	.58	.53
		2	.11	.22	.33	.50	.54	.53	.49	.45
	1.5	.25	.00	.02	.06	.09	.11	.12	.12	.11
		.5	.00	.11	.38	.61	.73	.77	.75	.71
		.75	.00	.17	.56	.87	1.03	1.07	1.04	.98
		1	.01	.19	.59	.90	1.05	1.08	1.05	.98
		1.25	.02	.17	.55	.83	.97	.99	.96	.89
		1.5	.03	.12	.48	.74	.85	.87	.84	.78
		1.75	.04	.08	.40	.63	.74	.76	.73	.68
		2	.04	.16	.31	.53	.63	.54	.63	.59
.7	.5	.25	.01	.25	.53	.67	.69	.66	.61	.55
		.5	.11	.53	1.13	1.42	1.49	1.43	1.33	1.21
		.75	.31	.37	.94	1.28	1.37	1.34	1.26	1.15
		1	.46	.77	.42	.92	1.07	1.07	1.02	.95
		1.25	.53	.98	.58	.52	.75	.80	.79	.74
		1.5	.55	1.05	.78	.28	.47	.58	.59	.57
		1.75	.53	1.04	.85	.50	.14	.38	.43	.44
		2	.50	1.00	.85	.57	.29	.20	.30	.33
	1	.25	.00	.08	.24	.35	.40	.40	.38	.36
		.5	.03	.30	.84	1.21	1.38	1.40	1.34	1.26
		.75	.13	.19	.83	1.28	1.47	1.50	1.45	1.37
		1	.23	.63	.41	1	1.23	1.29	1.26	1.19
		1.25	.29	.86	.61	.59	.91	1.01	1.01	.97
		1.5	.32	.98	.86	.33	.58	.74	.78	.77
		1.75	.32	1.02	.96	.61	.18	.50	.58	.59
		2	.31	1.01	.98	.71	.38	.26	.41	.45
	1.5	.25	.00	.02	.07	.13	.16	.18	.18	.17
		.5	.01	.11	.43	.76	.98	1.08	1.09	1.06
		.75	.03	.09	.56	1.01	1.30	1.43	1.45	1.42
		1	.07	.37	.32	.90	1.22	1.36	1.39	1.36
		1.25	.11	.57	.52	.58	.97	1.14	1.19	1.17
		1.5	.13	.71	.78	.34	.65	.87	.95	.96
		1.75	.14	.79	.92	.66	.21	.61	.73	.77
		2	.15	.81	.98	.79	.45	.32	.52	.59

Tables 5.3.3 and 5.3.4 evaluate cases of unrestricted misspecification. These tables show that higher D efficiencies occur with moderately overestimated β_2 and moderately underestimated β_1. Since the $k = \frac{\beta_1}{\beta_2} < 1$ case is currently addressed due to the duality between k > 1 and k < 1, it seems that slightly overestimating the larger coefficient while underestimating the smaller one might be relatively safer as far as preserving the efficiency is concerned. The performance can be poor when both β_1 and β_2 are underestimated, especially if $|\beta_0|$ is large and even over estimated. Otherwise, the efficiency is quite often at reasonable levels.

§5.4 Modified Ray Design

As in a regular ray design, the optimal ED's as well as the ray slope for a modified ray design are case-by-case solutions depending on the specific set of parameter values. Similar to the ray design case, it is assumed that the user implements the design according to the optimal ED's and ray slope conditioned on the specified parameters. The design matrix required to compute the design points can be given by the second and third columns of the model matrix in (2.6.2).

Let L_1^* and L_2^* denote the optimal logits and g the optimal ray slope due to the true parameters. Let $L_1^{*}{}'$, $L_2^{*}{}'$, and g' represent the corresponding attributes resulting from the specified parameters. The ratios m_0, m_1, m_2, and m_{12} defined in (5.1.1) and (5.2.1) will be used as measures of misspecification. Define a function S_m as

$$S_m(g', L) = \sqrt{\left(1 + \frac{m_1}{m_2} g'k\right)^2 + 4g'k \frac{m_{12}}{m_2^2} \frac{L - m_0\beta_0}{\Delta_\beta}} - \left(1 + \frac{m_1}{m_2} g'k\right) \qquad (5.4.1)$$

where $k = \dfrac{\beta_1}{\beta_2}$ and Δ_β was initially defined in (4.2.2) as $\Delta_\beta = \beta_0 - \beta_0^* = \dfrac{\beta_1\beta_2}{\beta_{12}}$, which represents the logit distance between the origin and the hyperbola center. The term $m_0\beta_0$ becomes β_0' when $\beta_0 = 0$. The actual logits of the design points obtained from the specified parameters are given by

$$L_1 = \beta_0(1-\frac{m_0}{m_{12}}) + \frac{L_1^{*\prime}}{m_{12}} + \frac{1}{2}\Delta_\beta\frac{m_2}{m_{12}}[(1-\frac{m_1}{m_{12}}) + \frac{g'}{k}(1-\frac{m_2}{m_{12}})]\, S_m(\frac{1}{g'}, L_1^{*\prime}),$$

$$L_2 = \beta_0(1-\frac{m_0}{m_{12}}) + \frac{L_2^{*\prime}}{m_{12}} + \frac{1}{2}\Delta_\beta\frac{m_2}{m_{12}}[(1-\frac{m_1}{m_{12}}) + \frac{g'}{k}(1-\frac{m_2}{m_{12}})]\, S_m(\frac{1}{g'}, L_2^{*\prime}),$$

$$L_3 = \beta_0(1-\frac{m_0}{m_{12}}) + \frac{L_1^{*\prime}}{m_{12}} + \frac{1}{2}\Delta_\beta\frac{m_2}{m_{12}}[(1-\frac{m_1}{m_{12}}) + \frac{1}{g'k}(1-\frac{m_2}{m_{12}})]\, S_m(g', L_1^{*\prime}),$$

$$L_4 = \beta_0(1-\frac{m_0}{m_{12}}) + \frac{L_2^{*\prime}}{m_{12}} + \frac{1}{2}\Delta_\beta\frac{m_2}{m_{12}}[(1-\frac{m_1}{m_{12}}) + \frac{1}{g'k}(1-\frac{m_2}{m_{12}})]\, S_m(g', L_2^{*\prime}),$$

(5.4.2)

where the term $\beta_0(1-\dfrac{m_0}{m_{12}})$ becomes $-\dfrac{\beta_0'}{m_{12}}$ when $\beta_0 = 0$.

The D efficiency of the design obtained from the specified parameters relative to the one from the true parameters is given by

$$\text{D-EFF} = F(\beta_0, \Delta_\beta, k, m_0, m_1, m_2, m_{12}, L_1^*, L_2^*, g, L_1^{*\prime}, L_2^{*\prime}, g')$$

$$= \left(\frac{m_2}{m_{12}}\right)^2 \left(\frac{[(g'^2-1)S_m(\frac{1}{g'}, L_1^{*\prime})S_m(\frac{1}{g'}, L_2^{*\prime}) + (\frac{1}{g'^2}-1)S_m(g', L_1^{*\prime})S_m(g', L_2^{*\prime})]}{[(g^2-1)S(\frac{1}{g}, L_1^*)S(\frac{1}{g}, L_2^*) + (\frac{1}{g^2}-1)S(g, L_1^*)S(g, L_2^*)]} \right.$$

$$\cdot\left(\frac{[S_m(g',L_2^*{}')-S_m(g',L_1^*{}')][S_m(\frac{1}{g'},L_2^*{}')-S_m(\frac{1}{g'},L_1^*{}')]}{[S(g,L_2^*)-S(g,L_1^*)][S(\frac{1}{g},L_2^*)-S(\frac{1}{g},L_1^*)]}\right)^{\frac{1}{2}}\frac{\left(\prod_{i=1}^{2}(2+e^{-L_i^*}+e^{L_i^*})\right)^{\frac{1}{2}}}{\left(\prod_{i=1}^{4}(2+e^{-L_i}+e^{L_i})\right)^{\frac{1}{4}}}$$

(5.4.3)

where the function S was initially defined in (2.6.1). An independent set of parameters in the D efficiency function F consists of β_0, β_0^*, k, m_0, m_1, m_2, and m_{12}. Among the other parameters, L_1^*, L_2^*, and g are determined by β_0, β_0^*, and k, while $L_1^*{}'$, $L_2^*{}'$, and g' are determined by the same along with the misspecification ratios.

Figure 5.4.1 illustrates an example of design shift due to parameter misspecification. The optimal design resides on the solid curves while the misplaced design on the dashed ones. Table 5.4.1 contains the relevant information concerning the parameters and the design points. The true parameters yield the optimal logits $L_1^* = -0.30$ and $L_2^* = 2.60$ as well as the optimal ray slopes 0.13 and 7.99. On the other hand, the misspecified parameters give the optimal logits $L_1^*{}' = 0.20$ and $L_2^*{}' = 3.06$ along with the optimal ray slopes 0.14 and 7.37, which are not optimal though with respect to the true parameters. The misplaced design does lie on the rays with slopes 0.14 and 7.37 and meanwhile falls on the parallel lines corresponding to $L_1^*{}'$ and $L_2^*{}'$ defined by the incorrect parameters. However, the actual logits of the design points are defined by the true parameters and are listed in table 5.4.1. Similar to what has been observed earlier in the ray designs, the rays here are affected very little by the parameter misspecification. The design is affected mostly by the wrong ED curves. The D efficiency of the incorrect design relative to the correct one in this particular case is 79.35%.

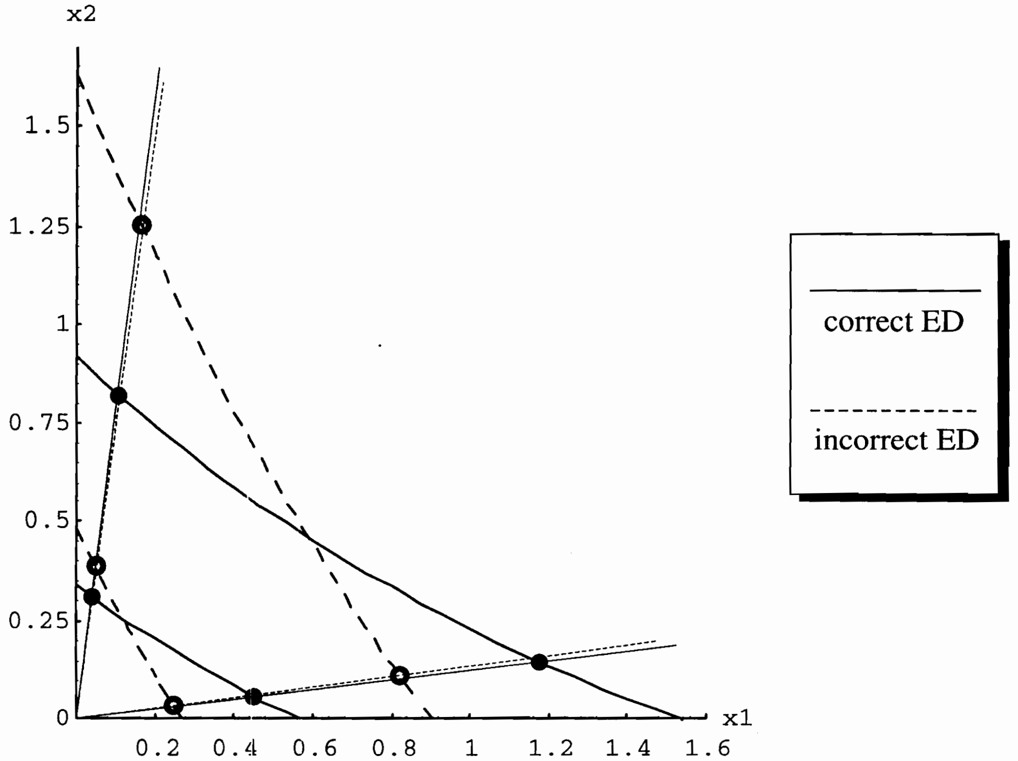


Figure 5.4.1 Design shift due to parameter misspecification, modified ray design

Table 5.4.1 Design shift due to parameter misspecification, modified ray design

Correct design ($\beta_0 = -2$, $\beta_1 = 3$, $\beta_2 = 5$, $\beta_{12} = 2$)		Misplaced design ($\beta_0' = -1$, $\beta_1' = 4.5$, $\beta_2' = 2.5$, $\beta_{12}' = 1$)	
(x_1, x_2)	Logit	(x_1, x_2)	Logit
(0.04, 0.31)	-0.30	(0.05, 0.38)	0.09
(0.10, 0.82)	2.60	(0.17, 1.24)	5.11
(0.45, 0.06)	-0.30	(0.25, 0.03)	-1.08
(1.17, 0.15)	2.60	(0.82, 0.11)	1.20

As the optimal solutions of a modified ray design depend on a number of model parameters jointly and the current study takes into account the wrong solutions due to each specific scenario of misspecification, the task is considerably increased on the computational side. Consequently, the numerical results are provided in a slightly reduced scope as compared to the previous sections.

The D efficiency under proportional misspecification ($m_0 = m_1 = m_2 = m_{12} = m$) as an over simplified case is first investigated in table 5.4.2. A general impression is that the designs appear slightly more robust to overestimation than to underestimation of the parameters. Recall that the modified ray design applies to synergistic situations where the center of the hyperbola ED's lies in the third quadrant. The efficiency slightly increases as Δ_β decreases. This means that the design tends to be slightly more robust as the synergistic interaction gets stronger, or the hyperbola center moves closer to the origin in terms of the logit distance. Another temperate increasing trend is found when the response probability at the zero doses becomes lower or β_0 gets negatively larger. Meanwhile the robustness property seems quite insensitive to the ratio k.

Table 5.4.2 D efficiency under parameter misspecification, modified ray design, $m_0 = m_1 = m_2 = m_{12} = m$

			m							
β_0	Δ_β	k	.25	.5	.75	1	1.25	1.5	1.75	2
-1	1	.3	.050	.566	.918	1	.958	.878	.792	.715
		.7	.050	.567	.919	1	.959	.878	.794	.716
	2	.3	.046	.557	.916	1	.957	.873	.785	.704
		.7	.047	.559	.916	1	.957	.874	.786	.706
-2	1	.3	.061	.596	.926	1	.963	.892	.817	.747
		.7	.061	.597	.927	1	.963	.893	.818	.748
	2	.3	.057	.588	.924	1	.962	.889	.811	.739
		.7	.058	.589	.924	1	.962	.889	.812	.741

Proportional misspecification except for β_0 ($m_1 = m_2 = m_{12} = m$) is evaluated in table 5.4.3. The efficiency appears to be relatively better when all the parameters are moderately overestimated as opposed to misspecification in opposite directions.

Table 5.4.3 D efficiency under parameter misspecification, modified ray design, $m_1 = m_2 = m_{12} = m$

β_0	Δ_β	k	m_0	m: .25	.5	.75	1	1.25	1.5	1.75	2
-1	1	.3	.5	.04	.57	.94	.99	.89	.76	.64	.54
			1	.02	.47	.89	1	.94	.82	.70	.59
			1.5	.01	.35	.80	.98	.97	.88	.76	.66
		.7	.5	.04	.57	.94	.99	.89	.76	.64	.54
			1	.02	.47	.89	1	.94	.82	.70	.60
			1.5	.01	.35	.80	.98	.97	.88	.77	.66
	2	.3	.5	.04	.56	.94	.99	.89	.76	.63	.53
			1	.02	.46	.89	1	.94	.82	.69	.58
			1.5	.01	.34	.80	.98	.97	.87	.76	.65
		.7	.5	.04	.56	.94	.99	.89	.76	.64	.53
			1	.02	.46	.89	1	.94	.82	.70	.59
			1.5	.01	.34	.80	.98	.97	.87	.76	.65
-2	1	.3	.5	.03	.60	.95	.92	.76	.60	.46	.37
			1	.00	.35	.85	1	.92	.77	.62	.49
			1.5	.00	.12	.57	.89	.97	.89	.77	.64
		.7	.5	.03	.60	.95	.92	.76	.60	.46	.37
			1	.00	.35	.85	1	.92	.77	.62	.50
			1.5	.00	.12	.57	.89	.97	.89	.77	.64
	2	.3	.5	.03	.59	.95	.92	.76	.59	.46	.36
			1	.00	.35	.85	1	.92	.76	.61	.49
			1.5	.00	.12	.57	.90	.97	.89	.76	.63
		.7	.5	.03	.59	.95	.92	.76	.59	.46	.36
			1	.00	.35	.85	1	.92	.76	.61	.49
			1.5	.00	.12	.57	.90	.97	.89	.76	.63

Table 5.4.4 examines a number of cases with less restricted misspecification. When β_0 and the linear coefficients are both underestimated, overestimating the interaction coefficient tends to result in higher efficiencies; when β_0 and the linear coefficients are both overestimated, underestimating the interaction coefficient tends to give better results. In general, the efficiency quite often achieves reasonable levels.

Table 5.4.4 D efficiency under parameter misspecification, modified ray design, $m_1 = m_2 = m$, k=0.5

				m_{12}							
β_0	Δ_β	m	m_0	.25	.5	.75	1	1.25	1.5	1.75	2
-1	1	.5	.5	.40	.57	.62	.64	.63	.62	.61	.59
			1	.31	.47	.52	.54	.54	.53	.52	.51
			1.5	.21	.35	.40	.42	.42	.42	.41	.40
		1	.5	.97	1.02	1.01	.99	.96	.93	.89	.87
			1	.93	1.01	1.02	1	.98	.95	.92	.89
			1.5	.86	.96	.99	.98	.96	.94	.92	.89
		1.5	.5	.96	.92	.87	.83	.80	.76	.73	.71
			1	1.00	.97	.93	.89	.86	.82	.79	.77
			1.5	1.02	1.01	.98	.94	.91	.88	.85	.82
	2	.5	.5	.42	.56	.62	.64	.64	.64	.63	.62
			1	.33	.46	.52	.54	.55	.55	.54	.53
			1.5	.23	.34	.39	.42	.43	.43	.43	.42
		1	.5	1.01	1.02	1.01	.99	.96	.94	.92	.89
			1	.99	1.01	1.01	1	.98	.96	.94	.92
			1.5	.93	.97	.99	.98	.97	.95	.94	.92
		1.5	.5	.91	.87	.84	.81	.78	.76	.74	.72
			1	.96	.93	.90	.87	.84	.82	.80	.77
			1.5	1.00	.98	.95	.92	.90	.87	.85	.83
-2	1	.5	.5	.41	.60	.65	.67	.66	.65	.63	.61
			1	.21	.35	.41	.43	.44	.43	.43	.42
			1.5	.06	.12	.16	.17	.18	.18	.18	.18
		1	.5	.96	.99	.96	.92	.88	.84	.81	.77
			1	.89	1.00	1.02	1	.97	.94	.91	.88
			1.5	.64	.81	.88	.89	.89	.88	.86	.84
		1.5	.5	.84	.78	.72	.67	.63	.60	.57	.54
			1	.98	.94	.89	.85	.80	.77	.73	.70
			1.5	1.00	1.01	.99	.96	.93	.89	.86	.83
	2	.5	.5	.44	.59	.65	.67	.67	.66	.65	.64
			1	.22	.35	.41	.43	.44	.44	.44	.43
			1.5	.07	.12	.16	.17	.18	.18	.19	.18
		1	.5	.99	.98	.96	.92	.89	.86	.83	.80
			1	.97	1.01	1.01	1	.98	.95	.93	.91
			1.5	.75	.84	.88	.90	.89	.89	.87	.86
		1.5	.5	.77	.73	.69	.65	.62	.59	.57	.55
			1	.94	.90	.86	.82	.79	.76	.73	.71
			1.5	1.01	1.00	.97	.94	.92	.90	.86	.84

In this chapter we investigated the robustness property under parameter misspecification for several types of D optimal designs, including the parallel-line, hyperbola-based, ray, and modified ray designs. The study shows that these designs all have reasonable efficiencies and do not appear to be severely affected when subject to moderate misspecification of the parameters. Still, driven by the conventional optimization criterion, these designs are intrinsically designated to be most competitive only for the set of parameter values used to construct the design, namely when the researcher knows the true parameters. As a result, the efficiency tends to decrease as the parameter guesses deviate from the true values. Conceptually, to take into account the uncertainty involved in the parameter knowledge, one could adopt the Bayesian design approach, which addresses various possible parameter values in a systematic way.

CHAPTER 6

BAYESIAN DESIGNS

§6.1 Bayesian Design Procedure

If the guessed parameters used in constructing an optimal design differ from the true values, the constructed design will be different and less efficient than the design which makes use of "true" parameters. Such behavior of the design, referred to as its robustness to parameter misspecification, was studied in the previous chapter. In fact, the designs discussed earlier in this dissertation are from the very beginning driven by such criteria that render the design to be the best only for the single set of parameter values that are used in constructing the design. The following idea then arises: improving the design robustness begins with using a criterion that takes into account various possible values of the parameters, not merely the construction values. To gear a design to various possible parameter values, one relevant approach is to use an overall criterion that is a "weighted average" of the individual criterion values over a range of parameter values surrounding the construction values. The "weight" represents the likelihood of the true parameter values falling on each individual case. All the criteria addressed earlier in this work put all the weight only on the construction values. Conceptually, distributing the weight should yield a design which is less good for the construction values but improved for other values. Such a design possesses a balance in its ability to cope with various possible states of nature.

The above description is indeed the core rationale of the Bayesian design procedure. For the one-variable logistic model, a Bayesian procedure was developed by

Chaloner and Larntz (1989) which utilizes a density function to represent the experimenter's *a priori* belief about the parameters. In general, the Bayesian design procedure is given by the following expression

$$\underset{\mathcal{D}}{\text{Min}} \int_{\theta} R(\delta, \theta)\, \pi(\theta)\, d\theta \qquad (6.1.1)$$

where θ is a vector of model parameters, π(θ) is the prior density of θ, R(δ, θ) is the expression for any design optimality criterion of choice, and δ is any design from the set $\mathcal{D}$ of candidate designs. Expression 6.1.1 comes from Bayesian decision theory where R(δ,θ) represents the risk. Unlike Bayesian estimation, R(δ, θ) implicitly contains a fixed estimator. The Bayesian optimal design minimizes the expected risk with respect to the designs rather than estimators.

The Bayesian D optimal criterion is given by

$$\underset{\mathcal{D}}{\text{Max}} \int_{\beta} |I(\beta)|\, \pi(\beta)\, d\beta \qquad (6.1.2)$$

where β is the vector of model parameters. The experimenter would need to select a prior distribution π(β) which reflects his or her knowledge about the parameters. In this dissertation, two common situations are investigated: normal and uniform prior distributions. In both cases, the joint distribution assumes no correlation. The design scenario considered is the parallel-line design for the no-interaction model. The designs are found using a numerical expectation and the Nelder-Mead algorithm which does function minimization through a simplex method. The integration is approximated by a 30 point segmentation of the density in the uniform case and a 30 point segmentation over the ±3 standard deviations in the normal case. A 40 point segmentation was initially explored within the ±4 standard deviations for the normal prior. Negligible differences were found between the 30 and 40 point approximations. Given the three-dimensional parameter space for the two-variable model, even the 30 point approach consumes very large amounts of CPU time. Consequently, the 30 point segmentation is used merely to avoid unnecessary computational intensity.

§6.2 Normal Prior Distribution

The normal prior distribution of the parameters is represented as follows:

$$\begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\end{bmatrix} \sim N\left(\begin{bmatrix}\mu_0\\ \mu_1\\ \mu_2\end{bmatrix}, \begin{bmatrix}\tau_0^2 & 0 & 0\\ 0 & \tau_1^2 & 0\\ 0 & 0 & \tau_2^2\end{bmatrix}\right). \qquad (6.2.1)$$

Initially a six-point parallel-line design is assumed to be constructed on three ED lines defined by the mean of the prior distribution with the lower and upper logits $L_1 = \mu_0 + \mu_1x_1 + \mu_2x_2$ and $L_2 = \mu_0 + \mu_1x_1 + \mu_2x_2$ and the middle logit being $(L_1 + L_2)/2$. The sample proportions are denoted r_1, r_2, and r_3 on the lower, middle, and upper ED's respectively where $r_1 + r_2 + r_3 = 1$ and the two points on the same ED always have equal sample sizes. It would be more convenient to reexpress the design logits in terms of a midpoint logit S and a shift T below and above the midpoint, or

$$S = \frac{L_1 + L_2}{2} \qquad \text{and} \qquad T = \frac{L_2 - L_1}{2}. \qquad (6.2.2)$$

Thus the design logits are given by S - T, S, and S + T. A symmetric design about ED_{50} is indicated by S = 0. Let L_i, i = 1, ..., 6, denote the actual logits of the design points when the parameters falls on a particular point $(\beta_0, \beta_1, \beta_2)'$ of the prior distribution, where points 1 and 4 reside on the lower ED, 2 and 5 on the middle ED, and 3 and 6 on the upper ED. These actual logits are given by

$$L_1 = \mu_0(\alpha_0 - \alpha_2) + \alpha_2 S + (L_X - L_D)(\alpha_1 - \alpha_2) - T\,\frac{\alpha_1 k^2 + \alpha_2}{k^2 + 1},$$

$$L_2 = \mu_0(\alpha_0 - \alpha_2) + \alpha_2 S + (L_X - L_D)(\alpha_1 - \alpha_2),$$

$$L_3 = \mu_0(\alpha_0 - \alpha_2) + \alpha_2 S + (L_X - L_D)(\alpha_1 - \alpha_2) + T\,\frac{\alpha_1 k^2 + \alpha_2}{k^2+1},$$

$$L_4 = \mu_0(\alpha_0 - \alpha_2) + \alpha_2 S + (L_X + L_D)(\alpha_1 - \alpha_2) - T\,\frac{\alpha_1 k^2 + \alpha_2}{k^2+1},$$

$$L_5 = \mu_0(\alpha_0 - \alpha_2) + \alpha_2 S + (L_X + L_D)(\alpha_1 - \alpha_2),$$

$$L_6 = \mu_0(\alpha_0 - \alpha_2) + \alpha_2 S + (L_X + L_D)(\alpha_1 - \alpha_2) + T\,\frac{\alpha_1 k^2 + \alpha_2}{k^2+1},$$

(6.2.3)

where $\mu_0 \neq 0$, $\alpha_0 = \frac{\beta_0}{\mu_0}$, $\alpha_1 = \frac{\beta_1}{\mu_1}$, $\alpha_2 = \frac{\beta_2}{\mu_2}$, $L_x = \mu_1 x_{10}$, $L_D = \mu_1 D_1$, and $k = \frac{\mu_1}{\mu_2}$, where x_{10} and D_1 are indices of the design center and length pertaining to the variable x_1 as illustrated in figure 2.1.1. The logit type quantities L_x and L_D serve as scale free indicators of the location and size of the design, more details of which were discussed in §4.3 and §5.1. When $\mu_0 = 0$, the term $\mu_0(\alpha_0 - \alpha_2)$ in the logit expression in (6.2.3) should be replaced by β_0.

Let $P_i = \frac{1}{1+e^{-L_i}}$ and $W_i = P_i(1 - P_i)$, i = 1, ..., 6. The Bayesian D optimal criterion can be expressed as

$$\frac{N^3}{\mu_1^2\mu_2^2}\frac{(L_D T)^2}{2}\cdot\int_{\alpha_2}\int_{\alpha_1}\int_{\alpha_0}[(r_1W_1 + r_2W_2 + r_3W_3)(4r_1r_3W_4W_6 + r_2W_5(r_1W_4 + r_3W_6))$$

$$+ (r_1W_4 + r_2W_5 + r_3W_6)(4r_1r_3W_1W_3 + r_2W_2(r_1W_1 + r_3W_3))]\, f(\alpha_0, \alpha_1, \alpha_2)\, d\alpha_0 d\alpha_1 d\alpha_2$$

(6.2.4)

where $f(\alpha_0, \alpha_1, \alpha_2)$ is the density corresponding to

$$\begin{bmatrix}\alpha_0\\ \alpha_1\\ \alpha_2\end{bmatrix} \sim N\left(\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}\gamma_0^2 & 0 & 0\\ 0 & \gamma_1^2 & 0\\ 0 & 0 & \gamma_2^2\end{bmatrix}\right). \qquad (6.2.5)$$

The standard deviation γ_j's are indeed the coefficients of variation in the original prior, i.e.

$$\gamma_j = \frac{\tau_j}{|\mu_j|}, \qquad j = 0, 1, 2. \tag{6.2.6}$$

When $\mu_0 = 0$, α_0 is not defined and it is assumed $\beta_0 \sim N(0, \gamma_0^2)$ where $\gamma_0 = \tau_0$.

Optimization of the Bayesian criterion in (6.2.4) appears to be dependent on μ_0, k, γ_0, γ_1, γ_2, L_x, L_D, S, T, r_1, r_2, and r_3. Further evaluation of the criterion suggests that the criterion is monotonically increasing with the scale free design length L_D. The Bayesian designs will be found by optimizing S, T, r_1, r_2, and r_3 for given values of μ_0, k, γ_0, γ_1, γ_2, L_x, and L_D, which must be supplied by the user in real applications. Among the user specifications, L_x and L_D represent the location and length of the design while μ_0, k, γ_0, γ_1, and γ_2 constitute a form of prior knowledge about the parameters.

Regarding μ_0, only the case $\mu_0 \leq 0$ will be evaluated as the criterion expression possesses a symmetry in the sense that $-\mu_0$ would yield the same criterion value as μ_0 if L_x and S also switch sign and r_1 and r_3 switch position. As to k where $k = \frac{\mu_1}{\mu_2}$, evaluating $|k| \leq 1$ would be sufficient since each $k > 1$ case corresponds to a $k < 1$ case if the two design variables switch position.

The coefficients of variation γ_0, γ_1, and γ_2 quantify the degree of uncertainty in the prior knowledge. Larger values of γ_j's indicate less precise knowledge while $\gamma_j = 0$ means that a point value is supplied for β_j as opposed to a nontrivial prior distribution. If the Bayesian design criterion in (6.2.4) involves integration over a range of the prior distribution which extends to both the negative and the positive regions, it can be shown that maximization of the Bayesian criterion can possibly require that the logit deviation T go to infinity. To ensure that a Bayesian design does exist, the integrated region of α_j must remain within the positive side, or $3\gamma_j < 1$. This requirement corresponds to the condition that within ±3 standard deviations from the mean of the prior, the coefficient does not switch sign, or the region $[\mu_j - 3\tau_j,\ \mu_j + 3\tau_j]$ does not contain zero. In practice, the experimenter may have a fairly wide range of ±3 standard deviations for a model effect due to lack of knowledge about its exact magnitude. However, it is not unreasonable to assume that the experimenter has some idea regarding the nature of the effect and hence is at least certain about the sign of the coefficient. Due to the above concern, which implies the assumption $3\gamma_j < 1$, the coefficient of variation γ_j is chosen to be evaluated at levels 0, 0.15, and 0.3.

Tables 6.2.1 and 6.2.2 give selected Bayesian D optimal designs for the cases $\mu_0 = 0$ and $\mu_0 = -2$ respectively. Even though the initial design layout assumes a three-line design, the optimal sample proportion r_2 for the middle ED line turns out to be zero for all the cases evaluated here. Consequently, all the Bayesian optimal designs end up to be two-line designs for the normal prior distributions considered here. In addition, the case $\gamma_0 = \gamma_1 = \gamma_2 = 0$, regardless of the values for other factors such as μ_0, k, L_x, L_D, and the type of prior, always yields the same non-Bayesian design which agrees with the results of parallel-line D optimal design obtained in §2.1.

For the cases in table 6.2.1 where $\mu_0 = 0$, the Bayesian designs all have equal sample sizes, indicated by $r_1 = r_3 = 0.5$, and symmetric ED's, indicated by the optimal central logit S equal to zero. The logit deviation T always increases as the uncertainty of parameter knowledge becomes larger, or γ_j increases. As a result, the design becomes

Table 6.2.1 Bayesian designs, normal priors, $\mu_0 = 0$, $|k| = 0.5$, $L_x = 0$, $L_D = 2$

γ_0	γ_1	γ_2	S	T	P_1	P_2	P_3	r_1	r_2	r_3	Criterion
0	0	0	0	1.223	.227	----	.773	.5	0	.5	.0324
0	0	.15	0	1.266	.219	----	.781	.5	0	.5	.0318
0	0	.3	0	1.360	.204	----	.796	.5	0	.5	.0298
0	.15	0	0	1.237	.225	----	.775	.5	0	.5	.0314
0	.15	.15	0	1.286	.217	----	.783	.5	0	.5	.0308
0	.15	.3	0	1.385	.200	----	.800	.5	0	.5	.0292
0	.3	0	0	1.281	.217	----	.783	.5	0	.5	.0288
0	.3	.15	0	1.325	.210	----	.790	.5	0	.5	.0285
0	.3	.3	0	1.448	.190	----	.810	.5	0	.5	.0278
.15	0	0	0	1.228	.227	----	.773	.5	0	.5	.0321
.15	0	.15	0	1.270	.219	----	.781	.5	0	.5	.0314
.15	0	.3	0	1.364	.204	----	.796	.5	0	.5	.0295
.15	.15	0	0	1.245	.224	----	.776	.5	0	.5	.0310
.15	.15	.15	0	1.288	.216	----	.784	.5	0	.5	.0305
.15	.15	.3	0	1.389	.200	----	.800	.5	0	.5	.0289
.15	.3	0	0	1.284	.217	----	.783	.5	0	.5	.0285
.15	.3	.15	0	1.330	.209	----	.791	.5	0	.5	.0283
.15	.3	.3	0	1.449	.190	----	.810	.5	0	.5	.0275
.3	0	0	0	1.242	.224	----	.776	.5	0	.5	.0313
.3	0	.15	0	1.282	.217	----	.783	.5	0	.5	.0307
.3	0	.3	0	1.370	.203	----	.797	.5	0	.5	.0289
.3	.15	0	0	1.255	.222	----	.778	.5	0	.5	.0303
.3	.15	.15	0	1.301	.214	----	.786	.5	0	.5	.0298
.3	.15	.3	0	1.402	.197	----	.803	.5	0	.5	.0283
.3	.3	0	0	1.296	.215	----	.785	.5	0	.5	.0279
.3	.3	.15	0	1.341	.207	----	.793	.5	0	.5	.0277
.3	.3	.3	0	1.458	.189	----	.811	.5	0	.5	.0269

Table 6.2.2 Bayesian designs, normal priors, $\mu_0 = -2$, $|k| = 0.5$, $L_x = 1$, $L_D = 2$

γ_0	γ_1	γ_2	S	T	P_1	P_2	P_3	r_1	r_2	r_3	Criterion
0	0	0	0	1.223	.227	----	.773	.5	0	.5	.0324
0	0	.15	.0752	1.265	.233	----	.793	.515	0	.485	.0315
0	0	.3	.1226	1.372	.223	----	.817	.506	0	.494	.0292
0	.15	0	.0157	1.245	.226	----	.779	.501	0	.499	.0311
0	.15	.15	.0681	1.295	.227	----	.796	.506	0	.494	.0303
0	.15	.3	.1764	1.404	.227	----	.829	.511	0	.489	.0283
0	.3	0	.0451	1.289	.224	----	.792	.504	0	.496	.0282
0	.3	.15	.1199	1.343	.227	----	.812	.510	0	.490	.0277
0	.3	.3	.2926	1.481	.234	----	.855	.519	0	.481	.0262
.15	0	0	0	1.239	.225	----	.775	.5	0	.5	.0313
.15	0	.15	.0381	1.284	.223	----	.789	.502	0	.498	.0305
.15	0	.3	.1135	1.387	.219	----	.818	.505	0	.495	.0283
.15	.15	0	.0127	1.260	.223	----	.781	.502	0	.498	.0301
.15	.15	.15	.0566	1.309	.222	----	.797	.503	0	.497	.0294
.15	.15	.3	.1631	1.426	.220	----	.831	.508	0	.492	.0274
.15	.3	0	.0614	1.310	.223	----	.798	.515	0	.485	.0273
.15	.3	.15	.1141	1.343	.226	----	.811	.506	0	.494	.0268
.15	.3	.3	.2725	1.491	.228	----	.854	.516	0	.484	.0255
.3	0	0	0	1.276	.218	----	.782	.5	0	.5	.0287
.3	0	.15	.0379	1.323	.217	----	.796	.506	0	.494	.0281
.3	0	.3	.0943	1.424	.209	----	.820	.503	0	.497	.0262
.3	.15	0	.0234	1.293	.219	----	.789	.506	0	.494	.0277
.3	.15	.15	.0509	1.342	.216	----	.801	.503	0	.497	.0271
.3	.15	.3	.1166	1.446	.209	----	.827	.502	0	.498	.0255
.3	.3	0	.0370	1.337	.214	----	.798	.502	0	.498	.0253
.3	.3	.15	.1120	1.385	.219	----	.817	.509	0	.491	.0249
.3	.3	.3	.2318	1.520	.216	----	.852	.510	0	.490	.0238

more spread out as the prior information gets vaguer with larger dispersion. This does agree with the intuition behind the Bayesian rationale.

When the prior mean μ_0 deviates from zero, as seen in the cases of table 6.2.2 where $\mu_0 = -2$, the two symmetric properties, i.e. equal sample sizes and symmetric ED's, are still concurrent with each other, but only occur when the prior distributions for both linear effects β_1 and β_2 are degenerated, or point priors. Otherwise, the optimal central logit S is found to be shifted always to the opposite direction from μ_0 while the optimal sample allocation is always weighted towards the same direction with μ_0, or weighted at the ED closer to μ_0. Currently given $\mu_0 < 0$, it is observed that $S > 0$, $r_1 > 0.5$, and $r_3 < 0.5$ for any case where γ_1 and γ_2 are not all zero. This implies that the design ED's fall more toward the upper probability tail with a greater sample weight at the lower ED. The shift in the ED's becomes more obvious as the prior knowledge about the linear effects involves more uncertainty, indicated by larger values of γ_1 and γ_2. However, the asymmetry in either the ED's or the sample weights does not seem to be severe in any of the cases and often appears moderate or slight. The logit deviation T is once again found to be increasing with γ_j, suggesting a wider design when the prior knowledge is vaguer.

The Bayesian D optimality criterion, apart from the factor $\frac{N^3}{\mu_1^2\mu_2^2}$, is given for each design in the tables. Inspection of these criterion values indicates that the Bayesian D optimality criterion decreases as the uncertainty in the parameter knowledge increases, characterized by larger values of γ_j. Meanwhile the Bayesian criterion also becomes worse as the prior mean μ_0 deviates from zero and thereby the design is forced to lose its symmetry. Nevertheless, the change in the criterion value never exhibits any large momentum, and the Bayesian design criterion still seems quite stable.

Bayesian designs are found for other values of μ_0, k, L_x, and L_D. No additional systematic features are observed, except that similar to the non-Bayesian situation, the Bayesian D optimality criterion increases as L_D becomes larger, representing a greater design length in the ED direction.

§6.3 Uniform Prior Distribution

The uniform prior distribution is assumed to take the form

$$\beta_j \sim \text{Uniform}(\mu_j - 3\delta_j,\ \mu_j + 3\delta_j),$$
$$\text{Cov}(\beta_j, \beta_{j'}) = 0,$$
$$j, j' = 0, 1, 2, \quad j \neq j'. \qquad (6.3.1)$$

Expressing the half range of the uniform distribution as $3\delta_j$ is meant to create an analogy to the normal prior case where the numerical integration extends to ±3 standard deviations from the mean. For the uniform prior in (6.3.1), the standard deviation of β_j is indeed $\sqrt{3}\,\delta_j$. Except γ_j's, all other parameters defined in the normal case can be used here with the uniform prior. The analogous γ_j's with the uniform prior given in the form of (6.3.1) would be defined by

$$\gamma_j = \frac{\delta_j}{|\mu_j|}, \qquad j = 0, 1, 2. \qquad (6.3.2)$$

It follows that

$$\alpha_j \sim \text{Uniform}(1 - 3\gamma_j,\ 1 + 3\gamma_j), \qquad j = 0, 1, 2, \qquad (6.3.3)$$

where $\alpha_j = \dfrac{\beta_j}{\mu_j}$ and the standard deviation of α_j is $\sqrt{3}\,\gamma_j$.

Similar to the normal prior case, a three-line design is initially assumed on the logits S - T, S, and S + T. Again, the notion of logit here does not represent the actual logit but rather a linear combination of the design variables via the mean of the prior, or $\mu_0 + \mu_1 x_1 + \mu_2 x_2$, which serves as a geometric characterization of the design.

The Bayesian D optimality criterion given the uniform prior takes the same form as in (6.2.4) for the normal prior except that the density $f(\alpha_0, \alpha_1, \alpha_2)$ is based on the uniform distribution in (6.3.3), or $f(\alpha_0, \alpha_1, \alpha_2) = \dfrac{1}{216\gamma_0\gamma_1\gamma_2}$. The design indicated by S, T, r_1, r_2, and r_3 is found for given values of μ_0, k, γ_0, γ_1, γ_2, L_X, and L_D by maximizing the Bayesian D criterion.

Similar to the situation of normal priors discussed in the previous section, selected Bayesian D optimal designs are found and given in tables 6.3.1 and 6.3.2 for the cases $\mu_0 = 0$ and $\mu_0 = -2$ respectively. Unlike the normal situation where the optimal designs in the selected cases are all found to have two levels, solutions under a uniform prior often leads to a three-line design. The optimal number of ED lines changes from two to three when the prior distribution becomes more spread out, or γ_j increases. Uniform priors lead to designs with more levels than normal priors due to the relatively heavier weights assigned to the extreme values by a uniform distribution.

For $\mu_0 = 0$, as shown in table 6.3.1, symmetric ED's, indicated by S = 0, is always accompanied by symmetric sample proportions, or $r_1 = r_3$. Whenever $S \neq 0$, there exist dual solutions. If the two solutions are denoted as $\{S, T, r_1, r_3\}$ and $\{S', T', r_1', r_3'\}$, they satisfy $S = -S'$, $T = T'$, $r_1 = r_3'$, and $r_3 = r_1'$. Dual solutions tend to give three-line designs and often occur when the prior distribution is wide. The cases with dual solutions are shaded in table 6.3.1.

Table 6.3.1 Bayesian designs, uniform priors, $\mu_0 = 0$, $|k| = 0.5$, $L_x = 0$, $L_D = 2$

γ_0	γ_1	γ_2	S	T	P_1	P_2	P_3	r_1	r_2	r_3	Criterion
0	0	0	0	1.223	.227	----	.773	.5	0	.5	.0324
0	0	.15	0	1.366	.203	----	.797	.5	0	.5	.0307
0	0	.3	-1.158	3.105	.014	.239	.875	.173	.389	.438	.0258
			1.158	3.105	.125	.761	.986	.438	.389	.173	.0258
0	.15	0	0	1.276	.218	----	.782	.5	0	.5	.0295
0	.15	.15	0	1.426	.194	----	.806	.5	0	.5	.0289
0	.15	.3	-.9132	3.525	.012	.286	.932	.232	.365	.403	.0284
			.9132	3.525	.068	.714	.988	.403	.365	.232	.0284
0	.3	0	0	1.390	.199	----	.801	.5	0	.5	.0230
0	.3	.15	0	1.724	.151	.5	.849	.437	.126	.437	.0243
0	.3	.3	0	4.709	.009	.5	.991	.354	.292	.354	.0365
.15	0	0	0	1.236	.225	----	.775	.5	0	.5	.0316
.15	0	.15	-.0561	1.371	.194	----	.788	.482	0	.518	.0300
			.0561	1.371	.212	----	.806	.518	0	.482	.0300
.15	0	.3	-1.182	3.133	.013	.235	.876	.174	.386	.440	.0254
			1.182	3.133	.124	.765	.987	.440	.386	.174	.0254
.15	.15	0	0	1.290	.216	----	.784	.5	0	.5	.0288
.15	.15	.15	0	1.445	.191	----	.809	.5	0	.5	.0283
.15	.15	.3	-.8954	3.531	.012	.290	.933	.238	.360	.402	.0279
			.8954	3.531	.067	.710	.988	.402	.360	.238	.0279
.15	.3	0	0	1.394	.199	----	.801	.5	0	.5	.0226
.15	.3	.15	0	1.734	.150	.5	.850	.438	.124	.438	.0239
.15	.3	.3	0	4.751	.009	.5	.991	.353	.294	.353	.0357
.3	0	0	0	1.272	.219	----	.781	.5	0	.5	.0294
.3	0	.15	0	1.413	.196	----	.804	.5	0	.5	.0281
.3	0	.3	-1.204	3.207	.012	.231	.881	.182	.379	.439	.0242
			1.204	3.207	.119	.769	.988	.439	.379	.182	.0242
.3	.15	0	0	1.335	.208	----	.792	.5	0	.5	.0270
.3	.15	.15	0	1.476	.186	----	.814	.5	0	.5	.0266
.3	.15	.3	-.7758	3.535	.013	.315	.940	.258	.347	.395	.0265
			.7758	3.535	.060	.685	.987	.395	.347	.258	.0265
.3	.3	0	0	1.424	.194	----	.804	.5	0	.5	.0215
.3	.3	.15	0	1.777	.145	.5	.855	.437	.126	.437	.0226
.3	.3	.3	0	4.873	.008	.5	.992	.350	.300	.350	.0337

Table 6.3.2 Bayesian designs, uniform priors, μ_0 = -2, |k| = 0.5, L_x = 1, L_D = 2

γ_0	γ_1	γ_2	S	T	P_1	P_2	P_3	r_1	r_2	r_3	Criterion
0	0	0	0	1.223	.227	----	.773	.5	0	.5	.0324
0	0	.15	.1426	1.377	.225	----	.820	.509	0	.491	.0301
0	0	.3	1.670	3.063	.199	.842	.991	.445	.329	.226	.0250
0	.15	0	.0475	1.286	.225	----	.791	.504	0	.496	.0289
0	.15	.15	.3053	1.479	.236	----	.856	.520	0	.480	.0277
0	.15	.3	2.386	3.611	.227	.916	.998	.442	.296	.262	.0264
0	.3	0	.1072	1.396	.216	----	.818	.508	0	.492	.0221
0	.3	.15	.5957	1.803	.230	.645	.917	.482	.107	.411	.0226
0	.3	.3	5.164	5.606	.391	.994	1.00	.429	.276	.295	.0314
.15	0	0	0	1.272	.219	----	.781	.5	0	.5	.0294
.15	0	.15	.1097	1.421	.212	----	.822	.504	0	.496	.0276
.15	0	.3	1.583	3.124	.176	.830	.991	.433	.328	.239	.0235
.15	.15	0	.0403	1.330	.216	----	.797	.503	0	.497	.0265
.15	.15	.15	.2383	1.505	.220	----	.851	.512	0	.488	.0256
.15	.15	.3	2.280	3.670	.199	.907	.997	.432	.302	.266	.0249
.15	.3	0	.0946	1.439	.207	----	.823	.513	0	.487	.0206
.15	.3	.15	.5393	1.906	.203	.632	.920	.451	.160	.389	.0211
.15	.3	.3	4.846	5.558	.329	.992	1.00	.425	.281	.294	.0298
.3	0	0	0	1.367	.203	----	.797	.5	0	.5	.0230
.3	0	.15	.0446	1.516	.187	----	.826	.500	0	.500	.0221
.3	0	.3	1.269	3.272	.119	.781	.989	.402	.327	.271	.0201
.3	.15	0	.0247	1.418	.199	----	.809	.501	0	.499	.0212
.3	.15	.15	.1275	1.669	.176	.532	.858	.466	.077	.457	.0208
.3	.15	.3	1.868	3.805	.126	.866	.997	.404	.317	.279	.0215
.3	.3	0	.0807	1.520	.192	----	.832	.503	0	.497	.0171
.3	.3	.15	.3975	2.117	.152	.598	.925	.402	.237	.361	.0178
.3	.3	.3	3.994	5.519	.179	.982	1.00	.407	.297	.296	.0258

When μ_0 is other than zero, uniform priors lead to design patterns similar to the normal prior case. As shown in table 6.3.2 where $\mu_0 = -2$, the design has symmetric ED's, or S = 0, and equal sample sizes only when both linear effects come with point priors, or $\gamma_1 = \gamma_2 = 0$. Otherwise the design ED's and sample weights both turn out to be unbalanced in the same fashion as found in the normal case, which is explained in §6.2. The only additional feature here is that for a uniform prior, the result is often a three- rather than two-line design.

The Bayesian optimality criterion, apart from the factor $\frac{N^3}{\mu_1^2 \mu_2^2}$, is given for each design in the tables. Similar to the normal case, the criterion decreases as the prior becomes more dispersed or as μ_0 deviates from zero. However, an increase in the criterion is often achieved whenever the optimal design increases from two to three levels due to any increase in the dispersion of the prior, represented by an increase in γ_j. If the increase in γ_j does not lead to an additional design level, then the criterion will decrease.

The Bayesian D optimality criterion in the uniform case tend to be smaller than that in the normal case when both types of priors lead to two-level designs. In the cases where the uniform prior leads to a three-level design whereas the normal prior gives a two-level solution, the criterion with the uniform prior is usually greater than that with the normal prior.

In general, compared to normal priors, uniform priors tend to give designs with more ED levels and with larger ED dispersions.

§6.4 Robustness to Parameter Misspecification

The experimenter is assumed to make guesses of the parameters through prior distributions. The Bayesian optimal design is found for the given prior distribution and constructed using the mean of the prior $(\mu_0, \mu_1, \mu_2)'$. If the guessed prior has the true parameter $(\beta_0, \beta_1, \beta_2)'$ as its mean, the design would be correctly constructed. Otherwise if the mean of the guessed prior differs from the true parameters, the design would be constructed from the wrong parameter values and is suspected to be often less than optimal. The robustness property of a Bayesian design can be evaluated by the D efficiency defined as

$$\text{D-EFF} = \left(\frac{|I(\beta)| \text{ due to given design constructed with } \mu}{|I(\beta)| \text{ due to given design constructed with } \beta} \right)^{\frac{1}{p}} \qquad (6.4.1)$$

where β denotes the true parameter $(\beta_0, \beta_1, \beta_2)'$, μ denotes the specified mean $(\mu_0, \mu_1, \mu_2)'$ of the prior, and p is the number of individual parameters, or $p = 3$.

The degree of misspecification can be measured with $m_0 = \frac{\mu_0}{\beta_0}$, $m_1 = \frac{\mu_1}{\beta_1}$, and $m_2 = \frac{\mu_2}{\beta_2}$ where $\beta_0 \neq 0$. For a six-point parallel-line design with the logits S - T, S, and S + T, the actual logits of the constructed design fall on

$$L_1 = \beta_0(1 - \frac{m_0}{m_2}) + \frac{1}{m_2} S + (L_X - L_D)(1 - \frac{m_1}{m_2}) - T\frac{m_1 k^2 + m_2}{m_1^2 k^2 + m_2^2},$$

$$L_2 = \beta_0(1 - \frac{m_0}{m_2}) + \frac{1}{m_2} S + (L_X - L_D)(1 - \frac{m_1}{m_2}),$$

$$L_3 = \beta_0(1 - \frac{m_0}{m_2}) + \frac{1}{m_2} S + (L_X - L_D)(1 - \frac{m_1}{m_2}) + T\frac{m_1 k^2 + m_2}{m_1^2 k^2 + m_2^2},$$

$$L_4 = \beta_0(1 - \frac{m_0}{m_2}) + \frac{1}{m_2}S + (L_X + L_D)(1 - \frac{m_1}{m_2}) - T\frac{m_1k^2 + m_2}{m_1^2k^2 + m_2^2},$$

$$L_5 = \beta_0(1 - \frac{m_0}{m_2}) + \frac{1}{m_2}S + (L_X + L_D)(1 - \frac{m_1}{m_2}),$$

$$L_6 = \beta_0(1 - \frac{m_0}{m_2}) + \frac{1}{m_2}S + (L_X + L_D)(1 - \frac{m_1}{m_2}) + T\frac{m_1k^2 + m_2}{m_1^2k^2 + m_2^2},$$

(6.4.2)

where $k = \frac{\beta_1}{\beta_2}$, $L_X = \beta_1 x_{10}$, and $L_D = \beta_1 D_1$. In the case of $\beta_0 = 0$, the term $\beta_0(1 - \frac{m_0}{m_2})$ in (6.4.2) should be replaced by $-\frac{\mu_0}{m_2}$.

Let S and T denote the central and deviation logits of the Bayesian D optimal design under consideration. The optimal sample proportions are denoted r1, r2, and r3. The logits for the optimal design can be denoted by $L_1^* = S - T$, $L_2^* = S$, and $L_3^* = S + T$. Let $P_i^* = \frac{1}{1+e^{-L_i^*}}$ and $W_i^* = P_i^*(1 - P_i^*)$, i = 1, 2, 3. Again, let P_i be the response probability at the i^{th} point of the constructed design, i = 1, ..., 6, or $P_i = \frac{1}{1+e^{-L_i}}$, where L_i are given in (6.4.2). Let $W_i = P_i(1 - P_i)$, i = 1, ..., 6. The D efficiency defined in (6.4.1) can then be expressed as

$$\text{D-EFF} = \{\frac{1}{m_2^2}[(r_1W_1 + r_2W_2 + r_3W_3)(4r_1r_3W_4W_6 + r_2W_5(r_1W_4 + r_3W_6))$$
$$+ (r_1W_4 + r_2W_5 + r_3W_6)(4r_1r_3W_1W_3 + r_2W_2(r_1W_1 + r_3W_3))]/$$
$$[2(r_1W_1^* + r_2W_2^* + r_3W_3^*)(4r_1r_3W_1^*W_3^* + r_2W_2^*(r_1W_1^* + r_3W_3^*))]\}^{\frac{1}{3}}.$$

(6.4.3)

The robustness property of the Bayesian designs with normal priors is illustrated through tables 6.4.1 and 6.4.2, each addressing a particular Bayesian design. Both designs achieve reasonable efficiencies most of the time, and are relatively more efficient when all parameters are moderately overestimated, especially for the latter design where $\beta_0 \neq 0$.

Table 6.4.1 contains efficiencies where $\beta_0 = 0$ and can roughly be compared to the top portion of table 5.1.2, which gives the efficiencies of the non-Bayesian parallel-line optimal design in similar situations of misspecification. The Bayesian design appears to be especially more robust than the conventional parallel-line optimal design under overestimation of the parameters.

Table 6.4.1 Bayesian design efficiency under parameter misspecification, normal prior, $\beta_0 = 0$, $|k| = 0.5$, $L_x = 0$, $L_D = 2$, $\gamma_0 = 0.3$, $\gamma_1 = 0.3$, $\gamma_2 = 0.3$

		m_2							
$\lvert\mu_0\rvert$	m_1	.25	.5	.75	1	1.25	1.5	1.75	2
0	.25	.05	.47	.76	.77	.70	.62	.56	.50
	.5	.29	.51	.80	.87	.81	.73	.65	.58
	.75	.43	.71	.87	.94	.91	.82	.74	.66
	1	.07	.81	.95	1	.98	.90	.82	.73
	1.25	.01	.53	.94	1.02	1.02	.96	.88	.80
	1.5	.00	.23	.76	.98	1.02	.99	.93	.85
	1.75	.00	.09	.51	.85	.98	.99	.95	.89
	2	.00	.03	.29	.66	.88	.95	.94	.90
.5	.25	.07	.49	.75	.76	.69	.62	.55	.50
	.5	.35	.54	.79	.85	.80	.72	.64	.58
	.75	.40	.73	.86	.92	.89	.81	.73	.66
	1	.10	.79	.93	.98	.96	.89	.81	.73
	1.25	.01	.54	.91	1.00	1.00	.95	.87	.79
	1.5	.00	.25	.75	.96	1.00	.97	.91	.84
	1.75	.00	.10	.51	.84	.96	.97	.94	.88
	2	.00	.04	.31	.66	.86	.93	.93	.89
1	.25	.12	.51	.70	.72	.67	.60	.55	.50
	.5	.27	.56	.75	.80	.77	.70	.63	.57
	.75	.24	.67	.80	.86	.84	.78	.71	.64
	1	.12	.66	.85	.91	.90	.85	.78	.71
	1.25	.02	.50	.83	.92	.93	.90	.84	.77
	1.5	.00	.28	.70	.89	.93	.92	.88	.81
	1.75	.00	.12	.51	.79	.90	.92	.89	.85
	2	.00	.46	.33	.64	.82	.89	.89	.86

Table 6.4.2 Bayesian design efficiency under parameter misspecification, normal prior, $\beta_0 = -2$, $|k| = 0.5$, $L_x = 1$, $L_D = 2$, $\gamma_0 = 0.3$, $\gamma_1 = 0.3$, $\gamma_2 = 0.3$

		m_2							
m_0	m_1	.25	.5	.75	1	1.25	1.5	1.75	2
.5	.25	.08	.47	.75	.78	.71	.64	.57	.52
	.5	.32	.48	.78	.87	.82	.73	.65	.58
	.75	.44	.68	.85	.93	.88	.80	.72	.64
	1	.07	.81	.92	.95	.91	.84	.76	.68
	1.25	.01	.56	.89	.93	.90	.84	.77	.70
	1.5	.00	.28	.72	.85	.85	.81	.76	.70
	1.75	.00	.13	.52	.72	.77	.75	.72	.68
	2	.00	.06	.34	.58	.67	.68	.67	.65
1	.25	.06	.43	.67	.74	.70	.64	.58	.52
	.5	.13	.52	.76	.85	.82	.75	.67	.60
	.75	.17	.67	.84	.93	.92	.85	.76	.68
	1	.12	.75	.93	1	.99	.92	.83	.74
	1.25	.02	.56	.95	1.03	1.03	.97	.88	.79
	1.5	.00	.25	.79	.99	1.02	.97	.90	.82
	1.75	.00	.09	.53	.86	.95	.94	.89	.82
	2	.00	.03	.32	.67	.84	.87	.85	.80
1.5	.25	.00	.15	.42	.58	.62	.61	.56	.52
	.5	.00	.21	.51	.68	.73	.71	.66	.60
	.75	.01	.30	.61	.78	.83	.81	.75	.68
	1	.03	.39	.72	.87	.92	.90	.83	.76
	1.25	.03	.41	.79	.94	.99	.97	.91	.83
	1.5	.01	.29	.73	.95	1.02	1.01	.95	.88
	1.75	.00	.13	.54	.86	1.00	1.01	.97	.90
	2	.00	.04	.32	.69	.90	.97	.96	.90

Table 6.4.2 contains efficiencies where $\beta_0 \neq 0$ and is comparable to the lower portion of table 5.1.4, which exhibits the robustness property of the non-Bayesian design in corresponding situations. The Bayesian design often has higher efficiencies and displays considerable improvement in robustness under overestimation. The Bayesian design only shows slightly lower efficiencies when subject to severely underestimated parameters.

For the Bayesian design with uniform priors, the robustness property is illustrated through tables 6.4.3 and 6.4.4, each evaluating a particular design. Both designs are fairly robust and especially so when the parameters are overestimated. The design in table 6.4.4, which is based on a broader prior, appears considerably more robust than the design in table 6.4.3, which results from a relatively tighter prior distribution.

Both table 6.4.3 and table 6.4.4 are comparable to the lower portion of table 5.1.4, which evaluates the robustness in the corresponding non-Bayesian situation. The Bayesian designs achieve better robustness overall than the conventional design, with substantial improvement in the presence of overestimation.

Table 6.4.3 Bayesian design efficiency under parameter misspecification, uniform prior, $\beta_0 = -2$, $|k| = 0.5$, $L_x = 1$, $L_D = 2$, $\gamma_0 = 0.3$, $\gamma_1 = 0.15$, $\gamma_2 = 0.15$

		m_2							
m_0	m_1	.25	.5	.75	1	1.25	1.5	1.75	2
.5	.25	.15	.47	.75	.80	.74	.67	.60	.54
	.5	.37	.50	.77	.87	.83	.75	.68	.61
	.75	.57	.68	.83	.92	.89	.82	.74	.66
	1	.10	.83	.91	.94	.92	.85	.77	.70
	1.25	.01	.60	.88	.93	.90	.85	.78	.72
	1.5	.00	.31	.73	.85	.85	.82	.77	.72
	1.75	.00	.15	.53	.72	.77	.76	.73	.69
	2	.00	.06	.36	.59	.68	.69	.68	.66
1	.25	.22	.50	.69	.76	.73	.67	.61	.55
	.5	.27	.57	.76	.86	.85	.78	.70	.63
	.75	.23	.69	.83	.94	.94	.87	.79	.71
	1	.14	.79	.93	1	1.00	.94	.86	.77
	1.25	.02	.61	.96	1.04	1.04	.98	.90	.82
	1.5	.00	.28	.83	1.01	1.03	.99	.92	.84
	1.75	.00	.10	.57	.88	.97	.95	.90	.84
	2	.00	.04	.35	.70	.86	.89	.86	.82
1.5	.25	.01	.23	.48	.62	.66	.64	.59	.55
	.5	.02	.31	.58	.72	.76	.74	.69	.63
	.75	.03	.40	.68	.82	.87	.84	.78	.71
	1	.05	.46	.78	.91	.95	.93	.87	.79
	1.25	.03	.45	.83	.97	1.02	1.00	.93	.86
	1.5	.01	.32	.77	.98	1.05	1.04	.98	.90
	1.75	.00	.13	.57	.90	1.02	1.04	1.00	.93
	2	.00	.04	.34	.73	.93	1.00	.98	.93

Table 6.4.4 Bayesian design efficiency under parameter misspecification, uniform prior, $\beta_0 = -2$, $|k| = 0.5$, $L_x = 1$, $L_D = 2$, $\gamma_0 = 0.3$, $\gamma_1 = 0.3$, $\gamma_2 = 0.15$

		m_2							
m_0	m_1	.25	.5	.75	1	1.25	1.5	1.75	2
.5	.25	.13	.51	.76	.86	.83	.76	.69	.63
	.5	.42	.56	.77	.90	.91	.85	.78	.71
	.75	.71	.70	.82	.94	.96	.91	.84	.77
	1	.18	.89	.91	.98	.99	.95	.88	.81
	1.25	.02	.77	.96	1.00	1.00	.96	.90	.84
	1.5	.00	.41	.85	.96	.98	.95	.90	.84
	1.75	.00	.18	.64	.85	.91	.91	.87	.83
	2	.00	.08	.44	.70	.81	.84	.83	.80
1	.25	.22	.54	.70	.79	.80	.76	.70	.64
	.5	.30	.59	.76	.87	.90	.86	.80	.73
	.75	.23	.71	.83	.94	.98	.95	.88	.81
	1	.16	.78	.92	1	1.04	1.01	.95	.87
	1.25	.04	.73	.99	1.06	1.08	1.06	1.00	.92
	1.5	.00	.40	.94	1.07	1.10	1.08	1.02	.95
	1.75	.00	.14	.71	.99	1.07	1.06	1.02	.96
	2	.00	.05	.44	.83	.98	1.01	1.00	.95
1.5	.25	.02	.31	.52	.64	.70	.70	.67	.63
	.5	.02	.39	.63	.74	.80	.80	.77	.72
	.75	.03	.46	.72	.84	.90	.90	.86	.80
	1	.04	.48	.80	.92	.98	.98	.94	.88
	1.25	.03	.43	.83	.98	1.05	1.05	1.01	.95
	1.5	.01	.36	.81	1.02	1.09	1.10	1.06	1.00
	1.75	.00	.20	.68	.98	1.09	1.11	1.09	1.03
	2	.00	.07	.45	.84	1.03	1.09	1.09	1.04

Comparing tables 6.4.1-2 for the normal priors to tables 6.4.3-4 for the uniform priors, it is found that uniform priors tend to yield more robust designs than normal priors. This is a result of larger weights given by uniform priors to the values near the margins.

Since Bayesian designs are less than optimal given perfect parameter knowledge, a higher robustness efficiency does not necessarily imply that the Bayesian design offers a larger, or better, criterion value than the conventional optimal design, either with or without parameter misspecification. To better understand the properties of Bayesian designs, it is found that under no parameter misspecification, the Bayesian designs examined in tables 6.4.1 through 6.4.4 have efficiencies of 98.0%, 96.6%, 94.9%, and 89.1% respectively, relative to the D optimal parallel-line design. These high efficiencies suggest that the above Bayesian designs are highly desirable from a practical point of view. In addition, the four designs, with their efficiencies in descending order, appear to have ascending degrees of robustness. This feature reflects the two conflicting design qualities: being efficient under good parameter knowledge versus being robust under poor parameter guesses.

In this chapter we have developed Bayesian D optimal designs with the parallel-line structure for the two-variable logistic model containing interaction. Construction of a Bayesian design is done by using the mean of the prior distribution in the same fashion for the parallel-line design as outlined in §2.1. Both two- and three-line designs are encountered in the optimal solutions. The three-line designs tend to be associated with prior distributions of larger dispersion. Similar to the results of other Bayesian optimal design work, it is found that the Bayesian designs developed here often appear more robust than the non-Bayesian parallel-line optimal design and achieve substantial improvement in certain situations.

CHAPTER 7

OPTIMAL DESIGNS FOR ESTIMATION OF INTERACTION

The optimal designs studied in the previous chapters focus on enhancing the overall estimation of the model parameters. In some dose-response studies, estimation of the interaction effect may be a special concern. To gear a design to the estimation of interaction, the optimization criterion would address the asymptotic variance of the interaction coefficient, which corresponds to a single element of the inverse of the information matrix I(β) rather than the entire information matrix addressed in D optimality. In this chapter we develop optimal designs that minimize the asymptotic variance of the interaction coefficient. The criterion is expressed as

$$\underset{\mathcal{D}}{\mathrm{Min}}\, \mathrm{Var}[\hat{\beta}_{12}]$$

where $\mathcal{D}$ is a set of candidate designs, $\hat{\beta}_{12}$ is the MLE for β_{12}, and $\mathrm{Var}[\hat{\beta}_{12}]$ denotes the asymptotic variance of $\hat{\beta}_{12}$.

The asymptotic variance $\mathrm{Var}[\hat{\beta}_{12}]$ is given by the appropriate diagonal element of the inverse of the information matrix I(β), where I(β) comes from fitting an interaction model. A hyperbola-based design is found by assuming that the dose-response relationship does contain interaction and have hyperbolic ED's. Similar to their approach used with D optimal designs, Brunden *et al.* (1988) derived a factorial design that minimizes $\mathrm{Var}[\hat{\beta}_{12}]$ by assuming a no-interaction kind of dose-response relationship with straight line ED's. Using Brunden's approach, a parallel-line design is also obtained mainly to provide design simplicity. The simple designs, namely the factorial and parallel-line designs, are then compared to the hyperbola-based design in terms of their relative efficiencies in the presence of interaction.

§7.1 Hyperbola-Based Design

The same hyperbola-based design structure as shown in figure 2.2.1 is used to generate a design that minimizes $Var[\hat{\beta}_{12}]$. Given this design structure, the asymptotic variance of $\hat{\beta}_{12}$ is shown to be

$$Var[\hat{\beta}_{12}] = \frac{\beta_{12}^2}{N}[\log(\frac{P_2}{1-P_2}) - \log(\frac{P_1}{1-P_1})]^{-2}[\frac{1}{rP_1(1-P_1)} + \frac{1}{(1-r)P_2(1-P_2)}] \tag{7.1.1}$$

where N is the total sample size, P_1 and P_2 are the ED levels of the design, and r is the sample proportion allocated to the points on ED_{100P_1}. Minimizing the $Var[\hat{\beta}_{12}]$ expression in (7.1.1) yields the optimal solutions

$$P_1 = 0.083222, \qquad P_2 = 0.916778, \qquad \text{and} \qquad r = 0.5. \tag{7.1.2}$$

While the D optimality criterion for a hyperbola-based design involves the central logit β_0^* and the ratio t of the design levels, the criterion $Var[\hat{\beta}_{12}]$ is not a function of β_0^* and t. The optimal design for $Var[\hat{\beta}_{12}]$ is therefore completely defined by (7.1.2) and independent of any additional constraint. This implies that theoretically, the experimenter has complete freedom in choosing where on $ED_{8.3}$ and $ED_{91.7}$ to put the design points, as long as the basic rule is followed that the two points on a common ED are symmetric about the hyperbola center.

Compared to the D optimal hyperbola-based design, which is given by $ED_{17.6}$ and $ED_{82.4}$, the optimal design for $Var[\hat{\beta}_{12}]$ has a larger dispersion. It seems that emphasizing the interaction coefficient in the design criterion tends to increase the design spread in probability space.

§7.2 Factorial Design

The 2×2 factorial design by Brunden *et al.* has the structure shown in figure 2.4.1. In this section we rederive the factorial design in a way that is more consistent with the rest of the design work in this research. Assuming no interaction in the true relationship, the design criterion is given by

$$\mathrm{Var}[\hat{\beta}_{12}] = \frac{16\beta_1^2\beta_2^2}{N}\ [\log(\frac{P_3}{1-P_3}) - \log(\frac{P_1}{1-P_1})]^{-4}$$

$$\cdot[\frac{1}{r_1P_1(1-P_1)} + \frac{4\left(\sqrt{P_1P_3}+\sqrt{(1-P_1)(1-P_3)}\right)^2}{(1-r_1-r_3)\sqrt{P_1P_3(1-P_1)(1-P_3)}} + \frac{1}{r_3P_3(1-P_3)}], \tag{7.2.1}$$

where N is the total sample size, P_1 and P_3 are the lower and upper design probabilities respectively, and r_1 and r_3 are the sample proportions allocated to the design points on the lower and upper ED's. The middle probability P_2 can be expressed as a function of P_1 and P_3 as given in (2.4.1). Minimizing the $\mathrm{Var}[\hat{\beta}_{12}]$ expression in (7.2.1) with respect to P_1, P_3, r_1, and r_3 yields the optimal solution

$$P_1 = 0.008173, \qquad P_2 = 0.5, \qquad P_3 = 0.991827,$$
$$r_1 = r_3 = 0.423704, \quad \text{and} \quad r_2 = 0.076296, \tag{7.2.2}$$

where r_2 is the sample proportion for each point on P_2.

Compared to the D optimal factorial design, which is given by $ED_{1.6}$, ED_{50}, and $ED_{98.4}$, the optimal design for $\mathrm{Var}[\hat{\beta}_{12}]$ has more extreme lower and upper probabilities.

This is in line with what has been observed from the hyperbola-based designs and once again suggests that emphasizing the interaction coefficient in the design criterion tends to stretch out the design to a wider range.

Unlike the D optimal design, which has equal sample sizes, the optimal design for $Var[\hat{\beta}_{12}]$ has almost 85% of the experimental runs allocated to the extreme ED's and only about 15% to the middle ED. This seems to be another consequence of the additional "stretching out" mentioned above.

§7.3 Parallel-Line Design

The parallel-line design for minimizing $Var[\hat{\beta}_{12}]$ assumes the same structure as shown in figure 2.1.1. Assuming no interaction in the true relationship, the design criterion is given by

$$Var[\hat{\beta}_{12}] = \frac{\beta_2^2}{ND_1^2} \frac{1}{f(k)^2} [\log(\frac{P_2}{1-P_2}) - \log(\frac{P_1}{1-P_1})]^{-2} [\frac{1}{rP_1(1-P_1)} + \frac{1}{(1-r)P_2(1-P_2)}], \tag{7.3.1}$$

where D_1 is a distance shown in figure 2.1.1 and characterizes the length of the ED edges, r is the sample proportion allocated to P_1, k is the slope of the ED lines given by $-\frac{\beta_1}{\beta_2}$, and the function f(k) is defined in (2.5.2). As explained in §2.5, a rectangular parallel-line design leads to a singular information matrix for fitting an interaction model when the linear effects are equal in the sense $k = \pm 1$. To avoid this, the convention set in §2.5 has been applied here. Namely, the design is rectangular when $k \neq \pm 1$ and has horizontal non-ED edges if $k = \pm 1$.

Minimizing the criterion in (7.3.1) gives the optimal results

$$P_1 = 0.083222, \qquad P_2 = 0.916778, \qquad \text{and} \qquad r = 0.5. \qquad (7.3.2)$$

These optimal solutions for the parallel-line design are identical to those for the hyperbola-based design. As far as optimization is concerned, the effective part of the criterion expression in (7.3.1) is identical to that in (7.1.1) for a hyperbola-based design. Consequently, both designs are characterized by the same optimal ED's while their structures are totally different.

Compared to the D optimal design, which is given by $ED_{17.6}$ and $ED_{82.4}$, the optimal design for $Var[\hat{\beta}_{12}]$ has more dispersed ED's. This has been observed in all three designs addressed in this chapter. There seems to be remarkable evidence that focusing on the estimation of interaction rather than balancing among all the parameters tends to stretch out the design to more extreme ED's. Namely, the interaction effect might be better estimated with a more dispersed design.

§7.4 Efficiency of Factorial Design

The factorial design by Brunden *et al.* is derived by assuming no interaction. When interaction is indeed present, the asymptotic variance $Var[\hat{\beta}_{12}]$ due to the design is no longer given by the expression in (7.2.1) and is expected to be higher in general. Comparing this asymptotic variance to that of the optimal hyperbola-based design, which is obtained by assuming the presence of interaction, leads to the relative efficiency defined as

$$\frac{Var(\hat{\beta}_{12}) \text{ from design B}}{Var(\hat{\beta}_{12}) \text{ from design A}} \qquad (7.4.1)$$

where design A is currently the factorial design and design B the hyperbola-based design.

Let P_i^* denote the actual response probability at the i^{th} point of the factorial design and r_i^* the sample proportion given to the i^{th} point. It can be shown that

$$\sum_{i=1}^{4} \frac{1}{r_i^* P_i^* (1-P_i^*)} = \frac{2}{r_1 r_2} + \sum_{s_1=-1,1} \sum_{s_2=-1,1} \left(\frac{1}{r_1} e^{s_1 s_2 L_0 + \frac{s_1}{\Delta_\beta}(L_x - \frac{s_2 L_0}{2})(\beta_0 + L_x + \frac{s_2 L_0}{2})} + \frac{1}{r_2} e^{\frac{s_1}{\Delta_\beta}(L_x + \frac{s_2 L_0}{2})(\beta_0 + L_x + \frac{s_2 L_0}{2})} \right) \quad (7.4.2)$$

where L_0 is the upper optimal logit 4.79871 for the factorial design, r_1 and r_2 are the optimal sample proportions given in (7.2.2), Δ_β is the logit distance between the origin and the hyperbola center given by (4.2.2), and L_x is a logit type index of the design location and was discussed in detail in §4.3.

The relative efficiency defined in (7.4.1) can be expressed as

$$EFF = F(\beta_0, \Delta_\beta, L_x) = \frac{L_0^4}{L^2 \Delta_\beta^2} \frac{\frac{1}{P(1-P)}}{\sum_{i=1}^{4} \frac{1}{r_i^* P_i^* (1-P_i^*)}} \quad (7.4.3)$$

where P is the upper probability 0.916778 for the hyperbola-based design and L =logit(P).

Due to the symmetries

(1) $F(\beta_0, \Delta_\beta, L_x) = F(-\beta_0, -\Delta_\beta, -L_x)$

and (2) $F(\beta_0, \Delta_\beta, L_x) = F(\beta_0, \Delta_\beta, -\beta_0 - L_x)$,

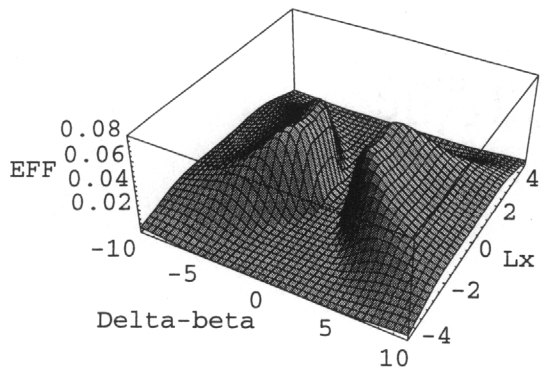
only cases of $\beta_0 < 0$ and $L_x \geq -\frac{\beta_0}{2}$ will be considered. The efficiency is evaluated in table 7.4.1. Conditions with maximum efficiencies are listed in table 7.4.2. The efficiency given each β_0 is plotted versus Δ_β and L_x in figure 7.4.1.

Table 7.4.1 Efficiency of factorial design for estimation of interaction

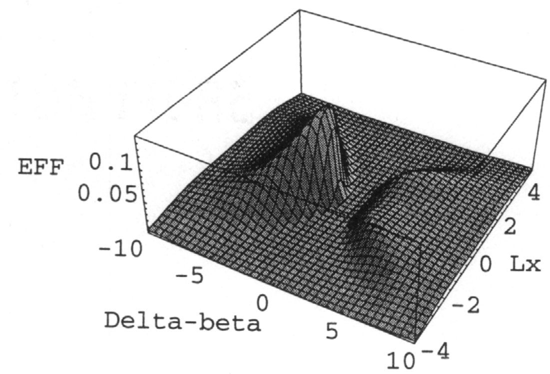
β_0	L_x	Δ_β -10	-5	-2	-1	1	2	5	10
0	0	.015	.042	.054	.012	.012	.054	.042	.015
	1	.016	.045	.040	.001	.001	.040	.045	.016
	2	.016	.036	.001	.000	.000	.001	.036	.016
	3	.013	.009	.000	.000	.000	.000	.009	.013
	4	.007	.001	.000	.000	.000	.000	.001	.007
-1	0.5	.018	.061	.136	.070	.001	.015	.026	.012
	1.5	.019	.063	.062	.001	.001	.019	.029	.013
	2.5	.019	.045	.002	.000	.000	.002	.029	.013
	3.5	.015	.012	.000	.000	.000	.000	.009	.011
	4.5	.008	.001	.000	.000	.000	.000	.001	.006
-2	1	.022	.082	.270	.168	.000	.003	.015	.010
	2	.022	.083	.097	.002	.000	.005	.017	.010
	3	.022	.056	.002	.000	.000	.002	.021	.011
	4	.018	.012	.000	.000	.000	.000	.009	.010
	5	.009	.001	.000	.000	.000	.000	.001	.006
-5	2.5	.028	.119	.081	.000	.000	.000	.001	.003
	3.5	.030	.127	.123	.000	.000	.000	.002	.004
	4.5	.032	.113	.030	.000	.000	.000	.003	.004
	5.5	.030	.033	.000	.000	.000	.000	.005	.005
	6.5	.017	.004	.000	.000	.000	.000	.002	.005

Table 7.4.2 Maximum efficiencies of factorial design for estimation of interaction

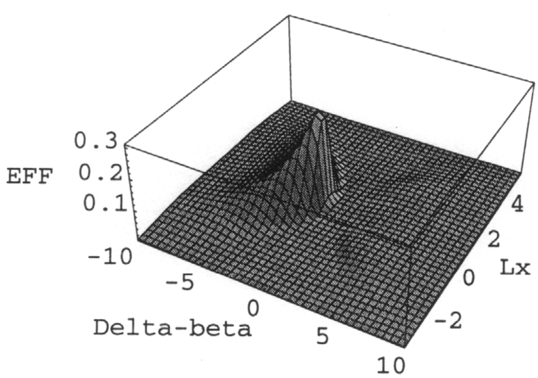
β_0	Max EFF	L_x	Δ_β
0	0.063371	$\pm$0.745772	$\pm$2.83113
-1	0.136175	0.5	-1.92800
-2	0.300293	1	-1.52606
-5	0.305351	1.65593, 3.34407	-2.70623
	Global Max:		
-3.28776	0.485127	1.64388	-1.86346



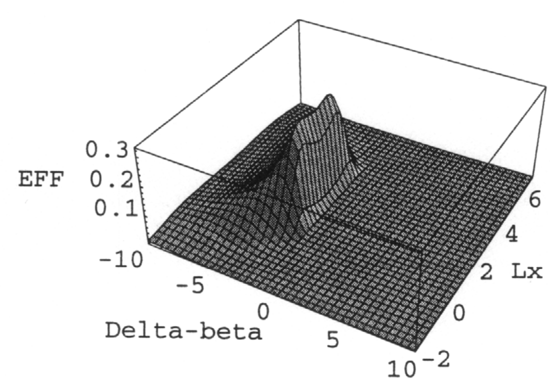
(1) βo = 0



(2) βo = −1



(3) βo = −2



(4) βo = −5

Figure7.4.1 Efficiency of factorial design for estimation of interaction

The above results indicate that the factorial design is very inefficient in estimating the interaction effect compared to the hyperbola-based design. The relative efficiency is often close to zero, while the best case is only near 0.5. Similar to what has been observed with the D optimal designs, the design is more inefficient when $\Delta_\beta>0$ than it is when $\Delta_\beta<0$, especially if $|\beta_0|$ is large. Given that $\beta_0<0$ is currently discussed, this means that the performance is worse when the design center lies further away from the hyperbola center. From the conditions associated with a maximum efficiency in table 7.4.2, the best location(s) for the design center is indicated by values of L_x either equal to or symmetric about $-\frac{\beta_0}{2}$. Regarding the above comments, similar discussion with more details is given in §4.3 which deals with the D optimal designs.

Due to the closeness to zero of its efficiencies, the factorial design for minimizing $Var[\hat{\beta}_{12}]$ does not seem to be an effective simple alternative to the hyperbola-based optimal design.

§7.5 Efficiency of Parallel-Line Design

Based on similar motivation, the relative efficiency of the parallel-line design can also be found using the definition in (7.4.1). The parallel-line design is assumed to have an equal area to that of the hyperbola-based design. Let P_i^* denote the actual response probability at the i^{th} point of the parallel-line design. It can be shown that

$$\sum_{i=1}^{4}\frac{1}{P_i^*(1-P_i^*)} = 8 + \sum_{s_1=-1,1}\sum_{s_2=-1,1}\sum_{s_3=-1,1} e^{s_1 s_3 L_0 + \frac{s_3}{\Delta_\beta}(L_x + \frac{s_1 L_0}{2}(1+f(k)) + s_2\Delta_\beta)(-\beta_0 - L_x + \frac{s_1 L_0}{2}(1-f(k)) - s_2\Delta_\beta)} \tag{7.5.1}$$

where L_0 is the upper logit of the parallel-line design, which is the same as that of the hyperbola-based design, Δ_β and L_x remain the same as in §7.4, and the function f(k) is given by (2.5.2) where $k = -\frac{\beta_1}{\beta_2}$. The efficiency of the parallel-line design is given by

$$EFF = F(k, \beta_0, \Delta_\beta, L_x) = f(k)^2 \frac{\frac{1}{P(1-P)}}{\frac{1}{4}\sum_{i=1}^{4}\frac{1}{P_i^*(1-P_i^*)}}. \qquad (7.5.2)$$

Due to the symmetries

(1) $F(k, \beta_0, \Delta_\beta, L_x) = F(k, -\beta_0, -\Delta_\beta, -L_x)$,

(2) $F(k, \beta_0, \Delta_\beta, L_x) = F(-k, \beta_0, \Delta_\beta, L_x)$,

and (3) for $k \neq \pm 1$, $F(k, \beta_0, \Delta_\beta, L_x) = F(\frac{1}{k}, \beta_0, \Delta_\beta, -\beta_0 - L_x)$,

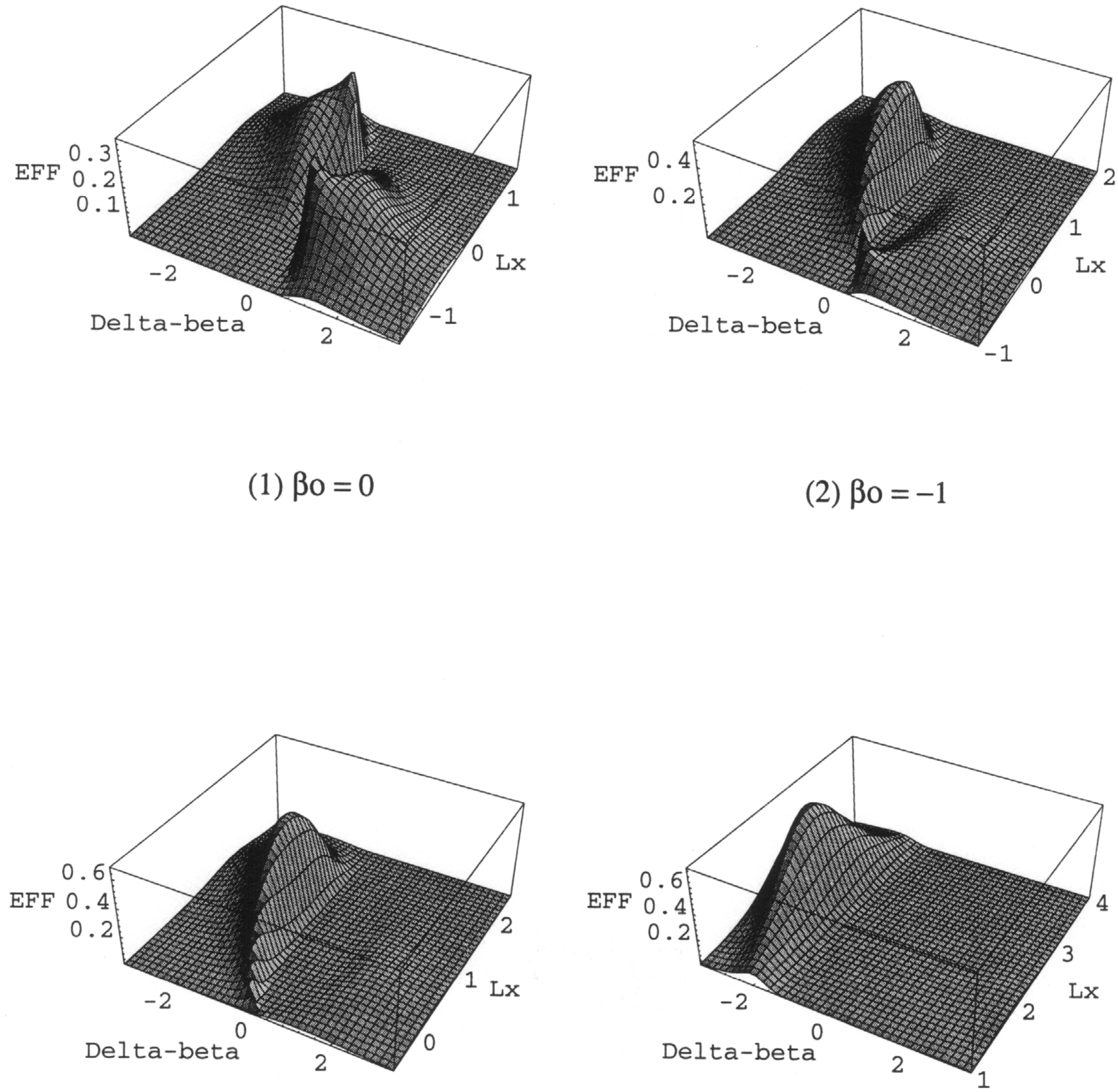
the efficiencies will be investigated for the restricted parameter space $|k| \in (0,1] \cap \beta_0 \in (-\infty, 0]$. The user selected design location is reflected by the scale free index L_x, which was discussed with more details in §4.3. As all previous efficiency analyses suggest that good results tend to occur at or near $L_x = -\frac{\beta_0}{2}$, the efficiency will be examined for a range of L_x symmetric about $-\frac{\beta_0}{2}$. Table 7.5.1 contains the efficiency for the above cases. Figure 7.5.1 shows the efficiency plots versus Δ_β and L_x for given β_0 at $|k| = 0.5$. The plots for other values of k look roughly the same. Situations producing maximum efficiencies are summarized in table 7.5.2.

Table 7.5.1 Efficiency of parallel-line design for estimation of interaction

$\lvert k \rvert$	β_0	L_x	Δ_β -3	-2	-1	-0.5	0.5	1	2	3
0.2	0	-1.5	.000	.000	.000	.000	.000	.005	.023	.019
		-0.5	.006	.012	.019	.010	.627	.458	.191	.069
		0	.023	.063	.159	.208	.208	.159	.063	.023
		0.5	.069	.191	.458	.627	.010	.019	.012	.006
		1.5	.019	.023	.005	.000	.000	.000	.000	.000
	-1	-1.0	.000	.000	.000	.000	.001	.019	.041	.025
		0	.009	.025	.077	.155	.121	.181	.129	.053
		0.5	.037	.126	.525	.921	.018	.057	.039	.017
		1.0	.098	.291	.553	.155	.001	.007	.008	.004
		2.0	.015	.015	.002	.000	.000	.000	.000	.000
	-2	-0.5	.001	.001	.001	.000	.030	.085	.067	.034
		0.5	.017	.066	.421	.295	.003	.040	.093	.046
		1.0	.069	.295	.973	.061	.001	.019	.031	.015
		1.5	.147	.404	.421	.021	.000	.004	.006	.003
		2.5	.013	.012	.001	.000	.000	.000	.000	.000
	-5	1.0	.011	.079	.007	.000	.000	.000	.005	.032
		2.0	.286	.419	.001	.000	.000	.000	.005	.041
		2.5	.706	.498	.001	.000	.000	.000	.006	.021
		3.0	.494	.478	.001	.000	.000	.000	.006	.006
		4.0	.024	.027	.001	.000	.000	.000	.000	.000
0.5	0	-1.5	.000	.000	.000	.000	.001	.025	.030	.014
		-0.5	.007	.019	.046	.040	.224	.261	.155	.055
		0	.025	.076	.176	.090	.090	.176	.076	.025
		0.5	.055	.155	.261	.224	.040	.046	.019	.007
		1.5	.014	.030	.025	.001	.000	.000	.000	.000
	-1	-1.0	.001	.001	.001	.000	.021	.050	.029	.013
		0	.012	.039	.173	.282	.015	.078	.099	.042
		0.5	.040	.147	.463	.287	.005	.056	.047	.019
		1.0	.082	.253	.495	.282	.003	.017	.012	.005
		2.0	.016	.028	.010	.000	.000	.000	.000	.000
	-2	-0.5	.001	.003	.007	.010	.024	.031	.026	.012
		0.5	.022	.100	.466	.035	.000	.013	.067	.036
		1.0	.075	.318	.556	.013	.000	.012	.036	.016
		1.5	.131	.386	.466	.011	.000	.007	.010	.005
		2.5	.018	.026	.007	.000	.000	.000	.000	.000
	-5	1.0	.018	.098	.001	.000	.000	.000	.002	.011
		2.0	.318	.232	.000	.000	.000	.000	.002	.029
		2.5	.610	.306	.000	.000	.000	.000	.004	.022
		3.0	.465	.320	.000	.000	.000	.000	.005	.008
		4.0	.038	.052	.000	.000	.000	.000	.000	.000

Table 7.5.1 (continued)

|k|	β_0	L_x	Δ_β -3	-2	-1	-0.5	0.5	1	2	3
0.8	0	-1.5	.000	.000	.000	.000	.006	.005	.002	.001
		-0.5	.003	.009	.024	.010	.013	.035	.023	.008
		0	.008	.023	.041	.010	.010	.041	.023	.008
		0.5	.008	.023	.035	.013	.010	.024	.009	.003
		1.5	.001	.002	.005	.006	.000	.000	.000	.000
	-1	-1.0	.000	.001	.002	.003	.001	.002	.001	.001
		0	.005	.018	.066	.035	.001	.010	.014	.006
		0.5	.012	.043	.091	.032	.001	.012	.015	.006
		1.0	.012	.041	.078	.035	.001	.008	.006	.002
		2.0	.001	.004	.011	.003	.000	.000	.000	.000
	-2	-0.5	.001	.002	.012	.020	.000	.001	.001	.001
		0.5	.009	.042	.082	.002	.000	.002	.010	.005
		1.0	.022	.083	.093	.001	.000	.002	.011	.005
		1.5	.021	.073	.082	.001	.000	.002	.005	.002
		2.5	.002	.008	.012	.001	.000	.000	.000	.000
	-5	1.0	.010	.019	.000	.000	.000	.000	.000	.001
		2.0	.091	.043	.000	.000	.000	.000	.000	.005
		2.5	.133	.060	.000	.000	.000	.000	.001	.006
		3.0	.105	.053	.000	.000	.000	.000	.001	.003
		4.0	.015	.017	.000	.000	.000	.000	.000	.000
1	0	-1.5	.016	.018	.003	.000	.000	.000	.000	.000
		-0.5	.066	.181	.444	.461	.005	.014	.010	.005
		0	.021	.056	.135	.183	.183	.135	.056	.021
		0.5	.005	.010	.014	.005	.461	.444	.181	.066
		1.5	.000	.000	.000	.000	.000	.003	.018	.016
	-1	-1.0	.013	.012	.001	.000	.000	.000	.000	.000
		0	.093	.270	.461	.090	.001	.005	.006	.004
		0.5	.034	.112	.468	1	.019	.049	.035	.015
		1.0	.008	.021	.057	.090	.180	.197	.124	.051
		2.0	.000	.000	.000	.000	.000	.012	.034	.023
	-2	-0.5	.011	.009	.001	.000	.000	.000	.000	.000
		0.5	.138	.368	.337	.019	.000	.003	.005	.003
		1.0	.063	.266	1	.089	.001	.019	.027	.013
		1.5	.015	.056	.337	.461	.005	.049	.092	.045
		2.5	.000	.001	.001	.000	.012	.066	.065	.035
	-5	1.0	.020	.021	.001	.000	.000	.000	.000	.000
		2.0	.461	.484	.001	.000	.000	.000	.006	.005
		2.5	.675	.537	.001	.000	.000	.000	.006	.020
		3.0	.258	.461	.001	.000	.000	.000	.005	.042
		4.0	.009	.064	.012	.000	.000	.000	.006	.038



(1) $\beta o = 0$

(2) $\beta o = -1$

(3) $\beta o = -2$

(4) $\beta o = -5$

Figure7.5.1 Efficiency of parallel-line design for estimation of interaction, $|k| = 0.5$

Table 7.5.2 Maximum efficiencies of parallel-line design for estimation of interaction

$\lvert k \rvert$	β_0	Max EFF	L_x	Δ_β
0.2	0	0.852060	0.287814	-0.259669
	-1	0.964767	0.594040	-0.603448
	-2	0.984445	1.03727	-1.06476
	-5	0.990601	2.50609	-2.53572
0.5	0	0.363788	0.739006	-0.401120
	-1	0.526365	0.840148	-0.865938
	-2	0.662670	1.08987	-1.26558
	-5	0.725659	2.51485	-2.65208
0.8	0	0.044105	0.133029	-1.21365
	-1	0.092291	0.539684	-1.10619
	-2	0.127309	1.01219	-1.38246
	-5	0.145704	2.50589	-2.72261
1	0	1*	0*	0*
	-1	1	0.5	-0.5
	-2	1	1	-1
	-5	1	2.5	-2.5

* : The values with asterisks are meaningful only in the limiting sense since Δ_β by definition is not allowed to be 0, i.e. for $\lvert k \rvert = 1$, the efficiency approaches 1 as β_0, L_x, and Δ_β all tend to zero in a manner such that $L_x = -\beta_0/2$ and $\Delta_\beta = \beta_0/2$.

Similar to every previous case in the efficiency studies of this research, the type of interaction with $\Delta_\beta<0$ can cause much better performance than the best cases with $\Delta_\beta>0$, especially for large $\lvert\beta_0\rvert$. Given $\beta_0<0$ is assumed, this roughly means that the parallel-line design would be less inferior when the kind of interaction renders the hyperbola-based design closer to the center of the parallel-line design. More insight regarding Δ_β can be found in §4.2 and §4.3. From table 7.5.2, the best value of L_x, which determines the design location, is either close or equal to $-\frac{\beta_0}{2}$. Based on the efficiencies, the parallel-line design does not seem to offer comparable performance with respect to the hyperbola-based design in estimating interaction.

Compared to the factorial design, the parallel-line design offers better efficiencies but over smaller regions of the parameter space. In conclusion, based on the low efficiencies of both simple designs, the overall impression suggests that the hyperbola-based design is still a superior choice over the factorial and parallel-line designs for estimation of interaction.

CHAPTER 8

Q OPTIMAL DESIGNS

The D optimal designs and the designs minimizing $\text{Var}[\hat{\beta}_{12}]$ studied in the previous chapters are both driven by parameter estimation. As discussed in the introduction, another goal of design optimization focuses on the prediction of response. One such criterion is Q optimality. A Q optimal design minimizes the average scaled variance of a predicted response. For the logistic model, Q optimality can be defined as either minimizing the average scaled asymptotic prediction variance of the response probability

$$\underset{\mathcal{D}}{\text{Min}}\, \text{APV}[\hat{p}]$$

or minimizing the average scaled asymptotic prediction variance of the logit

$$\underset{\mathcal{D}}{\text{Min}}\, \text{APV}[\text{logit}(\hat{p})].$$

In this chapter, optimal designs are derived by addressing the above two criteria separately for the two-variable logistic model containing no interaction. Relative efficiencies between the designs based on the logit and those based on the response probability show that the two types of designs are fairly close in their performance. Q optimal designs addressing the logit are also obtained for the interaction model.

§8.1 Parallel-Line Design Based on Logit

A Q optimal design minimizes the average scaled prediction variance over a preselected region in the design space and hence is region dependent. For the no-interaction model, consider a rectangular region R bounded by $ED_{100P_1'}$ and $ED_{100P_2'}$ and a rectangular parallel-line type design falling on ED_{100P_1} and ED_{100P_2}. The region and design arrangement are illustrated in figure 8.1.1. The region centers at (x_{1c}, x_{2c}) while the design at (x_{10}, x_{20}). The region center is a user's choice whereas the design center is subject to no initial constraint. Equal sample sizes are assumed for the design points on the same ED.

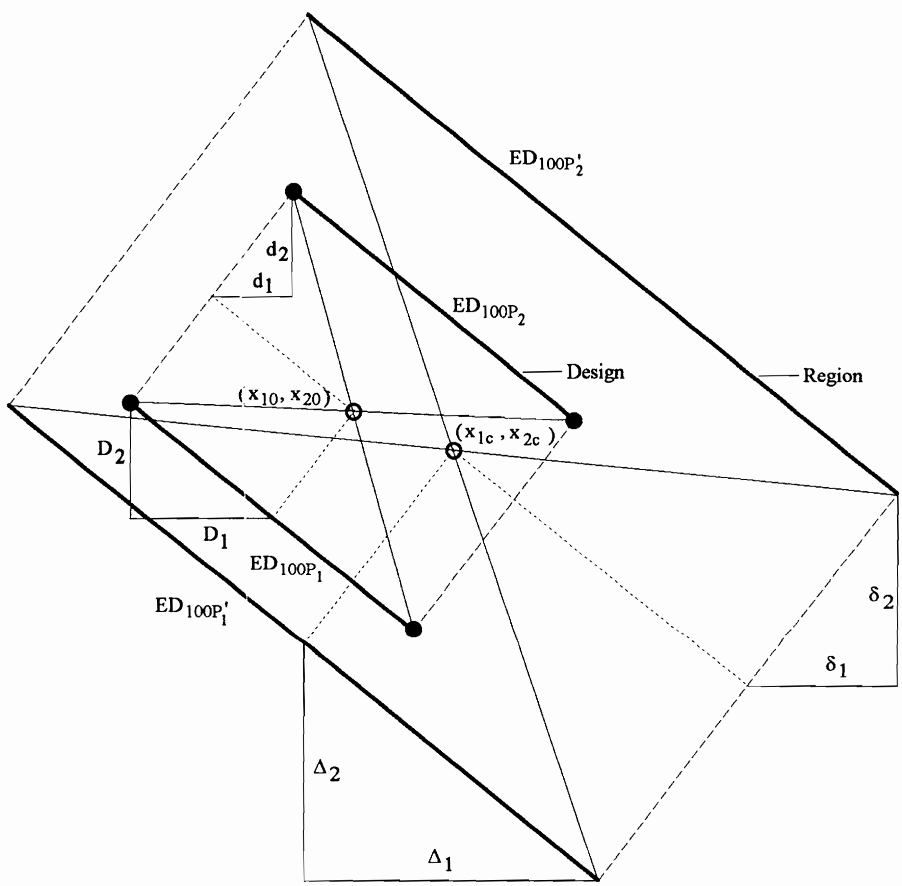


Figure 8.1.1 Region and design arrangement for Q optimal parallel-line design

The Q optimality criterion is derived in appendix A and is given below:

$$\mathrm{APV}[\mathrm{logit}(\hat{p})] = \frac{1}{(L_2 - L_1)^2}\left[\frac{(L_2 - \frac{L_1' + L_2'}{2})^2}{rP_1(1-P_1)} + \frac{(L_1 - \frac{L_1' + L_2'}{2})^2}{(1-r)P_2(1-P_2)}\right]$$

$$+ \frac{(L_2' - L_1')^2}{12(L_2 - L_1)^2}\left[\frac{1}{rP_1(1-P_1)} + \frac{1}{(1-r)P_2(1-P_2)}\right]$$

$$+ \frac{\frac{1}{D_1^2}[(x_{10} - x_{1c} - \frac{\beta_1}{\alpha}(\frac{L_1 + L_2}{2} - \frac{L_1' + L_2'}{2}))]^2 + \frac{\Delta_1^2}{3D_1^2}}{rP_1(1-P_1) + (1-r)P_2(1-P_2)}$$

(8.1.1)

where $L_i = \mathrm{logit}(P_i)$, $L_i' = \mathrm{logit}(P_i')$, i = 1, 2, $\alpha = \beta_1^2 + \beta_2^2$, r is the sample proportion on ED_{100P_1}, D_1 is a design distance defined in figure 8.1.1, and Δ_1 is the corresponding region attribute also shown in figure 8.1.1.

The first optimization result concerns with the design center. Minimizing $\mathrm{APV}[\mathrm{logit}(\hat{p})]$ in (8.1.1) immediately leads to

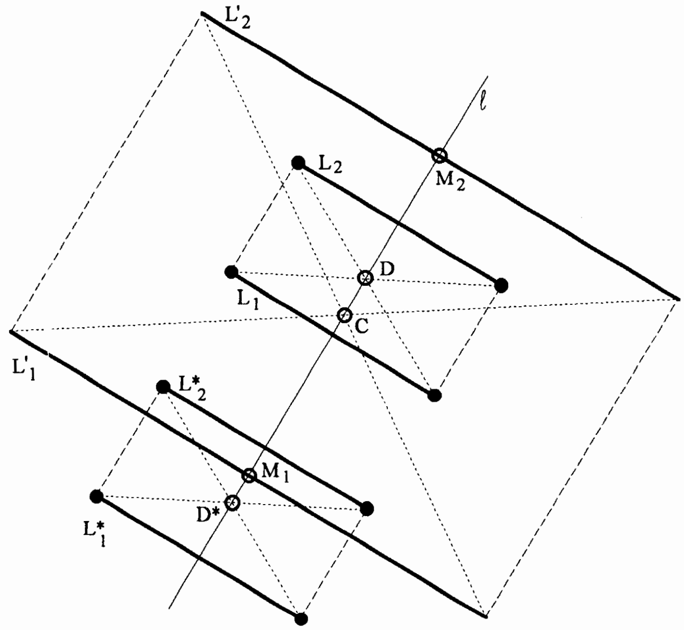
$$x_{10} = x_{1c} + \frac{\beta_1}{\alpha}(\frac{L_1 + L_2}{2} - \frac{L_1' + L_2'}{2}),$$

which further implies that

$$x_{20} = x_{2c} + \frac{\beta_2}{\alpha}(\frac{L_1 + L_2}{2} - \frac{L_1' + L_2'}{2}).$$

(8.1.2)

Regardless of what the design logits are, the optimal location for the design center must satisfy (8.1.2). In fact, it can be shown that the condition in (8.1.2) holds if and only if the center of the design falls on the straight line that goes through the region center (x_{1c}, x_{2c}) and the midpoints of the two region boundaries on $ED_{100P_1'}$ and $ED_{100P_2'}$. This feature is shown in figure 8.1.2. Two examples of optimal design center are given, denoted D and D*. Both points lie on line ℓ that goes through the middle of the region. As a result, the design center, initially free in the two-dimensional space, is now confined on a one-dimensional line. Within the central line of the region, its journey will end up at a point which has to be determined by the optimal design logits.



<u>*Legend*</u>

L_1', L_2' : region logits
L_1, L_2 : design logits
L_1^*, L_2^* : design logits
C : region center
M_1, M_2: midpoints of region edges on L_1' and L_2'
D : optimal design center for the design on L_1 and L_2
D* : optimal design center for the design on L_1^* and L_2^*
ℓ : line formed by points M_1, M_2, and C

Figure 8.1.2 Optimal location of design center for Q optimal parallel-line design

For a design with its center satisfying (8.1.2), the Q criterion reduces to

$\text{APV}[\text{logit}(\hat{p})] = Q(P_1, P_2, r \mid P_1', P_2', g)$

$$= \frac{1}{(L_2 - L_1)^2}\left[\frac{(L_2 - \frac{L_1' + L_2'}{2})^2}{rP_1(1-P_1)} + \frac{(L_1 - \frac{L_1' + L_2'}{2})^2}{(1-r)P_2(1-P_2)}\right]$$

$$+ \frac{(L_2' - L_1')^2}{12(L_2 - L_1)^2}\left[\frac{1}{rP_1(1-P_1)} + \frac{1}{(1-r)P_2(1-P_2)}\right] + \frac{1}{3g^2[rP_1(1-P_1) + (1-r)P_2(1-P_2)]} \tag{8.1.3}$$

where g is the ratio of the length of the ED edges for the design relative to that for the region, or $g = \frac{\sqrt{D_1^2 + D_2^2}}{\sqrt{\Delta_1^2 + \Delta_2^2}} = \left|\frac{D_1}{\Delta_1}\right| = \left|\frac{D_2}{\Delta_2}\right|$. In short, g is a ratio that describes the design width relative to the region width along the direction of the ED lines. Orthogonal to this direction, the size of the design relative to that of the region is characterized by the design logits given the region logits.

As anticipated from intuition, the expression in (8.1.3) indicates that the Q criterion always improves as g increases. When g goes to infinity, the Q criterion of a parallel-line design approaches that of a two-point design for the one-variable logistic model, which is given by

$$APV[\text{logit}(\hat{p})] = \frac{1}{(L_2 - L_1)^2}\left[\frac{(L_2 - \frac{L_1' + L_2'}{2})^2}{rP_1(1-P_1)} + \frac{(L_1 - \frac{L_1' + L_2'}{2})^2}{(1-r)P_2(1-P_2)}\right]$$

$$+ \frac{(L_2' - L_1')^2}{12(L_2 - L_1)^2}\left[\frac{1}{rP_1(1-P_1)} + \frac{1}{(1-r)P_2(1-P_2)}\right]. \tag{8.1.4}$$

The Q optimal designs are found for given P_1', P_2', and g, where P_1' and P_2' correspond to a user selected region of interest while g is a user preferred "design to region" ratio.

For symmetric regions, the Q optimal designs are given in table 8.1.1. These designs have symmetric ED's with equal sample sizes. When the region is broad, the design stays within the region. As the region shrinks, the design also becomes narrower but does not contract quickly enough to catch up with the region. Consequently, the design tends to cross over the region boundaries when the region becomes too small. Similar to the D optimal designs, the two-variable Q optimal design also has smaller dispersion than the corresponding one-variable design. As g increases, the design is stretched out and will eventually approach the one-variable solution. The Q criterion improves as the span of the region shrinks or the "design to region" ratio g increases.

For asymmetric regions, one needs to recognize a property of the Q criterion:

$$Q(P_1, P_2, r \mid P_1', P_2', g) = Q(1\text{-}P_2, 1\text{-}P_1, 1\text{-}r \mid 1\text{-}P_2', 1\text{-}P_1', g) . \qquad (8.1.5)$$

From the practical perspective, an asymmetric region of interest often spans over a relatively small range either above or below ED_{50}. Due to the property in (8.1.5), optimal designs are given only for lower asymmetric regions. These results appear in table 8.1.2. The designs still have symmetric ED's but with uneven sample allocation much more weighted toward the side of the region.

For asymmetric regions, conditional optimal designs are also found by restricting equal sample sizes. These designs are given in table 8.1.3. As expected intuitively, the ED's of these designs turn out to be asymmetric and shift toward the side of the region. The reason is that the sample allocation, which would otherwise have the priority, is not allowed to make the necessary adjustment required by the asymmetry of the region.

To construct a Q optimal parallel-line design, the design center first needs to be obtained from the optimal design logits L_1 and L_2 using the equations in (8.1.2). The design points can then be computed using the design matrix in (2.1.7), with -L and L replaced by L_1 and L_2 in case of asymmetric design ED's.

Table 8.1.1 Q optimal parallel-line designs based on logit, symmetric regions

Region	g	Design	Weighting(%)	APV[logit($\hat{p}$)]
0.01-0.99	1/3	0.188929-0.811071	50-50	47.7
	1/2	0.166422-0.833578	50-50	36.4
	1	0.145518-0.854482	50-50	28.4
	2	0.138550-0.861450	50-50	26.7
0.05-0.95	1/3	0.229003-0.770997	50-50	33.8
	1/2	0.204459-0.795541	50-50	24.0
	1	0.179999-0.820001	50-50	17.5
	2	0.171377-0.828623	50-50	15.8
0.10-0.90	1/3	0.255802-0.744198	50-50	28.4
	1/2	0.231136-0.768864	50-50	19.4
	1	0.205620-0.794380	50-50	13.6
	2	0.196353-0.803647	50-50	12.0
0.15-0.85	1/3	0.276962-0.723038	50-50	25.4
	1/2	0.252794-0.747206	50-50	16.9
	1	0.227152-0.772848	50-50	11.4
	2	0.217647-0.782353	50-50	10.0
0.20-0.80	1/3	0.296267-0.703733	50-50	23.3
	1/2	0.272957-0.727043	50-50	15.1
	1	0.247716-0.752284	50-50	9.9
	2	0.238202-0.761798	50-50	8.6
0.25-0.75	1/3	0.315279-0.684721	50-50	21.6
	1/2	0.293149-0.706851	50-50	13.8
	1	0.268756-0.731244	50-50	8.8
	2	0.259428-0.740572	50-50	7.5
0.30-0.70	1/3	0.335089-0.664911	50-50	20.2
	1/2	0.314505-0.685495	50-50	12.7
	1	0.291444-0.708556	50-50	7.9
	2	0.282505-0.717495	50-50	6.7

Table 8.1.2 Q optimal parallel-line designs based on logit, asymmetric regions

Region	g	Design	Weighting(%)	APV[logit($\hat{p}$)]
0.10-0.20	1/3	0.212136-0.787864	85.4-14.6	29.5
	1/2	0.181592-0.818408	89.9-10.1	19.0
	1	0.156318-0.843682	93.1- 6.9	11.9
	2	0.149966-0.850034	93.6- 6.4	9.9
0.30-0.40	1/3	0.338725-0.661275	90.6- 9.4	18.5
	1/2	0.329578-0.670422	89.8-10.2	11.1
	1	0.319657-0.680343	88.4-11.6	6.5
	2	0.315742-0.684258	87.8-12.2	5.3
0.05-0.25	1/3	0.200904-0.799096	80.5-19.5	33.6
	1/2	0.172204-0.827796	83.9-16.1	22.7
	1	0.144484-0.855516	86.8-13.2	15.2
	2	0.135500-0.864500	87.5-12.5	13.1
0.10-0.30	1/3	0.229930-0.770070	84.7-15.3	27.2
	1/2	0.204772-0.795228	87.3-12.7	17.4
	1	0.182802-0.817198	88.6-11.4	11.0
	2	0.175788-0.824212	88.8-11.2	9.3
0.20-0.40	1/3	0.291191-0.708809	86.5-13.5	21.2
	1/2	0.274995-0.725005	86.6-13.4	13.0
	1	0.258933-0.741067	85.9-14.1	7.9
	2	0.252998-0.747002	85.5-14.5	6.6
0.05-0.40	1/3	0.217413-0.782587	78.5-21.5	31.1
	1/2	0.191396-0.808604	80.6-19.4	20.9
	1	0.166158-0.833842	82.1-17.9	14.1
	2	0.157493-0.842507	82.4-17.6	12.2
0.05-0.50	1/3	0.226539-0.773461	76.2-23.8	30.2
	1/2	0.201421-0.798579	77.7-22.3	20.3
	1	0.176572-0.823428	78.6-21.4	13.8
	2	0.167832-0.832168	78.8-21.2	12.0

Table 8.1.3 Q optimal parallel-line designs based on logit, asymmetric regions, restricted weighting: 50%-50%

Region	g	Design	APV[logit($\hat{p}$)]
0.10-0.20	1/3	0.105384-0.387876	32.4
	1/2	0.083810-0.254532	21.1
	1	0.073824-0.225784	13.2
	2	0.070139-0.220717	11.1
0.30-0.40	1/3	0.275937-0.448919	18.6
	1/2	0.262076-0.450570	11.1
	1	0.246508-0.455550	6.5
	2	0.240352-0.457987	5.4
0.05-0.25	1/3	0.134647-0.601954	38.0
	1/2	0.068895-0.338729	28.3
	1	0.047408-0.236282	20.4
	2	0.043343-0.222482	18.1
0.10-0.30	1/3	0.126783-0.432345	29.1
	1/2	0.099213-0.349239	19.2
	1	0.081097-0.308734	12.6
	2	0.075471-0.299552	10.8
0.20-0.40	1/3	0.205389-0.457025	21.6
	1/2	0.183429-0.443735	13.3
	1	0.161456-0.436042	8.2
	2	0.153380-0.434092	6.9
0.05-0.40	1/3	0.137089-0.570642	33.9
	1/2	0.106282-0.511391	24.6
	1	0.071489-0.382338	18.4
	2	0.064242-0.355220	16.7
0.05-0.50	1/3	0.147533-0.587017	32.3
	1/2	0.123330-0.569796	23.1
	1	0.098736-0.518179	17.3
	2	0.089137-0.486875	15.8

§8.2 Parallel-Line Design Based on Response Probability

Given the region and design arrangement shown in figure 8.1.1, the Q optimality criterion based on the response probability is derived in appendix B. The result in appendix B indicates that the best location for the design center again satisfies the condition in (8.1.2). The same feature as illustrated in figure 8.1.2 applies here. Once the design center satisfies the condition, the Q optimality criterion becomes a function of the design attributes P_1, P_2, and r given the user specifications P_1', P_2', and g in the form

$$
\begin{aligned}
APV[\hat{p}] &= Q(P_1, P_2, r \mid P_1', P_2', g) \\
&= \frac{1}{3(L_2' - L_1')(L_2 - L_1)^2} \\
&\quad \cdot \Big[\frac{P_2'^2(3-2P_2')(L_2' - L_2)^2 - P_1'^2(3-2P_1')(L_1' - L_2)^2}{2rP_1(1-P_1)} \\
&\quad + \frac{P_2'^2(3-2P_2')(L_2' - L_1)^2 - P_1'^2(3-2P_1')(L_1' - L_1)^2}{2(1-r)P_2(1-P_2)} \\
&\quad + \frac{P_2'(1-P_2')(L_2' - L_2) - P_1'(1-P_1')(L_1' - L_2)}{rP_1(1-P_1)} \\
&\quad + \frac{P_2'(1-P_2')(L_2' - L_1) - P_1'(1-P_1')(L_1' - L_1)}{(1-r)P_2(1-P_2)} \\
&\quad - \left(\frac{L_2}{rP_1(1-P_1)} + \frac{L_1}{(1-r)P_2(1-P_2)}\right)\log\left(\frac{1-P_2'}{1-P_1'}\right) \\
&\quad - \left(\frac{1}{rP_1(1-P_1)} + \frac{1}{(1-r)P_2(1-P_2)}\right)\left(P_2' - P_1' + \int_{L_1'}^{L_2'} \frac{L}{1+e^{-L}}\,dL\right) \Big] \\
&\quad + \frac{P_2'^2(3-2P_2') - P_1'^2(3-2P_1')}{18g^2(L_2' - L_1')[rP_1(1-P_1) + (1-r)P_2(1-P_2)]}.
\end{aligned}
\tag{8.2.1}
$$

The criterion always improves as the "design to region" ratio g increases. Therefore optimal designs are found for selected values of g.

Table 8.2.1 gives Q optimal designs for symmetric regions. The designs have symmetric ED's with equal sample weights. Similar to the Q optimal designs based on the logit, these designs also become broader as the region broadens or the ratio g increases. However, unlike the logit case, the Q criterion APV[$\hat{p}$] will increase when the region begins to contract from the extreme (e.g. 0.01-0.99) and reach the worst point when the region is somewhere around 0.25-0.75 to 0.30-0.70. Additional study shows that APV[$\hat{p}$] will then start to improve as the region continues to decrease.

Since the Q criterion APV[$\hat{p}$] also possesses the property described by (8.1.5), only lower asymmetric regions are evaluated. The corresponding designs are given in table 8.2.2. These designs have symmetric ED's but uneven sample sizes weighted toward the side of the region. Designs subject to the constraint of equal sample sizes are given in table 8.2.3. These designs lose the symmetry in the ED's, which are shifted much closer to the regions. Restricting equal sample sizes only slightly hurts the Q criterion value.

Comparing the results of this section to those in §8.1, a design based on the response probability always falls inside the corresponding design based on the logit whenever the designs have symmetric ED's. For symmetric regions, the discrepancy is greater for larger regions. In the case of restricted sample sizes, the design based on the logit has slightly more extreme ED's toward the region than the corresponding design based on the response probability. In general, the two types of designs appear to be quite similar.

Table 8.2.1 Q optimal parallel-line designs based on probability, symmetric regions

Region	g	Design	Weighting(%)	APV[$\hat{p}$]
0.01-0.99	1/3	0.266305-0.733695	50-50	0.486
	1/2	0.241826-0.758174	50-50	0.327
	1	0.216171-0.783829	50-50	0.225
	2	0.206756-0.793244	50-50	0.198
0.05-0.95	1/3	0.271767-0.728233	50-50	0.727
	1/2	0.247433-0.752567	50-50	0.486
	1	0.221766-0.778234	50-50	0.331
	2	0.212297-0.787703	50-50	0.290
0.10-0.90	1/3	0.281589-0.718411	50-50	0.890
	1/2	0.257593-0.742407	50-50	0.588
	1	0.232004-0.767996	50-50	0.394
	2	0.222478-0.777522	50-50	0.344
0.15-0.85	1/3	0.293256-0.706744	50-50	0.996
	1/2	0.269789-0.730211	50-50	0.649
	1	0.244454-0.755546	50-50	0.428
	2	0.234929-0.765071	50-50	0.370
0.20-0.80	1/3	0.306496-0.693504	50-50	1.064
	1/2	0.283782-0.716218	50-50	0.683
	1	0.258943-0.741057	50-50	0.443
	2	0.249506-0.750494	50-50	0.380
0.25-0.75	1/3	0.321444-0.678556	50-50	1.104
	1/2	0.299763-0.700237	50-50	0.698
	1	0.275738-0.724262	50-50	0.444
	2	0.266509-0.733491	50-50	0.378
0.30-0.70	1/3	0.338521-0.661479	50-50	1.119
	1/2	0.318235-0.681765	50-50	0.697
	1	0.295448-0.704552	50-50	0.435
	2	0.286596-0.713404	50-50	0.368

Table 8.2.2 Q optimal parallel-line designs based on probability, asymmetric regions

Region	g	Design	Weighting(%)	APV[$\hat{p}$]
0.10-0.20	1/3	0.216340-0.783660	86.1-13.9	0.446
	1/2	0.186568-0.813432	90.6- 9.4	0.285
	1	0.163232-0.836768	93.3- 6.7	0.177
	2	0.157349-0.842651	93.7- 6.3	0.148
0.30-0.40	1/3	0.340411-0.659589	90.5- 9.5	0.948
	1/2	0.331294-0.668706	89.6-10.4	0.565
	1	0.321360-0.678640	88.3-11.7	0.332
	2	0.317432-0.682568	87.6-12.4	0.273
0.05-0.25	1/3	0.220958-0.779042	83.2-16.8	0.389
	1/2	0.194372-0.805628	86.1-13.9	0.253
	1	0.170436-0.829564	87.9-12.1	0.163
	2	0.162808-0.837192	88.2-11.8	0.139
0.10-0.30	1/3	0.243740-0.756260	85.7-14.3	0.587
	1/2	0.221148-0.778852	87.7-12.3	0.371
	1	0.201367-0.798633	88.4-11.6	0.231
	2	0.194821-0.805179	88.4-11.6	0.195
0.20-0.40	1/3	0.299439-0.700561	86.1-13.9	0.875
	1/2	0.283837-0.716163	86.0-14.0	0.535
	1	0.267985-0.732015	85.1-14.9	0.325
	2	0.262046-0.737954	84.6-15.4	0.271
0.05-0.40	1/3	0.258036-0.741964	79.8-20.2	0.566
	1/2	0.235727-0.764273	80.7-19.3	0.364
	1	0.213491-0.786509	80.8-19.2	0.234
	2	0.205497-0.794503	80.7-19.3	0.200
0.05-0.50	1/3	0.276110-0.723890	74.7-25.3	0.667
	1/2	0.254074-0.745926	74.9-25.1	0.429
	1	0.231029-0.768971	74.6-25.4	0.277
	2	0.222489-0.777511	74.4-25.6	0.238

Table 8.2.3 Q optimal parallel-line designs based on probability, asymmetric regions, restricted weighting: 50%-50%

Region	g	Design	APV[$\hat{p}$]
0.10-0.20	1/3	0.109430-0.374987	0.485
	1/2	0.090823-0.264587	0.312
	1	0.080396-0.237939	0.194
	2	0.076507-0.233153	0.163
0.30-0.40	1/3	0.278043-0.450684	0.952
	1/2	0.264196-0.452605	0.568
	1	0.248630-0.457783	0.334
	2	0.242472-0.460283	0.275
0.05-0.25	1/3	0.120302-0.458019	0.423
	1/2	0.087142-0.342813	0.287
	1	0.068933-0.290361	0.193
	2	0.063758-0.279305	0.168
0.10-0.30	1/3	0.142891-0.428458	0.615
	1/2	0.118043-0.373359	0.396
	1	0.098667-0.342781	0.254
	2	0.092281-0.335247	0.217
0.20-0.40	1/3	0.216399-0.464869	0.887
	1/2	0.194689-0.455400	0.546
	1	0.172496-0.450466	0.334
	2	0.164247-0.449349	0.280
0.05-0.40	1/3	0.164146-0.511976	0.587
	1/2	0.133923-0.479207	0.387
	1	0.108013-0.446569	0.260
	2	0.100004-0.436038	0.227
0.05-0.50	1/3	0.193931-0.560732	0.681
	1/2	0.163838-0.551092	0.445
	1	0.135746-0.537898	0.297
	2	0.126630-0.532341	0.258

§8.3 Comparison of Designs Based on Logit and on Response Probability

Q efficiencies can be computed to verify that the designs based on the logit and those based on the response probability indeed offer similar performance. The Q efficiency of a design based on the logit relative to the corresponding design based on the response probability is defined as

$$\frac{\text{APV}[\hat{p}] \text{ of the design minimizing APV}[\hat{p}]}{\text{APV}[\hat{p}] \text{ of the design minimizing APV}[\text{logit}(\hat{p})]}. \tag{8.3.1}$$

On the other hand, the Q efficiency of a design based on the response probability relative to the corresponding design based on the logit is defined as

$$\frac{\text{APV}[\text{logit}(\hat{p})] \text{ of the design minimizing APV}[\text{logit}(\hat{p})]}{\text{APV}[\text{logit}(\hat{p})] \text{ of the design minimizing APV}[\hat{p}]}. \tag{8.3.2}$$

Q efficiencies defined from the above two perspectives are evaluated in tables 8.3.1. and 8.3.2 respectively. Extremely high efficiencies are seen in most cases for both types of Q efficiency. The efficiency is often around 95% to 99%. The few worst cases still reach above 85%. A common trend revealed in both tables is that the efficiency increases as the region becomes smaller in its probability span. The efficiencies evaluated in both ways indicate that the parallel-line design based on the logit and the one based on the response probability perform roughly the same in terms of either APV[$\hat{p}$] or APV[logit($\hat{p}$)].

The prevailing high efficiencies in the comparison of the two types of designs provide convincing evidence that the Q optimality criterion based on the logit and the one based on the response probability are likely to produce fairly close results in design optimization for the logistic model. Using the logit instead of the response probability significantly reduces the complexity of the design work dealing with Q optimality, especially for a relatively complicated model such as the two-variable model containing interaction. Meanwhile the logit itself bears its practical importance as it intrinsically relates to the concept of tolerance in the study of binary responses. Q optimal designs for

Table 8.3.1 Q efficiency of design based on logit relative to design based on probability

		g: 1/3	1/2	1	2
Symmetric region	0.01-0.99	0.8935	0.8866	0.8863	0.8884
	0.05-0.95	0.9662	0.9627	0.9608	0.9607
	0.10-0.90	0.9877	0.9860	0.9848	0.9846
	0.15-0.85	0.9951	0.9944	0.9937	0.9935
	0.20-0.80	0.9981	0.9978	0.9974	0.9973
	0.25-0.75	0.9993	0.9992	0.9990	0.9990
	0.30-0.70	0.9998	0.9997	0.9997	0.9997
Asymmetric region	0.10-0.20	0.9990	0.9981	0.9961	0.9950
	0.30-0.40	0.9999	0.9998	0.9997	0.9997
	0.05-0.25	0.9840	0.9773	0.9650	0.9581
	0.10-0.30	0.9934	0.9899	0.9837	0.9807
	0.20-0.40	0.9980	0.9971	0.9957	0.9951
	0.05-0.40	0.9604	0.9468	0.9253	0.9148
	0.05-0.50	0.9496	0.9333	0.9091	0.8979
Asymmetric region (restricted weighting: 50%-50%)	0.10-0.20	0.9990	0.9962	0.9915	0.9896
	0.30-0.40	0.9999	0.9998	0.9997	0.9997
	0.05-0.25	0.9788	0.9878	0.9402	0.9194
	0.10-0.30	0.9948	0.9886	0.9761	0.9704
	0.20-0.40	0.9980	0.9969	0.9952	0.9944
	0.05-0.40	0.9730	0.9772	0.9592	0.9376
	0.05-0.50	0.9686	0.9726	0.9799	0.9761

Table 8.3.2 Q efficiency of design based on probability relative to design based on logit

		g: 1/3	1/2	1	2
Symmetric region	0.01-0.99	0.8595	0.8632	0.8749	0.8811
	0.05-0.95	0.9580	0.9560	0.9563	0.9572
	0.10-0.90	0.9855	0.9841	0.9833	0.9833
	0.15-0.85	0.9945	0.9938	0.9932	0.9930
	0.20-0.80	0.9979	0.9976	0.9973	0.9972
	0.25-0.75	0.9993	0.9991	0.9990	0.9989
	0.30-0.70	0.9998	0.9997	0.9997	0.9997
Asymmetric region	0.10-0.20	0.9989	0.9980	0.9960	0.9950
	0.30-0.40	0.9999	0.9998	0.9997	0.9997
	0.05-0.25	0.9809	0.9733	0.9621	0.9569
	0.10-0.30	0.9925	0.9890	0.9834	0.9809
	0.20-0.40	0.9979	0.9970	0.9957	0.9951
	0.05-0.40	0.9501	0.9382	0.9231	0.9166
	0.05-0.50	0.9364	0.9235	0.9076	0.9011
Asymmetric region (restricted weighting: 50%-50%)	0.10-0.20	0.9989	0.9963	0.9923	0.9906
	0.30-0.40	0.9999	0.9998	0.9997	0.9997
	0.05-0.25	0.9781	0.9874	0.9652	0.9564
	0.10-0.30	0.9938	0.9892	0.9804	0.9765
	0.20-0.40	0.9979	0.9969	0.9954	0.9947
	0.05-0.40	0.9630	0.9674	0.9705	0.9681
	0.05-0.50	0.9540	0.9591	0.9705	0.9746

the interaction model in the two-variable logistic case are derived based on the logit in the following section.

§8.4 Hyperbola-Based Design

Given the two-variable logistic model containing interaction, constant response probabilities fall on hyperbolic ED curves. Consider a region R which is a parallelogram with the vertices on $ED_{100P_1'}$ and $ED_{100P_2'}$. The opposite vertices on a common ED are symmetric about the hyperbola center. The preliminary design layout is a hyperbola-based design on ED_{100P_1} and ED_{100P_2}. The design points on a common ED are symmetric about the hyperbola center and assumed to have equal sample sizes. Figure 8.4.1 illustrates the region and the design in the centered space.

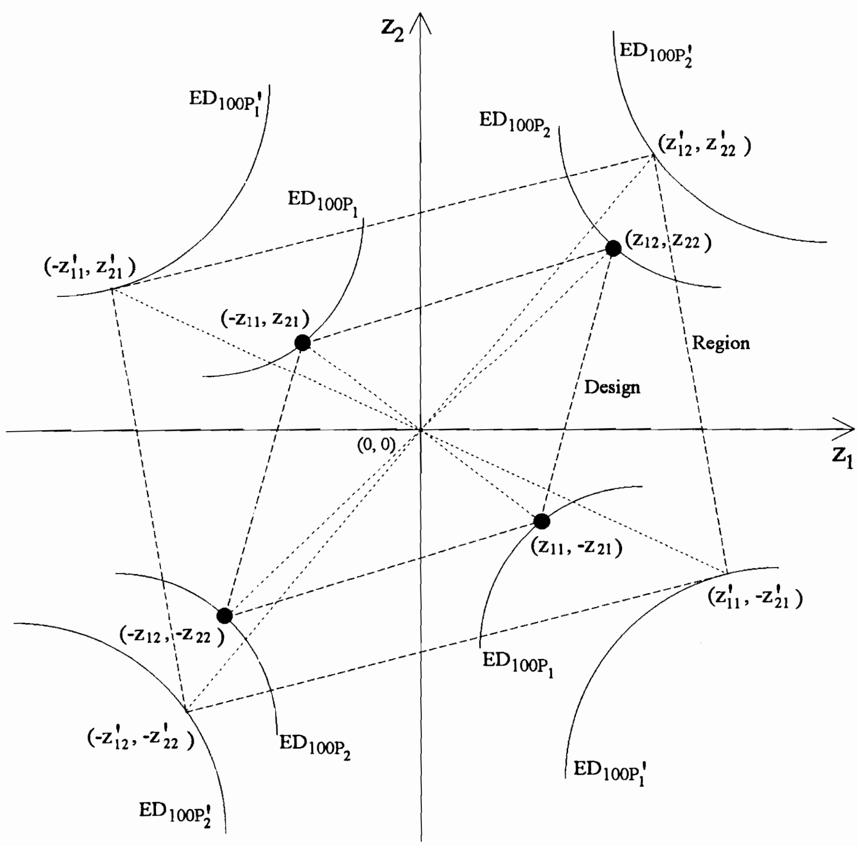


Figure 8.4.1 Region and design arrangement for Q optimal design, interaction model

Given the centered levels of the design and region vertices as shown in figure 8.4.1, define two sets of ratios:

$$(1) \quad t_1 = \left|\frac{z_{11}}{z_{12}}\right|, \quad t_1' = \left|\frac{z_{11}'}{z_{12}'}\right|, \quad g_1 = \left|\frac{z_{11}}{z_{11}'}\right|, \tag{8.4.1}$$

and

$$(2) \quad t_2 = \left|\frac{z_{21}}{z_{22}}\right|, \quad t_2' = \left|\frac{z_{21}'}{z_{22}'}\right|, \quad g_2 = \left|\frac{z_{21}}{z_{21}'}\right|. \tag{8.4.2}$$

The Q optimality criterion is derived in appendix C and is given by

$$\begin{aligned}
APV[\text{logit}(\hat{p})] = {} & \frac{N}{12[t(L_2 - \beta_0^*) - \frac{1}{t}(L_1 - \beta_0^*)]^2} \\
& \cdot[\frac{1}{\sigma_1^2}((\frac{t}{g}(L_2 - \beta_0^*) - \frac{g}{t}(L_1' - \beta_0^*))^2 + (\frac{t}{gt'}(L_2 - \beta_0^*) - \frac{gt'}{t}(L_2' - \beta_0^*))^2) \\
& + \frac{1}{\sigma_2^2}((\frac{1}{g}(L_1 - \beta_0^*) - g(L_1' - \beta_0^*))^2 + (\frac{1}{gt'}(L_1 - \beta_0^*) - gt'(L_2' - \beta_0^*))^2)] \\
& + \frac{N}{4(L_2 - L_1)^2}[\frac{1}{\sigma_1^2}(((L_2 - \beta_0^*) - \frac{1}{3}(L_1' - \beta_0^*))^2 + ((L_2 - \beta_0^*) - \frac{1}{3}(L_2' - \beta_0^*))^2) \\
& + \frac{1}{\sigma_2^2}(((L_1 - \beta_0^*) - \frac{1}{3}(L_1' - \beta_0^*))^2 + ((L_1 - \beta_0^*) - \frac{1}{3}(L_2' - \beta_0^*))^2)] \\
& + \frac{N}{6(L_2 - L_1)^2}(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2})[(L_1' + L_2' - 2\beta_0^*)^2 + (t'(L_2' - \beta_0^*) + \frac{1}{t'}(L_1' - \beta_0^*))^2]
\end{aligned} \tag{8.4.3}$$

where $\beta_0^* = \beta_0 - \frac{\beta_1\beta_2}{\beta_{12}}$, $L_1 = \text{logit}(P_1)$, $L_2 = \text{logit}(P_2)$, $L_1' = \text{logit}(P_1')$, and $L_2' = \text{logit}(P_2')$, $\sigma_1^2 = n_1P_1(1-P_1)$, $\sigma_2^2 = n_2P_2(1-P_2)$, and the ratios t, t′, and g can be given by either set (1) in (8.4.1), i.e. $t = t_1$, $t' = t_1'$, and $g = g_1$, or set (2) in (8.4.2), i.e. $t = t_2$, $t' = t_2'$, and $g = g_2$. Essentially, the Q criterion in (8.4.3) can be viewed as a function of the parameter β_0^*, the region attributes P_1', P_2', and t′, the design attributes P_1, P_2, r, and t, and the "design to region" ratio g, where r is the sample proportion allocated to ED_{100P_1}.

The Q criterion in (8.4.3), if denoted by $F(P_1, P_2, r, g, t, t', P_1', P_2', \beta_0^*)$, possesses the property

$$F(P_1, P_2, r, g, t, t', P_1', P_2', \beta_0^*) = F(1\text{-}P_2, 1\text{-}P_1, 1\text{-}r, \frac{t'}{t}g, \frac{1}{t}, \frac{1}{t'}, 1\text{-}P_2', 1\text{-}P_1', \text{-}\beta_0^*). \tag{8.4.4}$$

The "design to region" ratio g specifies the relative magnitude of the centered design level compared to the centered region level at the respective vertex. Algebraic work shows that an optimal solution for the ratio g does exit and can be expressed as

$$g = \left(\frac{(1+\frac{1}{t'^2})[\sigma_1^2(L_1-\beta_0^*)^2 + t^2\sigma_2^2(L_2-\beta_0^*)^2]}{(\sigma_1^2+\frac{\sigma_2^2}{t^2})[(L_1'-\beta_0^*)^2 + t'^2(L_2'-\beta_0^*)^2]} \right)^{\frac{1}{4}}. \tag{8.4.5}$$

On the other hand, numerical work confirms that minimizing APV[logit($\hat{p}$)] given β_0^*, P_1', P_2', and t' leads to extreme values of t, i.e. $t \to 0$ or $t \to \infty$. Similar to the D optimal design, the Q optimal design again will be assumed to have $t = 1$. To preserve the resemblance between the region and design structures, the region is assumed to have $t' = 1$. With $t = t' = 1$, both the region and the design possess the kind of geometry displayed in figure 2.2.2.

With $t = t' = 1$, the Q criterion is simplified to

$$\text{APV}[\text{logit}(\hat{p})] = Q(P_1, P_2, r, g \mid P_1', P_2', \beta_0^*)$$

$$= \frac{1}{6(L_2-L_1)^2}$$

$$\cdot\{\frac{1}{rP_1(1-P_1)}[(\frac{1}{g}(L_2 - \beta_0^*) - g(L_1' - \beta_0^*))^2 + (\sqrt{3}(L_2 - \beta_0^*) - \frac{1}{\sqrt{3}}(L_1' - \beta_0^*))^2$$

$$+(\frac{1}{g}(L_2-\beta_0^*)-g(L_2'-\beta_0^*))^2+(\sqrt{3}(L_2-\beta_0^*)-\frac{1}{\sqrt{3}}(L_2'-\beta_0^*))^2]$$

$$+\frac{1}{(1-r)P_2(1-P_2)}[(\frac{1}{g}(L_1-\beta_0^*)-g(L_1'-\beta_0^*))^2+(\sqrt{3}(L_1-\beta_0^*)-\frac{1}{\sqrt{3}}(L_1'-\beta_0^*))^2$$

$$+(\frac{1}{g}(L_1-\beta_0^*)-g(L_2'-\beta_0^*))^2+(\sqrt{3}(L_1-\beta_0^*)-\frac{1}{\sqrt{3}}(L_2'-\beta_0^*))^2]$$

$$+4(\frac{1}{rP_1(1-P_1)}+\frac{1}{(1-r)P_2(1-P_2)})(L_1'+L_2'-2\beta_0^*)^2\}. \tag{8.4.6}$$

The expression for the optimal g reduces to

$$g=\left(\frac{2[\sigma_1^2(L_1-\beta_0^*)^2+\sigma_2^2(L_2-\beta_0^*)^2]}{(\sigma_1^2+\sigma_2^2)[(L_1'-\beta_0^*)^2+(L_2'-\beta_0^*)^2]}\right)^{\frac{1}{4}}. \tag{8.4.7}$$

The property in (8.4.4) is rewritten as

$$Q(P_1, P_2, r, g\,|\,P_1', P_2', \beta_0^*) = Q(1-P_2, 1-P_1, 1-r, g\,|\,1-P_1', 1-P_2', -\beta_0^*). \tag{8.4.8}$$

The property suggests that solutions of optimal designs for $\beta_0^*<0$ can be easily extended to the solutions for $\beta_0^*<0$. Therefore the designs are found only for the $\beta_0^*\leq 0$ case. The selected values of β_0^* 0, -1, -2, and -5 correspond to the response probabilities 0.5, 0.2689, 0.1192, and 0.0067 at the hyperbola center, which are given by $P_0^* = \frac{1}{1+e^{-\beta_0^*}}$. These probabilities serve as easy references as to where the obtained optimal design falls with respect to the hyperbola center. Minimizing $Q(P_1, P_2, r, g\,|\,P_1', P_2', \beta_0^*)$ given the parameter β_0^* and the region specifications P_1' and P_2' yields the optimal design attributes P_1, P_2, r, and g. The optimal value for g from numerical minimization agrees with the analytical result given by (8.4.7).

Table 8.4.1 gives the optimal designs for symmetric regions. These designs have symmetric ED's but uneven sample sizes slightly weighted toward the direction of the response probability P_0^* at the hyperbola center. The APV improves as the region dispersion decreases or as β_0^* draws close to 0, which means that the centered axes approach ED_{50}.

The optimal designs for lower and upper asymmetric regions are given in tables 8.4.2 and 8.4.3 respectively. These designs also have symmetric ED's. The ED's for a lower tail region are not really the same as those for the opposite upper tail region. The sample weights are driven on one hand toward the direction of P_0^* and on the other hand toward the side of the region. Consequently, given $P_0^* \leq 0.5$, the lower asymmetric regions sometimes yield fairly unbalanced sample sizes weighted at the lower ED due to the combining forces of P_0^* and the region. However, the upper asymmetric regions often lead to only slightly uneven sample weights as the central probability P_0^* and the region at the opposite tail counteract each other in driving the sample weights. Only when $\beta_0^* = 0$, the lower and upper regions give equally shifted sample sizes weighted on opposite sides.

Designs subject to equal sample sizes are also obtained and given in tables 8.4.4 through 8.4.6 for the symmetric, lower asymmetric, and upper asymmetric regions respectively. In a majority of the cases, the designs have slightly asymmetric ED's that are shifted away from P_0^* and the region. In a few cases with lower tail regions where the optimal sample weights used to be the most unbalanced in the unrestricted minimization, the design under the equal sample size restriction does not follow the above pattern and settles down with highly asymmetric ED's in a neighborhood closely around the region. The few exceptions are possibly because the equal sample size condition deviates too far from the optimal structure for the design to maintain its balance by minimal adjustment

Table 8.4.1 Q optimal designs for interaction model, symmetric regions

β_0^*	Region	Design	Weighting (%)	g	APV
0	0.01-0.99	0.176430-0.823570	50-50	0.57905	27.4
	0.05-0.95	0.200315-0.799685	50-50	0.68567	18.2
	0.10-0.90	0.216949-0.783051	50-50	0.76431	14.5
	0.15-0.85	0.230692-0.769308	50-50	0.83327	12.3
	0.20-0.80	0.243834-0.756166	50-50	0.90355	10.8
	0.25-0.75	0.257457-0.742543	50-50	0.98191	9.5
	0.30-0.70	0.272509-0.727491	50-50	1.07652	8.4
-1	0.01-0.99	0.149156-0.850844	57.4-42.6	0.63161	36.0
	0.05-0.95	0.157610-0.842390	57.5-42.5	0.76455	26.9
	0.10-0.90	0.161748-0.838252	57.5-42.5	0.86180	23.4
	0.15-0.85	0.164303-0.835697	57.4-42.6	0.94267	21.5
	0.20-0.80	0.166157-0.833843	57.3-42.7	1.01763	20.2
	0.25-0.75	0.167586-0.832414	57.2-42.8	1.08959	19.2
	0.30-0.70	0.168708-0.831292	57.0-43.0	1.15847	18.5
-2	0.01-0.99	0.120036-0.879964	57.8-42.2	0.71934	59.5
	0.05-0.95	0.122222-0.877778	57.5-42.5	0.85296	50.0
	0.10-0.90	0.123105-0.876895	57.3-42.7	0.93383	46.5
	0.15-0.85	0.123584-0.876416	57.1-42.9	0.98958	44.6
	0.20-0.80	0.123897-0.876103	57.0-43.0	1.03244	43.4
	0.25-0.75	0.124116-0.875884	56.9-43.1	1.06642	42.5
	0.30-0.70	0.124274-0.875726	56.9-43.1	1.09324	41.9
-5	0.01-0.99	0.092834-0.907166	54.7-45.3	0.88325	211.9
	0.05-0.95	0.092998-0.907002	54.5-45.5	0.95602	200.7
	0.10-0.90	0.093053-0.906947	54.5-45.5	0.98562	196.9
	0.15-0.85	0.093080-0.906920	54.4-45.6	1.00136	195.1
	0.20-0.80	0.093096-0.906904	54.4-45.6	1.01139	194.0
	0.25-0.75	0.093107-0.906893	54.4-45.6	1.01826	193.2
	0.30-0.70	0.093115-0.906885	54.4-45.6	1.02310	192.7

Table 8.4.2 Q optimal designs for interaction model, lower asymmetric regions

β_0^*	Region	Design	Weighting (%)	g	APV
0	0.10-0.20	0.132271-0.867729	57.8-42.2	1.01189	35.5
	0.30-0.40	0.207121-0.792879	58.2-41.8	1.42163	11.5
	0.05-0.25	0.126277-0.873723	57.4-42.6	0.93296	42.8
	0.10-0.30	0.141688-0.858312	58.0-42.0	1.04008	29.0
	0.20-0.40	0.178102-0.821898	58.3-41.7	1.22365	16.0
	0.05-0.40	0.136386-0.863614	57.5-42.5	0.93711	34.0
	0.05-0.50	0.143863-0.856137	57.5-42.5	0.92555	29.7
-1	0.10-0.20	0.153315-0.846685	72.1-27.9	1.32102	17.5
	0.30-0.40	0.220697-0.779303	72.1-27.9	1.67359	9.1
	0.05-0.25	0.141689-0.858311	69.2-30.8	1.10881	23.3
	0.10-0.30	0.171822-0.828178	75.2-24.8	1.26910	13.8
	0.20-0.40	0.230403-0.769597	81.3-18.7	1.39256	8.2
	0.05-0.40	0.159028-0.840972	71.7-28.3	1.02987	18.2
	0.05-0.50	0.170662-0.829338	72.9-27.1	0.96220	16.3
-2	0.10-0.20	0.142236-0.857764	87.1-12.9	1.73908	12.8
	0.30-0.40	0.137027-0.862973	62.6-37.4	1.30056	27.1
	0.05-0.25	0.133563-0.866437	85.0-15.0	1.27624	16.0
	0.10-0.30	0.147457-0.852543	78.9-21.1	1.44948	15.7
	0.20-0.40	0.142346-0.857654	66.4-33.6	1.35196	22.7
	0.05-0.40	0.142014-0.857986	79.4-20.6	1.15070	17.8
	0.05-0.50	0.143111-0.856889	75.2-24.8	1.10058	19.8
-5	0.10-0.20	0.096432-0.903568	59.2-40.8	1.25413	108.8
	0.30-0.40	0.094354-0.905646	55.7-44.3	1.09478	158.4
	0.05-0.25	0.096614-0.903386	60.2-39.8	1.27197	101.9
	0.10-0.30	0.096017-0.903983	58.3-41.7	1.20346	119.2
	0.20-0.40	0.094883-0.905117	56.4-43.6	1.12352	145.7
	0.05-0.40	0.096163-0.903837	58.8-41.2	1.19782	115.1
	0.05-0.50	0.095827-0.904173	58.1-41.9	1.15941	123.5

Table 8.4.3 Q optimal designs for interaction model, upper asymmetric regions

β_0^*	Region	Design	Weighting (%)	g	APV
0	0.80-0.90	0.132271-0.867729	42.2-57.8	1.01189	35.5
	0.60-0.70	0.207121-0.792879	41.8-58.2	1.42163	11.5
	0.75-0.95	0.126277-0.873723	42.6-57.4	0.93296	42.8
	0.70-0.90	0.141688-0.858312	42.0-58.0	1.04008	29.0
	0.60-0.80	0.178102-0.821898	41.7-58.3	1.22365	16.0
	0.60-0.95	0.136386-0.863614	42.5-57.5	0.93711	34.0
	0.50-0.95	0.143863-0.856137	42.5-57.5	0.92555	29.7
-1	0.80-0.90	0.114885-0.885115	48.3-51.7	0.90366	64.5
	0.60-0.70	0.141320-0.858680	51.4-48.6	1.11429	29.6
	0.75-0.95	0.111959-0.888041	48.2-51.8	0.85929	73.9
	0.70-0.90	0.119142-0.880858	48.7-51.3	0.93021	55.3
	0.60-0.80	0.132880-0.867120	50.3-49.7	1.04065	36.4
	0.60-0.95	0.116808-0.883192	48.6-51.4	0.87771	61.6
	0.50-0.95	0.120205-0.879795	49.1-50.9	0.88320	55.2
-2	0.80-0.90	0.103674-0.896326	50.8-49.2	0.87473	106.3
	0.60-0.70	0.114919-0.885081	53.6-46.4	1.02209	59.5
	0.75-0.95	0.102188-0.897812	50.5-49.5	0.84375	118.0
	0.70-0.90	0.105709-0.894291	51.3-48.7	0.89748	94.3
	0.60-0.80	0.111665-0.888335	52.7-47.3	0.97676	69.0
	0.60-0.95	0.104577-0.895423	51.1-48.9	0.86405	102.1
	0.50-0.95	0.106174-0.893826	51.5-48.5	0.87367	93.6
-5	0.80-0.90	0.090240-0.909760	52.1-47.9	0.89230	314.7
	0.60-0.70	0.091993-0.908007	53.4-46.6	0.97526	230.3
	0.75-0.95	0.089948-0.910052	52.0-48.0	0.87533	333.9
	0.70-0.90	0.090604-0.909396	52.4-47.6	0.90792	294.0
	0.60-0.80	0.091545-0.908455	53.1-46.9	0.95292	248.5
	0.60-0.95	0.090395-0.909605	52.3-47.7	0.89242	306.7
	0.50-0.95	0.090673-0.909327	52.5-47.5	0.90212	291.7

Table 8.4.4 Q optimal designs for interaction model, symmetric regions, restricted weighting: 50%-50%

β_0^*	Region	Design	g	APV
0	0.01-0.99	0.176430-0.823570	0.57905	27.4
	0.05-0.95	0.200315-0.799685	0.68567	18.2
	0.10-0.90	0.216949-0.783051	0.76431	14.5
	0.15-0.85	0.230692-0.769308	0.83327	12.3
	0.20-0.80	0.243834-0.756166	0.90355	10.8
	0.25-0.75	0.257457-0.742543	0.98191	9.5
	0.30-0.70	0.272509-0.727491	1.07652	8.4
-1	0.01-0.99	0.158915-0.866536	0.65461	36.6
	0.05-0.95	0.165564-0.856199	0.79456	27.5
	0.10-0.90	0.168347-0.850154	0.89579	23.9
	0.15-0.85	0.169903-0.846044	0.97943	21.9
	0.20-0.80	0.170956-0.842878	1.05664	20.6
	0.25-0.75	0.171728-0.840333	1.13057	19.6
	0.30-0.70	0.172312-0.838277	1.20121	18.8
-2	0.01-0.99	0.133601-0.896506	0.74425	60.6
	0.05-0.95	0.135488-0.893786	0.88143	50.8
	0.10-0.90	0.136122-0.892479	0.96409	47.2
	0.15-0.85	0.136424-0.891702	1.02098	45.3
	0.20-0.80	0.136601-0.891166	1.06468	44.0
	0.25-0.75	0.136715-0.890774	1.09931	43.2
	0.30-0.70	0.136791-0.890484	1.12664	42.5
-5	0.01-0.99	0.100526-0.914887	0.89360	213.1
	0.05-0.95	0.100470-0.914502	0.96684	201.7
	0.10-0.90	0.100439-0.914360	0.99664	197.9
	0.15-0.85	0.100422-0.914287	1.01247	196.0
	0.20-0.80	0.100410-0.914242	1.02256	194.9
	0.25-0.75	0.100402-0.914212	1.02948	194.1
	0.30-0.70	0.100396-0.914191	1.03435	193.6

Table 8.4.5 Q optimal designs for interaction model, lower asymmetric regions, restricted weighting: 50%-50%

β_0^*	Region	Design	g	APV
0	0.10-0.20	0.146911-0.884892	1.01430	36.2
	0.30-0.40	0.179695-0.751582	1.40151	11.7
	0.05-0.25	0.140403-0.889785	0.93454	43.5
	0.10-0.30	0.155718-0.875488	1.04341	29.6
	0.20-0.40	0.176495-0.819571	1.22267	16.5
	0.05-0.40	0.150294-0.880123	0.93950	34.7
	0.05-0.50	0.156759-0.871908	0.92830	30.3
-1	0.10-0.20	0.183284-0.906176	1.50169	20.6
	0.30-0.40	0.130650-0.557858	1.61849	9.9
	0.05-0.25	0.171086-0.907615	1.22149	26.1
	0.10-0.30	0.192550-0.889727	1.51179	17.1
	0.20-0.40	0.127237-0.445787	1.28620	8.8
	0.05-0.40	0.184294-0.897857	1.17676	21.3
	0.05-0.50	0.187788-0.882456	1.12362	19.5
-2	0.10-0.20	0.049178-0.232565	1.35819	16.9
	0.30-0.40	0.157988-0.891840	1.38688	28.4
	0.05-0.25	0.042787-0.214400	0.92936	22.9
	0.10-0.30	0.188093-0.931842	1.75324	20.0
	0.20-0.40	0.167834-0.896818	1.47985	24.7
	0.05-0.40	0.183085-0.933857	1.39445	22.5
	0.05-0.50	0.179010-0.921518	1.28755	23.8
-5	0.10-0.20	0.112083-0.919398	1.28498	111.1
	0.30-0.40	0.103841-0.915182	1.11057	159.7
	0.05-0.25	0.113825-0.920803	1.30677	104.6
	0.10-0.30	0.109965-0.918069	1.22958	121.3
	0.20-0.40	0.105539-0.915842	1.14183	147.2
	0.05-0.40	0.111030-0.918859	1.22572	117.3
	0.05-0.50	0.109474-0.917947	1.18403	125.6

Table 8.4.6 Q optimal designs for interaction model, upper asymmetric regions, restricted weighting: 50%-50%

β_0^*	Region	Design	g	APV
0	0.80-0.90	0.115108-0.853089	1.01430	36.2
	0.60-0.70	0.248418-0.820305	1.40151	11.7
	0.75-0.95	0.110215-0.859597	0.93454	43.5
	0.70-0.90	0.124512-0.844282	1.04341	29.6
	0.60-0.80	0.180429-0.823505	1.22267	16.5
	0.60-0.95	0.119877-0.849706	0.93950	34.7
	0.50-0.95	0.128092-0.843241	0.92830	30.3
-1	0.80-0.90	0.111624-0.881748	0.89922	64.6
	0.60-0.70	0.143842-0.861533	1.12056	29.6
	0.75-0.95	0.108441-0.884411	0.85480	74.0
	0.70-0.90	0.116653-0.878268	0.92659	55.4
	0.60-0.80	0.133412-0.867697	1.04169	36.4
	0.60-0.95	0.114170-0.880446	0.87414	61.7
	0.50-0.95	0.118359-0.877854	0.88059	55.3
-2	0.80-0.90	0.105105-0.897825	0.87718	106.3
	0.60-0.70	0.121563-0.892352	1.03655	59.7
	0.75-0.95	0.103142-0.898809	0.84531	118.0
	0.70-0.90	0.107970-0.896673	0.90149	94.3
	0.60-0.80	0.116635-0.893688	0.98678	69.2
	0.60-0.95	0.106499-0.897445	0.86732	102.1
	0.50-0.95	0.108823-0.896629	0.87828	93.7
-5	0.80-0.90	0.093702-0.913232	0.89693	315.1
	0.60-0.70	0.097577-0.913604	0.98344	231.0
	0.75-0.95	0.093116-0.913228	0.87949	334.2
	0.70-0.90	0.094474-0.913276	0.91319	294.4
	0.60-0.80	0.096543-0.913465	0.96007	249.1
	0.60-0.95	0.094056-0.913276	0.89733	307.1
	0.50-0.95	0.094661-0.913325	0.90751	292.1

without a drastic change in its ED's. In virtually every case, restricting equal sample sizes only causes minimal or negligible increase in the APV.

Either restricting or not restricting equal sample sizes, the APV is considerably worse for the upper asymmetric regions than for the lower ones. Given $\beta_0^* \leq 0$, this means that a region close to the centered axes in terms of the logit or ED's generally allows better possible results. In all cases, the optimal "design to region" ratio g is usually found in a vicinity around 1.

To construct a hyperbola-based Q optimal design, the equidistant design level needs to be obtained from (8.4.1) or (8.4.2) using the region vertices and the optimal ratio g. The design points can then be computed using the design matrix in (2.2.7), with -L and L replaced by L_1 and L_2 in case of asymmetric design ED's.

CHAPTER 9

SUMMARY AND FUTURE RESEARCH

The primary goal of this research was to develop and study efficient and practical experimental design procedures for fitting the two-variable logistic regression models.

D and Q optimalities were addressed in this work. Within each criterion, the presence or absence of interaction plays a role in separating the design work into two categories. For each model, unlike the one-variable case, different design structures can be assumed in the two-dimensional design space. For example, the hyperbola-based design and the modified ray design were proposed for the interaction model. The equivalence property of a D optimal design was studied in several cases. The efficiencies of various designs, including some by other authors, were investigated versus the hyperbola-based design in the presence of interaction.

Designs were also developed to optimize specifically the estimation of the interaction coefficient. Comparison of different designs were presented through their relative efficiencies.

Robustness to parameter misspecification was studied for various D optimal designs. In an effort to better cope with poor parameter guesses, Bayesian design procedures were explored using D optimality. For a reasonable cost of losing some advantage under perfect parameter knowledge, the Bayesian designs were found to have improved robustness over the conventional D optimal designs.

Design work dealing with two variables and their possible interaction for a nonlinear model such as the logistic case may involve diverse issues. Some areas of extension and interest arising from this research are given below.

(1) In addition to the three level designs considered with the Bayesian procedure, designs with more than two ED levels can also be studied more thoroughly in the non-Bayesian environment. As suggested by other optimal design work, designs with multiple levels tend to be more robust and are often more appealing to practitioners. Nevertheless, for the no-interaction model, a four-point design on two ED's in the two-variable case seems to have better value compared to a two-point one-variable design. The two-variable design, even with only two ED's, still offers one lack-of-fit degree of freedom for testing interaction.

(2) In this research, D optimality received primary attention in the robustness analysis and the development of the Bayesian procedures. Bayesian designs can further be explored for the interaction model. In addition, the same directions of robustness analysis and robust design procedures can also be pursued with other design criteria such as Q optimality, which applies to the designs in chapter 8.

(3) Additional design criteria could be studied. For instance, G optimality optimizes the prediction variance through a minimax approach rather than averaging. However, it is anticipated that the G optimal design work is likely to require more constraints and may have to involve frequent numerical approximation.

(4) As stated earlier, a one stage design at its best can only achieve a customized trade-off between favoring good parameter knowledge and being insensitive to poor parameter guesses. One approach that does so is the Bayesian design procedure. To better avoid the trade-off, two-stage designs in the two-variable case might be considered.

Conceptually, a two-stage design procedure should more effectively separate the goals addressing good versus poor parameter knowledge. The first stage emphasizes robustness to parameter misspecification while the second stage strives at being highly efficient given the relatively more reliable knowledge about the parameters obtained from the first stage data. The remaining compromise is partly reflected in the sample allocation between the two stages. A two-stage procedure gains its advantage over the one-stage design by alleviating some of the dependence on the parameter guesses. Based upon other two-stage design work for the one-variable logistic model by Myers, Myers, and Carter (1994), a two-stage procedure in general is expected to yield efficient designs which are often considerably more robust than the single stage designs. With the two-variable logistic models, detailed constraints might be required in a study of two-stage designs.

(5) The two-variable design work should be extended to other nonlinear models. The primary interest normally lies in the generalized linear models for responses from the exponential families such as Poisson or exponential distributions.

BIBLIOGRAPHY

Abdelbasit, K. M. and Plackett, R. L. (1983), "Experimental Design for Binary Data," *Journal of the American Statistical Association*, **78**, 90-98.

Atkinson, A. C. (1982), "Development in the Design of Experiments," *International Statistical Review*, **50**, 161-177.

Box, G. E. P. and Draper, N. R. (1959), "A Basis for the Selection of a Response Surface Design," *Journal of the American Statistical Association*, **54**, 622-654.

Brunden, M. N., Vidmar, T. J., and McKean J. W. (1988), *Drug Interaction and Lethality Analysis*, Boca Raton: CRC Press.

Chaloner, K. and Larntz, K. (1989), "Optimal Bayesian Design Applied to Logistic Regression Experiments," *Journal of Statistical Planning and Inference*, **21**, 191-208.

Durkin, P.R. (1981), "Approach to the Analysis of Toxicant Interactions in the Aquatic Environment," *American Society for Testing and Materials* STP 737, 388-401.

Federov, V. V. (1972), *Theory of Optimal Experiments*, New York: Academic Press.

Finney, D. J. (1978), *Statistical Method in Biological Assays*, London: Charles Griffen.

Kalish, L. A. and Rosenberger, J. L. (1978), "Optimal Designs for the Estimation of the Logistic Function," Technical Report 33, Department of Statistics, Pennsylvania State University.

Kiefer, J. and Wolfowitz, J. (1959), "Optimal Designs in Regression Problems," *Annals of Mathematical Statistics*, **30**, 271-294.

Lehmann, E. L. (1983), *Theory of Point Estimation*, New York: John Wiley & Sons.

Letsinger, W. C. (1995), "Optimal One and Two-Stage Designs for the Logistic Regression Model," unpublished Ph.D. dissertation, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, Department of Statistics.

Mantel, N. (1958), "An Experimental Design in Combination Chemotherapy," *Annals of the New York Academy of Sciences*, **76**, 909-931.

Minkin, S. (1987), "Optimal Designs for Binary Data," *Journal of the American Statistical Association*, **82**, 1098-1103.

Myers, R. H. (1976), *Response Surface Methodology*, Blacksburg, Virginia: Author.

Myers, R. H., Khuri, A. I., and Carter, W. H., Jr. (1989), "Response Surface Methodology: 1966-1988," *Technometrics*, **31**, 137-157.

Myers, R. H. and Montgomery, D. C. (1995), *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, New York: John Wiley & Sons.

Myers, W. R. (1991), "Optimal Experimental Design for Fitting the Logistic Regression Model," unpublished Ph.D. dissertation, Virginia Commonwealth University, Richmond, Virginia, Department of Biostatistics.

Myers, W. R., Myers, R. H., and Carter, W. H., Jr. (1994), "Some Alphabetic Optimal Designs for the Logistic Regression Model," *Journal of Statistical Planning and Inference*, **42**, 57-77.

Myers, W. R., Myers, R. H., Carter, W. H., Jr., and White, K. L. (1996), "Two-Stage Designs for the logistic Regression Model in Single Agent Bioassay," *Journal of Biopharmaceutical Statistics*, Vol. **6**, NO. **4**, to appear.

Nelder, J. A. and Mead, R. (1965), "A Simplex Method for Function Minimization," *Computer Journal*, **7**, 308-313.

Sitter, R. R. (1992), "Robust Designs for Binary Data," *Biometrics*, **48**, 1145-1155.

White, L. V. (1973), "An Extension of the General Equivalence Theorem to Nonlinear Models," *Biometrika*, **60**, 345-348.

Appendix A Q Optimality Criterion Based on Logit for Parallel-Line Design

Given the parallel-line design in figure 8.1.1, the model matrix for fitting the no-interaction model is given by

$$X = ZA \tag{A1}$$

where

$$Z = \begin{bmatrix} 1 & -1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 1 & x_{10} & x_{20} \\ 0 & D_1 & D_2 \\ 0 & d_1 & d_2 \end{bmatrix}, \tag{A2}$$

where the design center (x_{10}, x_{20}) and the distances D_1, D_2, d_1, and d_2 are as shown in figure 8.1.1. The variance-covariance matrix of the response vector from the four design points is given by

$$\Sigma = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_1^2 & 0 \\ 0 & 0 & 0 & \sigma_2^2 \end{bmatrix} \tag{A3}$$

where $\sigma_1^2 = n_1P_1(1\text{-}P_1)$ and $\sigma_2^2 = n_2P_2(1\text{-}P_2)$, where n_1 and n_2 are the sample sizes at each point on ED_{100P_1} and ED_{100P_2} respectively.

Since $\text{logit}(\hat{p}) = x'\hat{\beta}$ and $I(\beta) = X'\Sigma X$, the Q criterion can be written as

$$APV[\text{logit}(\hat{p})] = \frac{N}{K}\int_R x'(X'\Sigma X)^{-1}x \, dx = N\cdot\text{Trace}[I(\beta)^{-1}S] \tag{A4}$$

where N is the total sample size, R is the region of interest, K is the area of R, and S is the region moment matrix given by

$$S = (s_{ij})_{3\times 3} = \frac{1}{K}\int_R x\,x'dx = \frac{1}{K}\iint_{(R)} \begin{bmatrix} 1 & x_1 & x_2 \\ x_1 & x_1^2 & x_1x_2 \\ x_2 & x_1x_2 & x_2^2 \end{bmatrix} dx_1dx_2\,. \tag{A5}$$

Due to (A1), $I(\beta)^{-1}$ can be written as

$$I(\beta)^{-1} = (V_{ij})_{3\times 3} = A^{-1}(Z'\Sigma Z)^{-1}(A^{-1})'. \tag{A6}$$

It can be shown that

$$A^{-1} = \begin{bmatrix} 1 & -(\frac{\beta_2^2 x_{10}}{\alpha D_1} + \frac{\beta_1^2 x_{20}}{\alpha D_2}) & -\frac{L_1 + L_2 - 2\beta_0}{L_2 - L_1} \\ 0 & \frac{\beta_2^2}{\alpha D_1} & \frac{2\beta_1}{L_2 - L_1} \\ 0 & \frac{\beta_1^2}{\alpha D_2} & \frac{2\beta_2}{L_2 - L_1} \end{bmatrix} \tag{A7}$$

where $L_1 = \text{logit}(P_1)$, $L_2 = \text{logit}(P_2)$, and $\alpha = \beta_1^2 + \beta_2^2$ and that

$$(Z'\Sigma Z)^{-1} = \frac{1}{8}\begin{bmatrix} \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} & 0 & \frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2} \\ 0 & \frac{4}{\sigma_1^2 + \sigma_2^2} & 0 \\ \frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2} & 0 & \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \end{bmatrix}. \tag{A8}$$

Substituting (A7) and (A8) into (A6) yields the elements of $I(\beta)^{-1}$ given by

$$V_{11} = \frac{1}{2(L_2 - L_1)^2}[\frac{(L_2 - \beta_0)^2}{\sigma_1^2} + \frac{(L_1 - \beta_0)^2}{\sigma_2^2}]$$
$$+ \frac{[x_{10} - \frac{\beta_1}{\alpha}(\frac{L_1 + L_2}{2} - \beta_0)]\,[x_{20} - \frac{\beta_2}{\alpha}(\frac{L_1 + L_2}{2} - \beta_0)]}{2(\sigma_1^2 + \sigma_2^2)D_1 D_2},$$

$$V_{22} = \frac{\beta_1^2}{2(L_2 - L_1)^2}(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}) + \frac{\beta_2^4}{2(\sigma_1^2 + \sigma_2^2)\alpha^2 D_1^2},$$

$$V_{33} = \frac{\beta_2^2}{2(L_2 - L_1)^2}(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}) + \frac{\beta_1^4}{2(\sigma_1^2 + \sigma_2^2)\alpha^2 D_2^2},$$

$$V_{12} = V_{21} = -\frac{\beta_1}{2(L_2 - L_1)^2}\left(\frac{L_2 - \beta_0}{\sigma_1^2} + \frac{L_1 - \beta_0}{\sigma_2^2}\right) - \frac{\beta_2^2[x_{10} - \frac{\beta_1}{\alpha}(\frac{L_1 + L_2}{2} - \beta_0)]}{2(\sigma_1^2 + \sigma_2^2)\alpha D_1^2},$$

$$V_{13} = V_{31} = -\frac{\beta_2}{2(L_2 - L_1)^2}\left(\frac{L_2 - \beta_0}{\sigma_1^2} + \frac{L_1 - \beta_0}{\sigma_2^2}\right) - \frac{\beta_1^2[x_{20} - \frac{\beta_2}{\alpha}(\frac{L_1 + L_2}{2} - \beta_0)]}{2(\sigma_1^2 + \sigma_2^2)\alpha D_2^2},$$

$$V_{23} = V_{32} = \frac{\beta_1\beta_2}{2(L_2 - L_1)^2}\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right) + \frac{\beta_1^2\beta_2^2}{2(\sigma_1^2 + \sigma_2^2)\alpha^2 D_1 D_2}.$$

(A9)

To obtain the region moment matrix S, which involves integration over the region R, define the transformation

$$\begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} \beta_1 & \beta_2 \\ -\beta_2 & \beta_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \quad \text{(A10)}$$

According to the theory of calculus, the integration of the function $f(x_1, x_2)$ over the region R in the (x_1, x_2) space is equivalent to the integration of the function $f(x_1(s, t), x_2(s, t))$ over the corresponding region R' in the (s, t) space, i.e.

$$\iint_{(R)} f(x_1, x_2)\, dx_1 dx_2 = \iint_{(R')} f(x_1(s, t), x_2(s, t))\, |J|^{-1}\, ds\, dt, \quad \text{(A11)}$$

where $|J|$ is the determinant of the Jacobian matrix given by

$$|J| = \begin{vmatrix} \frac{\partial s}{\partial x_1} & \frac{\partial s}{\partial x_2} \\ \frac{\partial t}{\partial x_1} & \frac{\partial t}{\partial x_2} \end{vmatrix} = \begin{vmatrix} \beta_1 & \beta_2 \\ -\beta_2 & \beta_1 \end{vmatrix} = \alpha \quad \text{(A12)}$$

and $x_1(s, t)$ and $x_2(s, t)$ are given by the inverse transformation

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{\alpha} \begin{bmatrix} \beta_1 & -\beta_2 \\ \beta_2 & \beta_1 \end{bmatrix} \begin{bmatrix} s \\ t \end{bmatrix}. \quad \text{(A13)}$$

Let

$$t_c = -\beta_2 x_{1c} + \beta_1 x_{2c} \quad \text{and} \quad t_\Delta = -\beta_2\Delta_1 + \beta_1\Delta_2\,, \tag{A14}$$

where the region center (x_{1c}, x_{2c}) and the region distances Δ_1 and Δ_2 are as shown in figure 8.1.1. In fact, $t_\Delta = -\frac{\Delta_1}{\beta_2}\alpha = \frac{\Delta_2}{\beta_1}\alpha$. Let

$$s_1 = L_1' - \beta_0 \quad \text{and} \quad s_2 = L_2' - \beta_0 \tag{A15}$$

where L_1' and L_2' are the logits on the region boundaries, i.e. $L_1' = \text{logit}(P_1')$ and $L_2' = \text{logit}(P_2')$. Let

$$t_1 = t_c - t_\Delta \quad \text{and} \quad t_2 = t_c + t_\Delta\,. \tag{A16}$$

The region R′ in the (s, t) space is given by $R' = \{(s, t)\colon s\in[s_1, s_2] \cap t\in[t_1, t_2]\}$. The equation in (A11) can be rewritten as

$$\iint_{(R)} f(x_1, x_2)\, dx_1dx_2 = \frac{1}{\alpha}\int_{t_1}^{t_2}\int_{s_1}^{s_2} f(x_1(s, t), x_2(s, t))\, ds\, dt\,. \tag{A17}$$

With (A17), the area K of the region R is found to be

$$K = \iint_{(R)} dx_1dx_2 = \frac{1}{\alpha}\int_{t_1}^{t_2}\int_{s_1}^{s_2} ds\, dt = -\frac{2\Delta_1}{\beta_2}(L_2' - L_1') = \frac{2\Delta_2}{\beta_1}(L_2' - L_1'). \tag{A18}$$

The elements of the region moment matrix S are given by the following integrations:

$$s_{11} = \frac{1}{K}\iint_{(R)} dx_1dx_2\,,$$

$$s_{12} = s_{21} = \frac{1}{K}\iint_{(R)} x_1dx_1dx_2 = \frac{1}{\alpha^2K}\int_{t_1}^{t_2}\int_{s_1}^{s_2} (\beta_1 s - \beta_2 t)\, ds\, dt\,,$$

$$s_{13} = s_{31} = \frac{1}{K}\iint_{(R)} x_2dx_1dx_2 = \frac{1}{\alpha^2K}\int_{t_1}^{t_2}\int_{s_1}^{s_2} (\beta_2 s + \beta_1 t)\, ds\, dt\,,$$

$$s_{22} = \frac{1}{K}\iint_{(R)} x_1^2dx_1dx_2 = \frac{1}{\alpha^3K}\int_{t_1}^{t_2}\int_{s_1}^{s_2} (\beta_1 s - \beta_2 t)^2\, ds\, dt\,,$$

$$s_{33} = \frac{1}{K}\iint_{(R)} x_2^2dx_1dx_2 = \frac{1}{\alpha^3K}\int_{t_1}^{t_2}\int_{s_1}^{s_2} (\beta_2 s + \beta_1 t)^2\, ds\, dt\,,$$

$$s_{23} = s_{32} = \frac{1}{K} \iint_{(R)} x_1 x_2 dx_1 dx_2 = \frac{1}{\alpha^3 K} \int_{t_1}^{t_2} \int_{s_1}^{s_2} (\beta_1 s - \beta_2 t)(\beta_2 s + \beta_1 t)\, ds\, dt . \tag{A19}$$

The integrations in (A19) yield the following results:

$$s_{11} = 1 ,$$

$$s_{12} = s_{21} = \frac{\beta_1}{2\alpha}(s_1 + s_2) - \frac{\beta_2}{\alpha} t_c ,$$

$$s_{13} = s_{31} = \frac{\beta_2}{2\alpha}(s_1 + s_2) + \frac{\beta_1}{\alpha} t_c ,$$

$$s_{22} = \frac{\beta_1^2}{3\alpha^2}(s_1^2 + s_1 s_2 + s_2^2) - \frac{\beta_1\beta_2}{\alpha^2} t_c(s_1 + s_2) + \frac{\beta_2^2}{\alpha^2} t_c^2 + \frac{1}{3}\Delta_1^2 ,$$

$$s_{33} = \frac{\beta_2^2}{3\alpha^2}(s_1^2 + s_1 s_2 + s_2^2) + \frac{\beta_1\beta_2}{\alpha^2} t_c(s_1 + s_2) + \frac{\beta_1^2}{\alpha^2} t_c^2 + \frac{1}{3}\Delta_2^2 ,$$

$$s_{23} = s_{32} = \frac{\beta_1\beta_2}{3\alpha^2}(s_1^2 + s_1 s_2 + s_2^2) + \frac{\beta_1^2 - \beta_2^2}{2\alpha^2} t_c(s_1 + s_2) - \frac{\beta_1\beta_2}{\alpha^2} t_c^2 + \frac{1}{3}\Delta_1\Delta_2 . \tag{A20}$$

Let $I(\beta)^{-1}S = (m_{ij})_{3\times3}$. Multiplying the elements of $I(\beta)^{-1}$ given in (A9) by the elements of S leads to the diagonal elements of $I(\beta)^{-1}S$ expressed in the form

$$m_{11} = \frac{1}{2(L_2 - L_1)^2}\left[\frac{(L_2 - \beta_0)^2}{\sigma_1^2} + \frac{(L_1 - \beta_0)^2}{\sigma_2^2}\right] - \frac{s_1 + s_2}{4(L_2 - L_1)^2}\left(\frac{L_2 - \beta_0}{\sigma_1^2} + \frac{L_1 - \beta_0}{\sigma_2^2}\right)$$
$$+ \frac{1}{2(\sigma_1^2 + \sigma_2^2)D_1^2}\left[(x_{10} - \frac{\beta_1}{\alpha}(\frac{L_1 + L_2}{2} - \beta_0))^2 + \frac{\beta_2}{\alpha} t_c(x_{10} - \frac{\beta_1}{\alpha}(\frac{L_1 + L_2}{2} - \beta_0))\right] ,$$

$$m_{22} + m_{33} = - \frac{s_1 + s_2}{4(L_2 - L_1)^2}\left(\frac{L_2 - \beta_0}{\sigma_1^2} + \frac{L_1 - \beta_0}{\sigma_2^2}\right) + \frac{s_1^2 + s_1 s_2 + s_2^2}{6(L_2 - L_1)^2}\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)$$
$$+ \frac{1}{2(\sigma_1^2 + \sigma_2^2)D_1^2}\left[\frac{\beta_2}{\alpha} t_c(x_{10} - \frac{\beta_1}{\alpha}(\frac{L_1 + L_2}{2} - \beta_0)) + \frac{\beta_2^2}{\alpha^2} t_c^2 + \frac{1}{3}\Delta_1^2\right] . \tag{A21}$$

The trace of $I(\beta)^{-1}S$ is given by

$$\text{Trace}[I(\beta)^{-1}S] = m_{11} + m_{22} + m_{33}$$

$$= \frac{1}{2(L_2 - L_1)^2}[\frac{(L_2 - \beta_0)^2}{\sigma_1^2} + \frac{(L_1 - \beta_0)^2}{\sigma_2^2}]$$

$$- \frac{s_1 + s_2}{2(L_2 - L_1)^2}(\frac{L_2 - \beta_0}{\sigma_1^2} + \frac{L_1 - \beta_0}{\sigma_2^2}) + \frac{s_1^2 + s_1 s_2 + s_2^2}{6(L_2 - L_1)^2}(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2})$$

$$+ \frac{1}{2(\sigma_1^2 + \sigma_2^2)D_1^2}[((x_{10} - \frac{\beta_1}{\alpha}(\frac{L_1 + L_2}{2} - \beta_0)) + \frac{\beta_2}{\alpha}t_c)^2 + \frac{1}{3}\Delta_1^2].$$

(A22)

Substituting s_1,s_2, and t_c with their expressions in (A14) and (A15) leads to $\text{Trace}[I(\beta)^{-1}S]$ expressed solely in the design and region parameters as

$$\text{Trace}[I(\beta)^{-1}S] = \frac{1}{2(L_2 - L_1)^2}[\frac{(L_2 - \frac{L_1' + L_2'}{2})^2}{\sigma_1^2} + \frac{(L_1 - \frac{L_1' + L_2'}{2})^2}{\sigma_2^2}]$$

$$+ \frac{(L_2' - L_1')^2}{24(L_2 - L_1)^2}[\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}]$$

$$+ \frac{1}{2(\sigma_1^2 + \sigma_2^2)D_1^2}[(x_{10} - x_{1c} - \frac{\beta_1}{\alpha}(\frac{L_1 + L_2}{2} - \frac{L_1' + L_2'}{2}))^2 + \frac{1}{3}\Delta_1^2].$$

(A23)

Substituting (A23) into (A4) leads to the Q criterion expressed as

$$\text{APV}[\text{logit}(\hat{p})] = \frac{1}{(L_2 - L_1)^2}[\frac{(L_2 - \frac{L_1' + L_2'}{2})^2}{rP_1(1 - P_1)} + \frac{(L_1 - \frac{L_1' + L_2'}{2})^2}{(1 - r)P_2(1 - P_2)}]$$

$$+ \frac{(L_2' - L_1')^2}{12(L_2 - L_1)^2}[\frac{1}{rP_1(1 - P_1)} + \frac{1}{(1 - r)P_2(1 - P_2)}]$$

$$+ \frac{\frac{1}{D_1^2}[(x_{10} - x_{1c} - \frac{\beta_1}{\alpha}(\frac{L_1 + L_2}{2} - \frac{L_1' + L_2'}{2}))]^2 + \frac{\Delta_1^2}{3D_1^2}}{rP_1(1 - P_1) + (1 - r)P_2(1 - P_2)}$$

(A24)

where r is the sample proportion allocated to ED_{100P_1}, or $r = 2n_1/N$.

Appendix B Q Optimality Criterion Based on Probability for Parallel-Line Design

Using the Delta method, the Q criterion based on the response probability can be written as

$$\mathrm{APV}[\hat{p}] = \frac{N}{K}\int_R p^2(1-p)^2\, x'(X'\Sigma X)^{-1}x\, dx = N\cdot \mathrm{Trace}[I(\beta)^{-1}S^*] \tag{B1}$$

where p is the response probability at the point x and S* is a matrix containing integrations over the region R in the form

$$S^* = (s^*_{ij})_{3\times 3} = \frac{1}{K}\int_R p^2(1-p)^2\, x\, x'\, dx$$

$$= \frac{1}{K}\iint_{(R)} \begin{bmatrix} p^2(1-p)^2 & p^2(1-p)^2 x_1 & p^2(1-p)^2 x_2 \\ p^2(1-p)^2 x_1 & p^2(1-p)^2 x_1^2 & p^2(1-p)^2 x_1 x_2 \\ p^2(1-p)^2 x_2 & p^2(1-p)^2 x_1 x_2 & p^2(1-p)^2 x_2^2 \end{bmatrix} dx_1 dx_2 .$$

(B2)

Applying the transformation in (A10) leads to

$$p = \frac{1}{1+e^{-(\beta_0+s)}} . \tag{B3}$$

Using the relationship in (A17), the elements of S* can be written as integrations in the (s, t) space:

$$s^*_{11} = \frac{1}{\alpha K}\int_{t_1}^{t_2}\int_{s_1}^{s_2} p^2(1-p)^2\, ds\, dt ,$$

$$s^*_{12} = s^*_{21} = \frac{1}{\alpha^2 K}\int_{t_1}^{t_2}\int_{s_1}^{s_2} p^2(1-p)^2\, (\beta_1 s - \beta_2 t)\, ds\, dt ,$$

$$s^*_{13} = s^*_{31} = \frac{1}{\alpha^2 K}\int_{t_1}^{t_2}\int_{s_1}^{s_2} p^2(1-p)^2\, (\beta_2 s + \beta_1 t)\, ds\, dt ,$$

$$s^*_{22} = \frac{1}{\alpha^3 K}\int_{t_1}^{t_2}\int_{s_1}^{s_2} p^2(1-p)^2\, (\beta_1 s - \beta_2 t)^2\, ds\, dt ,$$

$$s_{33}^* = \frac{1}{\alpha^3 K} \int_{t_1}^{t_2} \int_{s_1}^{s_2} p^2(1-p)^2 (\beta_2 s + \beta_1 t)^2 \, ds \, dt ,$$

$$s_{23}^* = s_{32}^* = \frac{1}{\alpha^3 K} \int_{t_1}^{t_2} \int_{s_1}^{s_2} p^2(1-p)^2 (\beta_1 s - \beta_2 t)(\beta_2 s + \beta_1 t) \, ds \, dt , \tag{B4}$$

where s_1, s_2, t_1, t_2, α, and K are as defined in appendix A. In finding the elements of S*, the following indefinite integrations are frequently used:

$$\int p^2(1-p)^2 \, ds = \frac{1}{6} p^2(3-2p) ,$$

$$\int p^2(1-p)^2 s \, ds = \frac{1}{6} [p^2(3-2p) s + p(1-p) + \log(1-p)] ,$$

and $$\int p^2(1-p)^2 s^2 \, ds = \frac{1}{6} [p^2(3-2p) s^2 + 2p(1-p) s - 2p - 2\beta_0 \log(1-p) - 2\int \frac{L}{1+e^{-L}} dL] . \tag{B5}$$

Define $$h = P_2'^2 (3 - 2P_2') - P_1'^2 (3 - 2P_1') ,$$

$$h_1 = s_2 P_2'^2 (3 - 2P_2') - s_1 P_1'^2 (3 - 2P_1') ,$$

and $$h_2 = s_2^2 P_2'^2 (3 - 2P_2') - s_1^2 P_1'^2 (3 - 2P_1') . \tag{B6}$$

Making use of (B5), it is shown that

$$s_{11}^* = \frac{h}{6(L_2' - L_1')} ,$$

$$s_{12}^* = s_{21}^* = \frac{\beta_1}{6\alpha(L_2' - L_1')} [h_1 + P_2'(1-P_2') - P_1'(1-P_1') + \log(\frac{1-P_2'}{1-P_1'}) - \frac{\beta_2}{\beta_1} t_c h] ,$$

$$s_{13}^* = s_{31}^* = \frac{\beta_2}{6\alpha(L_2' - L_1')} [h_1 + P_2'(1-P_2') - P_1'(1-P_1') + \log(\frac{1-P_2'}{1-P_1'}) + \frac{\beta_1}{\beta_2} t_c h] ,$$

$$\begin{aligned} s_{22}^* = \frac{\beta_1^2}{6\alpha^2 (L_2' - L_1')} & [h(\frac{\beta_2^2}{\beta_1^2})(t_c^2 - \frac{\alpha^2 \Delta_1 \Delta_2}{3\beta_1\beta_2}) \\ & - 2 t_c(\frac{\beta_2}{\beta_1})(h_1 + P_2'(1-P_2') - P_1'(1-P_1') + \log(\frac{1-P_2'}{1-P_1'})) + h_2 \\ & + 2s_2 P_2'(1-P_2') - 2s_1 P_1'(1-P_1') - 2P_2' + 2P_1' - 2\beta_0 \log(\frac{1-P_2'}{1-P_1'}) - 2\int_{L_1'}^{L_2'} \frac{L}{1+e^{-L}} dL] , \end{aligned}$$

$$s_{33}^* = \frac{\beta_2^2}{6\alpha^2(L_2' - L_1')}[\, h(\frac{\beta_1^2}{\beta_2^2})(t_c^2 - \frac{\alpha^2\Delta_1\Delta_2}{3\beta_1\beta_2})$$

$$+ 2\, t_c(\frac{\beta_1}{\beta_2})(h_1 + P_2'(1-P_2') - P_1'(1-P_1') + \log(\frac{1-P_2'}{1-P_1'})) + h_2$$

$$+ 2s_2 P_2'(1-P_2') - 2s_1 P_1'(1-P_1') - 2P_2' + 2P_1' - 2\beta_0 \log(\frac{1-P_2'}{1-P_1'}) - 2\int_{L_1'}^{L_2'} \frac{L}{1+e^{-L}} dL\,] ,$$

$$s_{23}^* = s_{32}^* = \frac{\beta_1\beta_2}{6\alpha^2(L_2' - L_1')}[\, - h(t_c^2 - \frac{\alpha^2\Delta_1\Delta_2}{3\beta_1\beta_2})$$

$$+ 2\, t_c(\frac{\beta_1}{\beta_2} - \frac{\beta_2}{\beta_1})(h_1 + P_2'(1-P_2') - P_1'(1-P_1') + \log(\frac{1-P_2'}{1-P_1'})) + h_2$$

$$+ 2s_2 P_2'(1-P_2') - 2s_1 P_1'(1-P_1') - 2P_2' + 2P_1' - 2\beta_0 \log(\frac{1-P_2'}{1-P_1'}) - 2\int_{L_1'}^{L_2'} \frac{L}{1+e^{-L}} dL\,] ,$$

(B7)

where $L_1' = \text{logit}(P_1')$, $L_2' = \text{logit}(P_2')$, t_c is defined in (A14), and Δ_1 and Δ_2 are region distances as shown in figure 8.1.1.

Let $I(\beta)^{-1}S^* = (m_{ij})_{3\times3}$. Multiplying the elements of $I(\beta)^{-1}$ given in (A9) by the elements of S^* yields the diagonal elements of $I(\beta)^{-1}S^*$ expressed as

$$m_{11} = \frac{h}{12(L_2' - L_1')(L_2 - L_1)^2}[\frac{(L_2-\beta_0)^2}{\sigma_1^2} + \frac{(L_1-\beta_0)^2}{\sigma_2^2}]$$

$$- \frac{\frac{L_2-\beta_0}{\sigma_1^2} + \frac{L_1-\beta_0}{\sigma_2^2}}{12(L_2' - L_1')(L_2 - L_1)^2}[h_1 + P_2'(1-P_2') - P_1'(1-P_1') + \log(\frac{1-P_2'}{1-P_1'})]$$

$$+ \frac{h}{12(L_2' - L_1')(\sigma_1^2 + \sigma_2^2)D_1^2}$$

$$\cdot [(x_{10} - \frac{\beta_1}{\alpha}(\frac{L_1+L_2}{2} - \beta_0))^2 + \frac{\beta_2}{\alpha} t_c(x_{10} - \frac{\beta_1}{\alpha}(\frac{L_1+L_2}{2} - \beta_0))] ,$$

$$m_{22} = -\frac{\beta_1^2(\frac{L_2-\beta_0}{\sigma_1^2}+\frac{L_1-\beta_0}{\sigma_2^2})}{12\alpha(L_2'-L_1')(L_2-L_1)^2}$$

$$\cdot [h_1 + P_2'(1-P_2') - P_1'(1-P_1') + \log(\frac{1-P_2'}{1-P_1'}) - \frac{\beta_2}{\beta_1}t_c h]$$

$$+ \frac{\beta_1^2(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2})}{12\alpha(L_2'-L_1')(L_2-L_1)^2}[h_2 + 2s_2 P_2'(1-P_2') - 2s_1 P_1'(1-P_1')$$

$$- 2P_2' + 2P_1' - 2\beta_0\log(\frac{1-P_2'}{1-P_1'}) - 2\int_{L_1'}^{L_2'}\frac{L}{1+e^{-L}}dL$$

$$- 2t_c(\frac{\beta_2}{\beta_1})(h_1 + P_2'(1-P_2') - P_1'(1-P_1') + \log(\frac{1-P_2'}{1-P_1'}))]$$

$$+ \frac{\beta_2^2}{12\alpha(L_2'-L_1')(\sigma_1^2+\sigma_2^2)D_1^2}[h(\frac{\beta_2^2}{\alpha^2}t_c^2 + \frac{1}{3}\Delta_1^2)$$

$$+ \frac{\beta_2}{\alpha}t_c h(x_{10} - \frac{\beta_1}{\alpha}(\frac{L_1+L_2}{2} - \beta_0))$$

$$- \frac{\beta_1}{\alpha}((x_{10} - \frac{\beta_1}{\alpha}(\frac{L_1+L_2}{2} - \beta_0)) + \frac{\beta_2}{\alpha}t_c)$$

$$\cdot (h_1 + P_2'(1-P_2') - P_1'(1-P_1') + \log(\frac{1-P_2'}{1-P_1'}))],$$

$$m_{33} = -\frac{\beta_2^2(\frac{L_2-\beta_0}{\sigma_1^2}+\frac{L_1-\beta_0}{\sigma_2^2})}{12\alpha(L_2'-L_1')(L_2-L_1)^2}$$

$$\cdot [h_1 + P_2'(1-P_2') - P_1'(1-P_1') + \log(\frac{1-P_2'}{1-P_1'}) + \frac{\beta_1}{\beta_2}t_c h]$$

$$+ \frac{\beta_2^2(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2})}{12\alpha(L_2'-L_1')(L_2-L_1)^2}[h_2 + 2s_2 P_2'(1-P_2') - 2s_1 P_1'(1-P_1')$$

$$- 2P_2' + 2P_1' - 2\beta_0 \log(\frac{1-P_2'}{1-P_1'}) - 2\int_{L_1'}^{L_2'} \frac{L}{1+e^{-L}} dL$$

$$+ 2t_c(\frac{\beta_1}{\beta_2})(h_1 + P_2'(1-P_2') - P_1'(1-P_1') + \log(\frac{1-P_2'}{1-P_1'}))]$$

$$+ \frac{\beta_1^2}{12\alpha(L_2' - L_1')(\sigma_1^2 + \sigma_2^2)D_2^2}[h(\frac{\beta_1^2}{C_\beta^2}t_c^2 + \frac{1}{3}\Delta_2^2)$$

$$- \frac{\beta_1}{\alpha} t_c h(x_{20} - \frac{\beta_2}{\alpha}(\frac{L_1+L_2}{2} - \beta_0))$$

$$- \frac{\beta_2}{\alpha}((x_{20} - \frac{\beta_2}{\alpha}(\frac{L_1+L_2}{2} - \beta_0)) - \frac{\beta_1}{\alpha} t_c)$$

$$\cdot (h_1 + P_2'(1-P_2') - P_1'(1-P_1') + \log(\frac{1-P_2'}{1-P_1'}))].$$

(B8)

It follows from (B8) that the trace of $I(\beta)^{-1}S^*$ is

$$\text{Trace}[I(\beta)^{-1}S^*] = m_{11} + m_{22} + m_{33}$$

$$= \frac{h}{12(L_2' - L_1')(L_2 - L_1)^2}[\frac{(L_2-\beta_0)^2}{\sigma_1^2} + \frac{(L_1-\beta_0)^2}{\sigma_2^2}]$$

$$- \frac{\frac{L_2-\beta_0}{\sigma_1^2} + \frac{L_1-\beta_0}{\sigma_2^2}}{6(L_2' - L_1')(L_2 - L_1)^2}[h_1 + P_2'(1-P_2') - P_1'(1-P_1') + \log(\frac{1-P_2'}{1-P_1'})]$$

$$+ \frac{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}{12(L_2' - L_1')(L_2 - L_1)^2}[h_2 + 2s_2 P_2'(1-P_2') - 2s_1 P_1'(1-P_1') - 2P_2' + 2P_1'$$

$$- 2\beta_0 \log(\frac{1-P_2'}{1-P_1'}) - 2\int_{L_1'}^{L_2'} \frac{L}{1+e^{-L}} dL]$$

$$+ \frac{h}{12(L_2' - L_1')(\sigma_1^2 + \sigma_2^2)D_1^2}[((x_{10} - \frac{\beta_1}{\alpha}(\frac{L_1+L_2}{2} - \beta_0)) + \frac{\beta_2}{\alpha} t_c)^2 + \frac{1}{3}\Delta_1^2].$$

(B9)

Simplifying (B9) leads to

$$\text{Trace}[I(\beta)^{-1}S^*] = \frac{1}{6(L_2' - L_1')(L_2 - L_1)^2}$$

$$\cdot\, [\frac{a_2(L_2' - L_2)^2 - a_1(L_1' - L_2)^2}{2\sigma_1^2} + \frac{a_2(L_2' - L_1)^2 - a_1(L_1' - L_1)^2}{2\sigma_2^2}$$

$$+ \frac{b_2(L_2' - L_2) - b_1(L_1' - L_2)}{\sigma_1^2} + \frac{b_2(L_2' - L_1) - b_1(L_1' - L_1)}{\sigma_2^2}$$

$$- (\frac{L_2}{\sigma_1^2} + \frac{L_1}{\sigma_2^2})\log(\frac{1 - P_2'}{1 - P_1'}) - (\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2})(P_2' - P_1' + \int_{L_1'}^{L_2'} \frac{L}{1 + e^{-L}} dL)\,]$$

$$+ \frac{a_2 - a_1}{12(L_2' - L_1')(\sigma_1^2 + \sigma_2^2)D_1^2} [(x_{10} - x_{1c} - \frac{\beta_1}{\alpha}(\frac{L_1 + L_2}{2} - \frac{L_1' + L_2'}{2}))^2 + \frac{1}{3}\Delta_1^2] \qquad \text{(B10)}$$

where $\quad a_1 = P_1'^2 (3\text{-}2P_1'), \qquad a_2 = P_2'^2 (3\text{-}2P_2'),$

$\qquad b_1 = P_1'(1\text{-}P_1'), \qquad$ and $\qquad b_2 = P_2'(1\text{-}P_2').$ (B11)

The Q criterion is then given by (B10) multiplied by N. Minimizing this criterion requires that the best location for the design center satisfy

$$x_{10} - x_{1c} - \frac{\beta_1}{C_\beta}(\frac{L_1 + L_2}{2} - \frac{L_1' + L_2'}{2}) = 0, \qquad \text{(B12)}$$

which corresponds to the condition in (8.1.2). This condition guarantees the presence of the same feature as illustrated in figure 8.1.2 regarding the design center. When the above condition holds, the Q criterion is reduced to

$$\text{APV}[\hat{p}] = Q(P_1, P_2, r \mid P_1', P_2', g)$$

$$= \frac{1}{3(L_2' - L_1')(L_2 - L_1)^2}$$

$$\cdot\, [\frac{P_2'^2(3 - 2P_2')(L_2' - L_2)^2 - P_1'^2(3 - 2P_1')(L_1' - L_2)^2}{2rP_1(1 - P_1)}$$

$$+\frac{P_2'^2(3-2P_2')(L_2'-L_1)^2-P_1'^2(3-2P_1')(L_1'-L_1)^2}{2(1-r)P_2(1-P_2)}$$

$$+\frac{P_2'(1-P_2')(L_2'-L_2)-P_1'(1-P_1')(L_1'-L_2)}{rP_1(1-P_1)}$$

$$+\frac{P_2'(1-P_2')(L_2'-L_1)-P_1'(1-P_1')(L_1'-L_1)}{(1-r)P_2(1-P_2)}$$

$$-\left(\frac{L_2}{rP_1(1-P_1)}+\frac{L_1}{(1-r)P_2(1-P_2)}\right)\log\left(\frac{1-P_2'}{1-P_1'}\right)$$

$$-\left(\frac{1}{rP_1(1-P_1)}+\frac{1}{(1-r)P_2(1-P_2)}\right)\left(P_2' - P_1' + \int_{L_1'}^{L_2'}\frac{L}{1+e^{-L}}dL\right)\Big]$$

$$+\frac{P_2'^2(3-2P_2')-P_1'^2(3-2P_1')}{18g^2(L_2'-L_1')[rP_1(1-P_1)+(1-r)P_2(1-P_2)]}.$$

(B13)

When the region is symmetric, the Q criterion can be further simplified to

$APV[\hat{p}] = Q_1(P_1, P_2, r \mid P', g)$

$$=\frac{1}{12L'(L_2-L_1)^2}\Big[(2P'-1)(1+2P'-2P'^2)\left(\frac{L_2^2}{rP_1(1-P_1)}+\frac{L_1^2}{(1-r)P_2(1-P_2)}\right)$$

$$+\left(\frac{1}{rP_1(1-P_1)}+\frac{1}{(1-r)P_2(1-P_2)}\right)$$

$$\cdot\left(L'^2(2P'-1)(1+2P'-2P'^2)+4L'P'(1-P')-2(2P'-1)-2\int_{L_1'}^{L_2'}\frac{L}{1+e^{-L}}dL\right)\Big]$$

$$+\frac{(2P'-1)(1+2P'-2P'^2)}{36g^2L'[rP_1(1-P_1)+(1-r)P_2(1-P_2)]}$$

(B14)

where P' is the upper probability of the region and $L' = \text{logit}(P')$. Using (B14) to find the Q optimal design for a symmetric region can greatly reduce computational intensity.

Appendix C Q Optimality Criterion Based on Logit for Hyperbola-Based Design

Similar to (A4), the Q criterion for the interaction model can be written as

$$\mathrm{APV}[\mathrm{logit}(\hat{p})] = N \cdot \mathrm{Trace}[I(\beta)^{-1} S] \tag{C1}$$

where $I(\beta) = X'\Sigma X$ and S is the region moment matrix given by

$$S = (s_{ij})_{4\times 4} = \frac{1}{K}\int_R x\, x' dx = \frac{1}{K}\iint_{(R)} \begin{bmatrix} 1 & x_1 & x_2 & x_1x_2 \\ x_1 & x_1^2 & x_1x_2 & x_1^2x_2 \\ x_2 & x_1x_2 & x_2^2 & x_1x_2^2 \\ x_1x_2 & x_1^2x_2 & x_1x_2^2 & x_1^2x_2^2 \end{bmatrix} dx_1 dx_2. \tag{C2}$$

where K is the area of the region R.

The model matrix X contained in $I(\beta)$ can be written as

$$X = ZC \tag{C3}$$

where

$$Z = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \tag{C4}$$

and

$$C = \begin{bmatrix} 1 & x_{10} & x_{20} & x_{10}x_{20} + \frac{z_{12}z_{22} - z_{11}z_{21}}{2} \\ 0 & \frac{z_{11}+z_{12}}{2} & \frac{z_{22}-z_{21}}{2} & \frac{z_{22}-z_{21}}{2}x_{10} + \frac{z_{11}+z_{12}}{2}x_{20} \\ 0 & \frac{z_{12}-z_{11}}{2} & \frac{z_{21}+z_{22}}{2} & \frac{z_{21}+z_{22}}{2}x_{10} + \frac{z_{12}-z_{11}}{2}x_{20} \\ 0 & 0 & 0 & \frac{z_{11}z_{21}+z_{12}z_{22}}{2} \end{bmatrix}, \tag{C5}$$

where (x_{10}, x_{20}) is the hyperbola center and z_{ij}, i, j = 1, 2, are the centered design levels as indicated in figure 8.4.1. It follows from (C3) that

$$I(\beta)^{-1} = (X'\Sigma X)^{-1} = C^{-1}(Z'\Sigma Z)^{-1}(C^{-1})' \tag{C6}$$

where

$$\Sigma = \begin{bmatrix} \sigma_2^2 & 0 & 0 & 0 \\ 0 & \sigma_1^2 & 0 & 0 \\ 0 & 0 & \sigma_1^2 & 0 \\ 0 & 0 & 0 & \sigma_2^2 \end{bmatrix} \text{ and } Z'\Sigma Z = 2\begin{bmatrix} \sigma_1^2+\sigma_2^2 & 0 & 0 & \sigma_2^2-\sigma_1^2 \\ 0 & \sigma_1^2+\sigma_2^2 & \sigma_2^2-\sigma_1^2 & 0 \\ 0 & \sigma_2^2-\sigma_1^2 & \sigma_1^2+\sigma_2^2 & 0 \\ \sigma_2^2-\sigma_1^2 & 0 & 0 & \sigma_1^2+\sigma_2^2 \end{bmatrix}, \quad \text{(C7)}$$

where $\sigma_1^2 = n_1P_1(1\text{-}P_1)$ and $\sigma_2^2 = n_2P_2(1\text{-}P_2)$. It is found that

$$C^{-1} = \begin{bmatrix} 1 & -\frac{x_{10}(z_{21}+z_{22})-x_{20}(z_{12}-z_{11})}{z_{11}z_{22}+z_{12}z_{21}} & \frac{x_{10}(z_{22}-z_{21})-x_{20}(z_{11}+z_{12})}{z_{11}z_{22}+z_{12}z_{21}} & \frac{2x_{10}x_{20}-(z_{12}z_{22}-z_{11}z_{21})}{z_{11}z_{21}+z_{12}z_{22}} \\ 0 & \frac{z_{21}+z_{22}}{z_{11}z_{22}+z_{12}z_{21}} & -\frac{z_{22}-z_{21}}{z_{11}z_{22}+z_{12}z_{21}} & -\frac{2x_{20}}{z_{11}z_{21}+z_{12}z_{22}} \\ 0 & -\frac{z_{12}-z_{11}}{2z_{11}z_{22}+z_{12}z_{21}} & \frac{z_{11}+z_{12}}{z_{11}z_{22}+z_{12}z_{21}} & -\frac{2x_{10}}{z_{11}z_{21}+z_{12}z_{22}} \\ 0 & 0 & 0 & \frac{2}{z_{11}z_{21}+z_{12}z_{22}} \end{bmatrix}$$

(C8)

and

$$(Z'\Sigma Z)^{-1} = \frac{1}{8\sigma_1^2\sigma_2^2}\begin{bmatrix} \sigma_1^2+\sigma_2^2 & 0 & 0 & \sigma_1^2-\sigma_2^2 \\ 0 & \sigma_1^2+\sigma_2^2 & \sigma_1^2-\sigma_2^2 & 0 \\ 0 & \sigma_1^2-\sigma_2^2 & \sigma_1^2+\sigma_2^2 & 0 \\ \sigma_1^2-\sigma_2^2 & 0 & 0 & \sigma_1^2+\sigma_2^2 \end{bmatrix}. \quad \text{(C9)}$$

Let $I(\beta)^{-1} = (V_{ij})_{4\times4}$. Substituting (C8) and (C9) into (C6) leads to $I(\beta)^{-1}$ expressed with its elements

$$V_{11} = \frac{1}{2(z_{11}z_{21}+z_{12}z_{22})^2}\left[\frac{(z_{12}z_{22}-x_{10}x_{20})^2}{\sigma_1^2} + \frac{(z_{11}z_{21}+x_{10}x_{20})^2}{\sigma_2^2}\right]$$

$$+ \frac{1}{2(z_{11}z_{22}+z_{12}z_{21})^2}\left[\frac{(x_{10}z_{22}-x_{20}z_{12})^2}{\sigma_1^2} + \frac{(x_{10}z_{21}+x_{20}z_{11})^2}{\sigma_2^2}\right],$$

$$V_{22} = \frac{x_{20}^2}{2(z_{11}z_{21}+z_{12}z_{22})^2}\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right) + \frac{1}{2(z_{11}z_{22}+z_{12}z_{21})^2}\left(\frac{z_{22}^2}{\sigma_1^2} + \frac{z_{21}^2}{\sigma_2^2}\right),$$

$$V_{33} = \frac{x_{10}^2}{2(z_{11}z_{21}+z_{12}z_{22})^2}\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}\right) + \frac{1}{2(z_{11}z_{22}+z_{12}z_{21})^2}\left(\frac{z_{12}^2}{\sigma_1^2}+\frac{z_{11}^2}{\sigma_2^2}\right),$$

$$V_{44} = \frac{1}{2(z_{11}z_{21}+z_{12}z_{22})^2}\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}\right),$$

$$V_{12} = V_{21} = \frac{x_{20}}{2(z_{11}z_{21}+z_{12}z_{22})^2}\left[\frac{z_{12}z_{22}-x_{10}x_{20}}{\sigma_1^2} - \frac{z_{11}z_{21}+x_{10}x_{20}}{\sigma_2^2}\right]$$
$$+ \frac{1}{2(z_{11}z_{22}+z_{12}z_{21})^2}\left[-\frac{z_{22}(x_{10}z_{22}-x_{20}z_{12})}{\sigma_1^2} - \frac{z_{21}(x_{10}z_{21}+x_{20}z_{11})}{\sigma_2^2}\right],$$

$$V_{13} = V_{31} = \frac{x_{10}}{2(z_{11}z_{21}+z_{12}z_{22})^2}\left[\frac{z_{12}z_{22}-x_{10}x_{20}}{\sigma_1^2} - \frac{z_{11}z_{21}+x_{10}x_{20}}{\sigma_2^2}\right]$$
$$+ \frac{1}{2(z_{11}z_{22}+z_{12}z_{21})^2}\left[\frac{z_{12}(x_{10}z_{22}-x_{20}z_{12})}{\sigma_1^2} - \frac{z_{11}(x_{10}z_{21}+x_{20}z_{11})}{\sigma_2^2}\right],$$

$$V_{14} = V_{41} = \frac{1}{2(z_{11}z_{21}+z_{12}z_{22})^2}\left[-\frac{z_{12}z_{22}-x_{10}x_{20}}{\sigma_1^2} + \frac{z_{11}z_{21}+x_{10}x_{20}}{\sigma_2^2}\right],$$

$$V_{23} = V_{32} = \frac{x_{10}x_{20}}{2(z_{11}z_{21}+z_{12}z_{22})^2}\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}\right)$$
$$+ \frac{1}{2(z_{11}z_{22}+z_{12}z_{21})^2}\left(-\frac{z_{12}z_{22}}{\sigma_1^2}+\frac{z_{11}z_{21}}{\sigma_2^2}\right),$$

$$V_{24} = V_{42} = -\frac{x_{20}}{2(z_{11}z_{21}+z_{12}z_{22})^2}\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}\right),$$

$$V_{34} = V_{43} = -\frac{x_{10}}{2(z_{11}z_{21}+z_{12}z_{22})^2}\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}\right). \tag{C10}$$

Let R′ denote the region in the centered (z_1, z_2) space corresponding to the region R in the (x_1, x_2) space. Define the transformation from the (z_1, z_2) to (u, v) space

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} z'_{21} + z'_{22} & z'_{11} - z'_{12} \\ z'_{21} - z'_{22} & z'_{11} + z'_{12} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \tag{C11}$$

where z'_{ij}, i, j = 1, 2, are the centered levels at the region vertices as displayed in figure 8.4.1. The Jacobian matrix has determinant

$$|J| = \begin{vmatrix} \frac{\partial u}{\partial z_1} & \frac{\partial u}{\partial z_2} \\ \frac{\partial v}{\partial z_1} & \frac{\partial v}{\partial z_2} \end{vmatrix} = \begin{vmatrix} z'_{21} + z'_{22} & z'_{11} - z'_{12} \\ z'_{21} - z'_{22} & z'_{11} + z'_{12} \end{vmatrix} = 2h \tag{C12}$$

where

$$h = z'_{11}z'_{22} + z'_{12}z'_{21}. \tag{C13}$$

The inverse transformation is given by

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \frac{1}{2h} \begin{bmatrix} z'_{11} + z'_{12} & z'_{12} - z'_{11} \\ z'_{22} - z'_{21} & z'_{21} + z'_{22} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}. \tag{C14}$$

The parallelogram region R′ in the (z_1, z_2) space is projected into a square region in the (u, v) space, which can be written as

$$R'' = \{(u, v): u \in [-h, h] \cap v \in [-h, h]\}. \tag{C15}$$

The property from calculus given in (A11) allows an integration over R′ in the (z_1, z_2) space to be expressed as an integration over R″ in the (u, v) space in the form

$$\iint_{(R')} f(z_1, z_2)\, dz_1 dz_2 = \frac{1}{2h} \int_{-h}^{h} \int_{-h}^{h} f(z_1(u, v), z_2(u, v))\, du\, dv. \tag{C16}$$

Using (C16), the area K of the region R is found to be

$$K = \iint_{(R)} dx_1 dx_2 = \iint_{(R')} dz_1 dz_2 = \frac{1}{2h} \int_{-h}^{h} \int_{-h}^{h} du\, dv = 2h\,. \tag{C17}$$

Also from (C16), it follows that any odd moment in z_1 and/or z_2 is zero, e.g.

$$\frac{1}{K} \iint_{(R')} z_1 dz_1 dz_2 = \frac{1}{K} \iint_{(R')} z_2 dz_1 dz_2 = \frac{1}{K} \iint_{(R')} z_1^2 z_2 dz_1 dz_2 = \frac{1}{K} \iint_{(R')} z_1 z_2^2 dz_1 dz_2 = 0. \tag{C18}$$

Using the centering relationship in (C3) and the property in (C18), the region moment matrix S defined in (C2) can be rewritten as

$$S = \frac{1}{K}\iint_{(R')}\begin{bmatrix} 1 & x_{10} & x_{20} & z_1z_2 + x_{10}x_{20} \\ & z_1^2 + x_{10}^2 & z_1z_2 + x_{10}x_{20} & x_{20}z_1^2 + 2x_{10}z_1z_2 + x_{10}^2x_{20} \\ & & z_2^2 + x_{20}^2 & x_{10}z_2^2 + 2x_{20}z_1z_2 + x_{10}x_{20}^2 \\ & & & z_1^2z_2^2 + 4x_{10}x_{20}z_1z_2 + x_{20}^2z_1^2 + x_{10}^2z_2^2 + x_{10}^2x_{20}^2 \end{bmatrix} dz_1dz_2$$

(C19)

where the lower diagonal elements are omitted due to the matrix being symmetric. Using (C16), the even moments involved in S are found to be

$$\frac{1}{K}\iint_{(R')} z_1^2dz_1dz_2 = \frac{1}{6}(z_{11}'^2 + z_{12}'^2),$$

$$\frac{1}{K}\iint_{(R')} z_2^2dz_1dz_2 = \frac{1}{6}(z_{21}'^2 + z_{22}'^2),$$

$$\frac{1}{K}\iint_{(R')} z_1z_2dz_1dz_2 = \frac{1}{6}(z_{12}'z_{22}' - z_{11}'z_{21}'),$$

$$\frac{1}{K}\iint_{(R')} z_1^2z_2^2dz_1dz_2 = \frac{1}{90}(6z_{11}'^2z_{21}'^2 + 6z_{12}'^2z_{22}'^2 + z_{11}'^2z_{22}'^2 + z_{12}'^2z_{21}'^2 - 4z_{11}'z_{12}'z_{21}'z_{22}').$$

(C20)

Substituting the region moments in (C20) into (C19) yields the elements of S

$$s_{11} = 1,$$

$$s_{22} = \frac{1}{6}(z_{11}'^2 + z_{12}'^2) + x_{10}^2,$$

$$s_{33} = \frac{1}{6}(z_{21}'^2 + z_{22}'^2) + x_{20}^2,$$

$$s_{44} = \frac{1}{90}(6z_{11}'^2z_{21}'^2 + 6z_{12}'^2z_{22}'^2 + z_{11}'^2z_{22}'^2 + z_{12}'^2z_{21}'^2 - 4z_{11}'z_{12}'z_{21}'z_{22}')$$
$$+ \frac{2}{3}x_{10}x_{20}(z_{12}'z_{22}' - z_{11}'z_{21}') + \frac{1}{6}x_{10}^2(z_{21}'^2 + z_{22}'^2) + \frac{1}{6}x_{20}^2(z_{11}'^2 + z_{12}'^2) + x_{10}^2x_{20}^2,$$

$$s_{12} = s_{21} = x_{10},$$

$$s_{13} = s_{31} = x_{20} ,$$

$$s_{14} = s_{41} = s_{23} = s_{32} = \frac{1}{6}(z'_{12}z'_{22} - z'_{11}z'_{21}) + x_{10}x_{20} ,$$

$$s_{24} = s_{42} = \frac{1}{6}x_{20}(z'^2_{11} + z'^2_{12}) + \frac{1}{3}x_{10}(z'_{12}z'_{22} - z'_{11}z'_{21}) + x^2_{10}x_{20} ,$$

$$s_{34} = s_{43} = \frac{1}{6}x_{10}(z'^2_{21} + z'^2_{22}) + \frac{1}{3}x_{20}(z'_{12}z'_{22} - z'_{11}z'_{21}) + x_{10}x^2_{20} . \tag{C21}$$

Multiplying $I(\beta)^{-1}$ given in (C10) and S gives the diagonal elements of $I(\beta)^{-1}S$, the sum of which leads to

$$\begin{aligned}\mathrm{Trace}[I(\beta)^{-1}S] = {} & \frac{1}{12(z_{11}z_{22}+z_{12}z_{21})^2}[\frac{1}{\sigma_1^2}((z'_{11}z_{22} + z'_{21}z_{12})^2 + (z'_{12}z_{22} - z'_{22}z_{12})^2) \\ & + \frac{1}{\sigma_2^2}((z'_{12}z_{21} + z'_{22}z_{11})^2 + (z'_{11}z_{21} - z'_{21}z_{11})^2)] \\ & + \frac{1}{4(z_{11}z_{21}+z_{12}z_{22})^2}[\frac{1}{\sigma_1^2}((z_{12}z_{22} + \frac{1}{3}z'_{11}z'_{21})^2 + (z_{12}z_{22} - \frac{1}{3}z'_{12}z'_{22})^2) \\ & + \frac{1}{\sigma_2^2}((z_{11}z_{21} + \frac{1}{3}z'_{12}z'_{22})^2 + (z_{11}z_{21} - \frac{1}{3}z'_{11}z'_{21})^2)] \\ & + \frac{1}{6(z_{11}z_{21}+z_{12}z_{22})^2}(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2})[(z'_{12}z'_{22}-z'_{11}z'_{21})^2+(z'_{11}z'_{22}- z'_{12}z'_{21})^2].\end{aligned} \tag{C22}$$

Based on §2.2, it can be seen that

$$z_{11}z_{21} = -\frac{L_1 - \beta_0^*}{\beta_{12}}, \qquad z_{12}z_{22} = \frac{L_2 - \beta_0^*}{\beta_{12}},$$

$$z'_{11}\, z'_{21} = -\frac{L'_1 - \beta_0^*}{\beta_{12}}, \qquad z'_{12}\, z'_{22} = \frac{L'_2 - \beta_0^*}{\beta_{12}}, \tag{C23}$$

where $\beta_0^* = \beta_0 - \frac{\beta_1\beta_2}{\beta_{12}}$, $L_1 = \text{logit}(P_1)$, $L_2 = \text{logit}(P_2)$, $L_1' = \text{logit}(P_1')$, and $L_2' = \text{logit}(P_2')$. Define two sets of ratios

$$(1) \qquad t_1 = \left|\frac{z_{11}}{z_{12}}\right|, \qquad t_1' = \left|\frac{z_{11}'}{z_{12}'}\right|, \qquad g_1 = \left|\frac{z_{11}}{z_{11}'}\right|, \tag{C24}$$

and

$$(2) \qquad t_2 = \left|\frac{z_{21}}{z_{22}}\right|, \qquad t_2' = \left|\frac{z_{21}'}{z_{22}'}\right|, \qquad g_2 = \left|\frac{z_{21}}{z_{21}'}\right|. \tag{C25}$$

The coordinates in Trace[$I(\beta)^{-1}S$] can be expressed in terms of the logits in (C23) and the ratios in (C24) or (C25). Substituting Trace[$I(\beta)^{-1}S$] into (C1) leads to the Q optimality criterion expresses as

$$\begin{aligned}
\text{APV}[\text{logit}(\hat{p})] = {} & \frac{N}{12[t(L_2-\beta_0^*)-\frac{1}{t}(L_1-\beta_0^*)]^2} \\
& \cdot[\frac{1}{\sigma_1^2}((\frac{t}{g}(L_2-\beta_0^*)-\frac{g}{t}(L_1'-\beta_0^*))^2+(\frac{t}{gt'}(L_2-\beta_0^*)-\frac{gt'}{t}(L_2'-\beta_0^*))^2) \\
& +\frac{1}{\sigma_2^2}((\frac{1}{g}(L_1-\beta_0^*)-g(L_1'-\beta_0^*))^2+(\frac{1}{gt'}(L_1-\beta_0^*)-gt'(L_2'-\beta_0^*))^2)] \\
& +\frac{N}{4(L_2-L_1)^2}[\frac{1}{\sigma_1^2}(((L_2-\beta_0^*)-\frac{1}{3}(L_1'-\beta_0^*))^2+((L_2-\beta_0^*)-\frac{1}{3}(L_2'-\beta_0^*))^2) \\
& +\frac{1}{\sigma_2^2}(((L_1-\beta_0^*)-\frac{1}{3}(L_1'-\beta_0^*))^2+((L_1-\beta_0^*)-\frac{1}{3}(L_2'-\beta_0^*))^2)] \\
& +\frac{N}{6(L_2-L_1)^2}(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2})[(L_1'+L_2'-2\beta_0^*)^2+(t'(L_2'-\beta_0^*)+\frac{1}{t'}(L_1'-\beta_0^*))^2]
\end{aligned} \tag{C26}$$

where the ratios t, t′, and g can be given by either set (1) in (C24), i.e. $t = t_1$, $t' = t_1'$, and $g = g_1$, or set (2) in (C25), i.e. $t = t_2$, $t' = t_2'$, and $g = g_2$.

VITA

Yan Jia, daughter of Guohua Yan and Shiqi Jia, was born on November 3, 1964 in Beijing, China. She enrolled in University of Science and Technology of China in 1982 and graduated in 1987 with a Bachelor of Science degree in System Science and Management Science. In 1989, she participated in the development of the Euro-Disney resort as an information analyst at Créativité et Développement, Société d'Exploitation de la Gaîte Lyrique in Paris, France. Her graduate studies at Virginia Polytechnic Institute and State University started in 1991. Yan was the recipient of the 1992 Boyd Harshbarger Award in recognition of superior scholarship during the first year of graduate studies in statistics. Later in 1992 she was inducted into Mu Sigma Rho, the National Statistical Honor Society, and received a Master of Science degree in Statistics. Yan continued her studies under the direction of Professor Raymond H. Myers and obtained her Ph.D. degree in Statistics in 1996. Upon graduation, she expected to join Bayer Corporation as a statistician. Yan Jia is a member of the American Statistical Association.

Yan Jia