D-Optimal Designs For Weighted Linear Regression and Binary Regression Models

Neclâ GÜNDÜZ

A Dissertation Submitted to
The University of Glasgow
For the Degree of

Doctor of Philosophy

UNIVERSITY
of
GLASGOW

DEPARTMENT OF STATISTICS

September 1999

Summary

This thesis is concerned with D-optimal designs primarly for binary response or weighted linear regression models. Its principle aim is to prove (using geometric and other arguments) that D-optimal designs have two support points for two parameter models depending on one design variable for all possible design intervals. We also extend established results, for Gamma, Beta and Normal density weight functions.

The first aim of this work is to prove Ford, Torsney and Wu (1992) conjectures for a variety of such models. We also extend these results to higher dimensions. This is based on a parameter-dependent transformation to a weighted regression model and results will be extended to other such models.

Chapter 1 mainly gives an introduction to the study for linear and nonlinear Optimal designs for regression models.

Chapter 2 leads on with D-optimal designs for binary regression models which depend on two parameters and one covariate x in a design region, say $\mathcal{X}$. It mainly deals with the following three cases: (a) $\mathcal{X}$ is a unbounded, (b) $\mathcal{X}$ is a bounded interval and (c) $\mathcal{X}$ is bounded at one end only. We first establish that only two support points are needed and then establish their values. The above conjecture is confirmed for most models using a transformation to weighted regression design.

Chapter 3 presents Weighted Linear Regression and D-optimal Designs for the particular case of a Three Parameter Model with two design variables under a transformation to a weighted linear regression when the design space is rectangular. We first show that we have a four-point design for many of the weight functions considered. We also have an explicit solution for the optimal weights. An appropriate extension of the above conjecture is confirmed.

Consideration of more realistic constraints on two design variables in **Chapter 4** leads, under a transformation, to bounded design spaces in the shape of polygons. We establish results about D-optimal designs for such spaces.

Chapter 5 widens the scope of the thesis, by considering more general models and, in particular, multiparameter binary regression models. Here again, we establish the existence of an explicit solution for the optimal weights for the rectangular case of the design space and further extensions of the conjecture.

Chapter 6 extends the ideas of **Chapter 2** by applying them to Contingent Valuation Studies. We illustrate one type of Contingent Valuation (CV) study, namely a dichotomous choice CV study with the design variable being a 'Bid' value. Respondents are asked if they are willing to pay this value for some service or amenity. We focus on both dichotomous choice (or single bounded) CV's and on double bounded CV's (in which a second bid is offered).

Finally, **Chapter 7** presents our conclusions and ideas for future work.

To My Mother.

Acknowledgements

I am deeply indebted to my supervisor Dr. Ben Torsney for his abundant patience, his support and his guidance. His willingness to share his vast experience played a central role throughout my research, and ultimately led to the successful completion of this Ph.D. thesis.

I also would like to thank Professor Adrian Bowman and all the staff members of the Statistics Department for providing me with every possible assistance, statistically or otherwise, and for maintaining a very friendly and intellectually conducive and stimulating research environment. I will live to remember their individual and collective positive influence on my research career.

Throughout my stay in Glasgow, I met so many good friends, and most particularly the colourful postgraduate students of the Statistics Department. They all inspired me positively in one way or another, and I herein wish to express my most sincere thanks to all of them for their unforgettable friendship.

My gratitute also goes to my friends in Ankara.

The U.K. Government ORS awards completed my Gazi University scholarship and made it possible for me to support myself financially throughoout my studies. Let this be an opportunity for me to express my deepest gratitude to them.

Finally, but most importantly, I would like to express my deeply heartfelt thanks to my family for their unlimited help, their unconditional love, and their constant encouragement and support at all times.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The last decade has seen an increasing level of work on optimal designs, with a particular focus on binary response models.

Throughout our research, the main focus has been on the exploration of some current ideas in this rapidly expanding field, with the aim of understanding and testing them in some specific contexts, and possibly extending them.

But before presenting our findings in detail, it is probably necessary to gain more insight into the building blocks of optimum design through a simple regression design example.

In a classical regression problem for example, the aim is to investigate the relation between a response variable and a set of explanatory variables. For such an investigation to be carried out, it is necessary to gather some values for the variables by making observations.

In some cases, it is possible for the experimenter to choose the values of the explanatory variables, which means that he/she can choose the situations in which the observations will be made. And here in lies a very significant aspect of the experiment, since the choice of the situation will in some way determine the quality of the design.

Some basic questions then naturally arise : **What is a design ? And how do we measure the quality of a design ?**
Although we shall later on cover both linear and nonlinear models in our study of optimum designs, we address the above questions through a simple linear model with n explanatory variables (van Berkum, 1995).

1.1 What is a design?

We consider a linear model with n explanatory variables $x_1, ..., x_n$. The notation for the model is as follows

$$Y = \beta_1 f_1(\underline{x}) + \beta_2 f_2(\underline{x}) + ... + \beta_k f_k(\underline{x}) + \varepsilon \; = \; \underline{\beta}^T \underline{f}(\underline{x}) + \varepsilon \; ,$$

with

$\underline{x} = (x_1, ..., x_n)^T$, the vector of n explanatory variables,

$\underline{x} \in \mathcal{X}$, the experimental region, $\quad \mathcal{X} \subset \mathcal{R}^n$,

$f_i : \mathcal{X} \to \mathcal{R}$, a continuous function from $\mathcal{X}$ into $\mathcal{R}$ $\;(i = 1, ..., k)$,

$\underline{f}(\underline{x}) = (f_1(\underline{x}), \cdots, f_k(\underline{x}))^T$,

and

ε a stochastic variable, the error term, independent of x

Y response, $Y \in \mathcal{R}$,

$\underline{\beta} = (\beta_1, ..., \beta_k)^T$, the vector of unknown parameters.

Let's assume in the above example that we are able to make N observations, and let's denote each observation by Y_i. For our assumed N observations, we shall therefore have a corresponding set of normally distributed random errors, each denoted by ε_i. In this particular case, we consider those errors ε_i to be independent and uncorrelated with zero mean and constant variance σ^2, that is $\mathbb{V}(\varepsilon_i) = \sigma^2 \quad \forall i = 1, \cdots, N$.

We now move on to the description of the design itself. We denote by $\underline{x}_1, \underline{x}_2, \cdots, \underline{x}_m$ the distinct m points ($m \leq n$) in the experimental region where observations will be taken. Here, each $\underline{x}_i = (x_{i1}, x_{i2}, \cdots, x_{in})^T \in \mathcal{X}$. We also define n_i as the number of observations made at the point $\underline{x}_i$, $\sum_{i=1}^{m} n_i = N$.

Let $\mathcal{E}$ denote the design of our experiment. For clarity, we shall sometimes refer to it as $\mathcal{E}(N)$ in order to express the fact that we make N observations in that particular design. In other words, a design $\mathcal{E}(N)$ can be fully described by specifying the above mentioned variables. We summarize the design as follows:

Design(General) : To obtain an observation on Y we need to choose a value for $\underline{x}$ in $\mathcal{X}$. We want good estimation of $\underline{\beta}$. Suppose the experimenter is allowed to take N independent observations on Y at vectors $\underline{x}_1, \underline{x}_2, \cdots, \underline{x}_N$ chosen by her/him from the set $\mathcal{X}$. The basic problem is : **How many observations should be taken at each point in $\mathcal{X}$ or what proportion of observations should be taken at each point in $\mathcal{X}$?** The basic idea is that we should choose $(\underline{x}_1, \underline{x}_2, \cdots, \underline{x}_N)$ to make the variance-covariance matrix of the estimators "as small as possible", or alternatively to make its inverse "as large as possible". Suppose that the N observations consist of n_i observations taken at $\underline{x}_i$, $i = 1, \cdots, m$. This is an exact design which is usually represented as follows

$$\begin{pmatrix} \underline{x}_1 & \underline{x}_2 & \cdots \underline{x}_m \\ n_1 & n_2 & \cdots n_m \end{pmatrix} \quad . \tag{1.1}$$

Model for an observation

Let Y_{il} denote the l^{th} observation obtained at $\underline{x}_i$. A model is

$Y_{il} = \beta_1 f_1(\underline{x}_i) + \cdots + \beta_k f_k(\underline{x}_i) + \varepsilon_{il} \qquad i = 1, \cdots, m$ and $l = 1, \cdots, n_i$

with $\mathbb{E}[\varepsilon_{il}] = 0$ and $\mathbb{V}[\varepsilon_{il}] = \sigma^2 \quad \forall \quad i, l.$

Throughout the rest of the text, we shall use the following notations:

Notation for the design The $N \times n$ matrix D of values of the explanatory variables can be given by the following :

$$D = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{11} & \cdots & x_{1n} \\ x_{21} & \cdots & x_{2n} \\ \vdots & & \vdots \\ x_{21} & \cdots & x_{2n} \\ \vdots & & \vdots \\ x_{m1} & \cdots & x_{mn} \\ \vdots & & \vdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix} \begin{matrix} \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} n_1 \text{ times} \\ \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} n_2 \text{ times} \\ \\ \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} n_m \text{ times} \end{matrix}$$

Notation for the Equation of the model : The general equation for the model under consideration can be written as :

$\underline{Y} = X\underline{\beta} + \underline{\varepsilon}$

Where $\underline{Y} = (Y_{11}, \cdots, Y_{1n_1}, \cdots, Y_{m1}, \cdots, Y_{mn_m})$ and

$\underline{\varepsilon} = (\varepsilon_{11}, \cdots, \varepsilon_{1n_1}, \cdots, \varepsilon_{m1}, \cdots, \varepsilon_{mn_m})$

The $N \times k$ matrix X with values $f_j(\underline{x}_i)$ is called the design matrix, and has the following form :

$$X = \begin{bmatrix} f_1(\underline{x}_1) & \dots & f_k(\underline{x}_1) \\ \vdots & & \vdots \\ f_1(\underline{x}_1) & \dots & f_k(\underline{x}_1) \\ \vdots & & \vdots \\ f_1(\underline{x}_m) & \dots & f_k(\underline{x}_m) \\ \vdots & & \vdots \\ f_1(\underline{x}_m) & \dots & f_k(\underline{x}_m) \end{bmatrix} \begin{matrix} \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} n_1 \text{ times} \\ \\ \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} n_m \text{ times} \end{matrix}$$

Note : In vector form, we can write the design matrix as : $X^T = \left[\underline{f}(\underline{x}_1) \cdots \underline{f}(\underline{x}_1) \cdots \underline{f}(\underline{x}_m) \cdots \underline{f}(\underline{x}_m)\right]^T$. For the least squares estimator $\underline{\hat{\beta}} = (\hat{\beta}_1, ..., \hat{\beta}_k)^T$ we have

$$\underline{\hat{\beta}} = (X^T X)^{-1} X^T \underline{Y} \ ,$$

where $\underline{Y}$ is the vector of observations,

$$\underline{Y} = (Y_1, ..., Y_N)^T \ .$$

For the variance-covariance matrix $\text{Cov}(\underline{\hat{\beta}})$ of the vector $\underline{\hat{\beta}}$ we have

$$\text{Cov}(\underline{\hat{\beta}}) = (X^T X)^{-1} \sigma^2 \ .$$

The predicted value of the response at $\underline{x}$ is

$$\begin{aligned} \hat{Y} &= \hat{\beta}_1 f_1(\underline{x}) + \cdots + \hat{\beta}_k f_k(\underline{x}) \\ &= \left(\underline{f}(\underline{x})\right)^T \underline{\hat{\beta}}, \end{aligned}$$

with

$$\begin{aligned} \underline{f}(\underline{x}) &= \left(f_1(\underline{x}), \cdots, f_k(\underline{x})\right)^T, \\ \mathbb{V}(\hat{Y}) &= \sigma^2 \left(\underline{f}(\underline{x})\right)^T (X^T X)^{-1} \underline{f}(\underline{x}), \end{aligned}$$

and the standardized variance is

$$\begin{aligned} d(x, \mathcal{E}(N)) &= \frac{\mathbb{V}(\hat{Y})}{\sigma^2} \\ &= \left(\underline{f}(\underline{x})\right)^T (X^T X)^{-1} \underline{f}(\underline{x}). \end{aligned}$$

The matrix X^TX is very important and is called the **information matrix** M. This matrix is also equal to

$$M = X^TX = \sum_{i=1}^{m} n_i \underline{f}(\underline{x}_i)(\underline{f}(\underline{x}_i))^T ,$$

with

$$\underline{f}(\underline{x}_i) = (f_1(\underline{x}_i), f_2(\underline{x}_i), ..., f_k(\underline{x}_i))^T .$$

This last notation for M will be useful in finding optimal designs. To emphasize the fact that the information matrix depends on the choice of the design we also write $M(\mathcal{E})$ or $M(\mathcal{E}(N))$. Now we have

$$Cov(\hat{\beta}) = \sigma^2 M^{-1}$$

1.2 How do we measure the quality of a design?

In the previous section, we gave a brief definition of a design, together with the definitions of some other fundamental concepts used in the study of optimal design of experiments.

We now focus on the criteria that are commonly used to measure the quality of a design. As we shall see later on, the choice of a design will always depend on our interest. If our interest is parameter estimation for instance, we therefore will want to choose a design that would **minimize** in some sense the variance-covariance of the parameter estimator of our model.

1.2.1 Criteria for the quality of a design

In the current statistics literature, there are many criteria used to measure the quality of a design, among which are

- **D-optimality:** a design is D-optimum if it maximizes the value of $det(M(\mathcal{E}(N)))$ or $\log det(M(\mathcal{E}(N)))$, i.e. the generalized variance of the parameter estimates is minimized,

- **c-optimality:** in c-optimality the interest is in estimating the linear combination of the parameters $\underline{c}^T\underline{\beta}$ with minimum variance. The criterion to be minimized is therefore

$$Var(\underline{c}^T\hat{\underline{\beta}}) \quad \propto \quad \underline{c}^T M^{-1}(\mathcal{E}(N))\underline{c}$$

  where c is $p \times 1$ (A disadvantage of c-optimum designs is that they are often singular.);

- **G-optimality:** a G-optimum design $\mathcal{E}^*(N)$ minimizes the maximum over the design region $\mathcal{X}$ of the standardized variance i.e. $\mathcal{E}^*(N)$ solves $minmax\, \underline{f}^T(\underline{x})M^{-1}(\mathcal{E}(N))\underline{f}(\underline{x}) \quad \underline{x} \in \mathcal{X}$ (this minimax value equals k);

- **A-optimality:** minimize the sum (or avarage) of the variances of the parameter estimates;

- **E-optimality:** minimises $\lambda_{max}(M^{-1})$, the maximum eigen-value of M^{-1}, where $M = M(\mathcal{E}(N))$.

All these criteria are functions of the variance-covariance matrix of the parameter estimates, and this justifies the central role that the information matrix plays in the determination of the optimal design.

For the purpose of our study, we shall be using D-Optimality.

1.2.2 Introducing D-Optimality

The study of D-optimality has been central to the work on optimum experimental design since the begining e.g. Kiefer (1959). Fedorov(1972), Silvey(1980), and Pazman (1986) likewise stress D-optimality. Farrell *et al.* (1967), give a summary of earlier work on D-optimality. This includes that of Kiefer and Wolfowitz (1961) and Kiefer (1961) which likewise concentrate on results for regression models, including extentions to D_s-optimality. [See Atkinson and Donev (1992)].

The D-criterion is known as the criterion based on the generalized variance of $\hat{\underline{\beta}}$, that is the determinant of the information matrix.
More precisely, D-optimality will consist in determining the design that maximizes the determinant, $|M(\mathcal{E}(N))|$, of the information matrix.
In fact, there is a relation between this determinant and a confidence region for the vector of unknown parameters. Assuming that ε is normal, the confidence region for β is defined by

$$(\underline{\beta} - \hat{\underline{\beta}})^T X^T X (\underline{\beta} - \hat{\underline{\beta}}) < k \ s^2 F_{\alpha;k,N-k} \ ,$$

where $F_{\alpha;k,N-k}$ is the critical value of the F distribution with k and $(N-k)$ degrees of freedom and where

$$s^2 = \frac{1}{N-k}(\underline{Y} - X\hat{\underline{\beta}})^T(\underline{Y} - X\hat{\underline{\beta}})$$

is an unbiased estimator for σ^2. This confidence region is an ellipsoid. The volume of this ellipsoid is proportional to $(\det(X^TX)^{-1})^{\frac{1}{2}}$. *So a D-optimal design is a design which minimises the volume of this ellipsoid since $M(\mathcal{E}(N)) = X^TX$.*
In the following we will consider D-optimality. An advantage of D-optimality is that the optimum designs for quantitative factors do not depend upon the scale of the variables. This criterion is invariant with respect to a linear transformation of the form $\Theta = A\underline{\beta}$. Except for G-optimality, the other criteria do not have this important property.

The value of det $(M(\mathcal{E}(N)))$ depends on the number of observations. If we have a design $\mathcal{E}(N)$, then we can easily improve the design by choosing a design $\mathcal{E}(2N)$ that consists is doubling the replication of the points of the design $\mathcal{E}(N)$. In this case we have

$$\det(M(\mathcal{E}(2N))) = 2^k \det(M(\mathcal{E}(N))) \ .$$

Therefore it is not useful to compare designs with different numbers of values of N. In the literature though, there is a special treatment of D-optimality that addresses the case of designs with a fixed number of observations. That criterion is called D_N-optimality.

The D-optimal criterion has been the most commonly used, and has dominated the literature of optimal designs; [see, Fedorov (1972), Silvey (1980)].
Let $\Phi(M) = \log\left[det(M)\right]$. The properties of the D-optimality criterion include :

1. $\Phi(M)$ is an increasing function over the set of positive definite symmetric matrices. That is for $M_1, M_2 \in \mathcal{M}$, then $\Phi(M_1 + M_2) \geq \Phi(M_1)$ where $\mathcal{M}$ is the set of all non negative definite symmetric matrices.

2. The function $\Phi(M(\mathcal{E}(N)))$, where $M(\mathcal{E}(N))$ is the information matrix of the design $\mathcal{E}(N))$, is a strictly concave function on the set $\mathcal{M}$ [see Fedorov (1972)]

3. $\Phi(M)$ is differentiable whenever it is finite.

4. D-optimal designs are invariant with respect to any non degenerate linear transformation of the estimated parameters. [see Fedorov (1972)]

1.2.3 Optimal designs with fixed N

Definition 1.1. D_N***-optimality :***

A design $\mathcal{E}^*(N)$ is D_N-optimal if $M(\mathcal{E}^*(N))$ maximizes $det(M(\mathcal{E}^*(N)))$ over all N point designs.

$$det(M(\mathcal{E}^*(N))) = max_{\mathcal{E}(N)}\ det(M(\mathcal{E}(N))) \ .$$

1.2.4 Normalized Designs

We now focus on the comparison of designs with different numbers of observations, and this is done by using normalized designs, which requires us to standardize designs. This will be discussed in the next section.

Consider a design with N observations. In the point $\underline{x}_i$ we have n_i observations. Another way to say this is that a fraction n_i/N of the total number of observations is taken at the point $\underline{x}_i$.

This consideration gives the following definition.

Definition 1.2. *Exact design : An exact normalized design p_N has the form*

$$p_N = (\underline{x}_1, ..., \underline{x}_m; p_1, ..., p_m) \ ,$$

and there exists integer N such that $p_i = n_i/N$ $(n_i \leq N)$.

Definition 1.3. *Discrete design : A discrete normalized design p has the form*

$$p = (\underline{x}_1, ..., \underline{x}_m; p_1, ..., p_m) \ ,$$

where $p_i \in \mathcal{R}, p_i > 0$ and $\sum_{i=1}^{m} p_i = 1$.

Definition 1.4. *Continuous Design : A continuous design is characterized by a measure ξ on the experimental region χ.*

The names exact, discrete and continuous are confusing. A measure ξ may also be discrete of course. The name exact has been chosen because an exact design can also be performed in practice.

A discrete design may also be exact (if $p_i \in$ irrational numbers for all i). Every discrete design can be approximated by an exact design with a large number of observations N.

We now define the per observation information matrix of a continuous design ξ to be

$$M(\xi) = \int_{\mathcal{X}} f(\underline{x})(f(\underline{x}))^T d\xi(x)$$

and in the case of an absolutely continous measure, i.e. a measure with a density $p(x)$ we define it to be

$$M(p) = \int_{\mathcal{X}} p(x)\ \underline{f}(\underline{x})(\underline{f}(\underline{x}))^T dx\ =\ E_p\{\underline{f}(\underline{x})\underline{f}^T(\underline{x})\}$$

with

$$\int_{\mathcal{X}} p(x)dx = 1\ .$$

If the design is discrete or exact, then we have

$$M(p) = \sum_{i=1}^{m} p_i\ \underline{f}(\underline{x}_i)(\underline{f}(\underline{x}_i))^T\ .$$

The standardized variance function $d(x,p)$ equals

$$d(\underline{x},p)\ =\ (\underline{f}(\underline{x}))^T M^{-1}(p)\underline{f}(\underline{x}) \tag{1.2}$$

Continuous designs can be useful to find optimal designs. They do not have any practical meaning. They are just useful in as much as they help in finding the optimal design analytically. We study only Discrete designs.

Theorem 1.1. *For any given design p there exists a discrete design $p' = (\underline{x}_1, ..., \underline{x}_l; p_1, ..., p_l)$ with*

$$M(p) = M(p')\ ,$$

and

$$l \leq \frac{(k+1)k}{2} + 1\ .$$

Proof: See Fedorov (1972) page 66-67.

Theorem 1.2. *Let p be a normalized design with variance function $d(\underline{x}, p)$. Then we have in the case of a discrete design*

$$\sum_{i=1}^{m} p_i \; d(\underline{x}_{i*}, p) = k \; ,$$

where $\underline{x}_{i*} = (x_{i1}, x_{i2}, \cdots, x_{in})^T \in \mathcal{X}$ *and in the case of a continuous design*

$$\int_{\mathcal{X}} d(\underline{x}, p) d\xi(\underline{x}) = k \; .$$

Proof. We can recall the variance function from equation 1.2:

$$d(\underline{x}_i, p) = d_i = (\underline{f}(\underline{x}_i))^T M^{-1}(p) \underline{f}(\underline{x}_i) \tag{1.3}$$

By equation 1.3

$$\begin{aligned}
\sum_{i=1}^{n} p_i d_i &= \sum_{i=1}^{n} p_i (\underline{f}(\underline{x}_i)^T M^{-1}(p) \underline{f}(\underline{x}_i) \\
&= tr\{M^{-1}(p) \sum_{i=1}^{n} p_i (\underline{f}(\underline{x}_i)) \underline{f}^T(\underline{x}_i)\} \\
&= tr\{M^{-1}(p) M(p)\} \\
&= tr I_k \\
&= k
\end{aligned}$$

which is what was required to prove. [See Fedorov (1972) page 68-69.] □

A proof for a continuous design is similar.

Theorem 1.3. *For the maximum value of the function $d(\underline{x}, p)$ we have*

$$\max_{\underline{x} \in \mathcal{X}} d(\underline{x}, p) \geq k \; .$$

Proof: See Fedorov (1972). We now define D-optimality independently of the number of observations.

Definition 1.5. *A design p' is called D-optimal if*

$$det(M(p')) = \max_p det(M(p)) \ ,$$

where maximization is with respect to all possible designs measures, descrete or continous.

1.2.5 Weighted Linear Regression Design

An example of a linear design problem can be a design for a weighted regression model. As we shall see later, the simplest case of Weighted Linear Regression plays a central role in our D-optimal design problems. See also Torsney and Musrati(1993).

- **Model :**

$$\begin{aligned} \mathbb{E}(y) &= \alpha + \beta z, \qquad z \in Z = [a, b] \\ \mathbb{V}(y) &= \frac{\sigma^2}{w(z)}, \end{aligned}$$

  for some weight function w(z). α, β and σ ($\sigma > 0$) are unknown parameters and Z is the design space.

- **Design :**

  Design points $z_1, z_2, \cdots, z_r$, $z_i \in Z$ with weights $p_1, p_2, \cdots, p_r$ where the variables p_i are nonnegative and sum to 1. i.e.

$$\sum_{i=1}^{r} p_i = 1, \qquad 0 \leq p_i \leq 1.$$

- **Information Matrix :**

  The informatiom matrix is of the form

$$\begin{aligned} M(p) &= \sum_{i=1}^{r} p_i w(z_i) \begin{pmatrix} 1 \\ z_i \end{pmatrix} (1 \ , z_i) \\ &= \mathbb{E}_p \left\{ w(z) \begin{pmatrix} 1 \\ z \end{pmatrix} (1 \ , z) \right\} \\ &= \mathbb{E}_p \left\{ \underline{g}\underline{g}^T \right\} \qquad \text{where} \quad \underline{g} = (g_1, g_2)^T \text{and } g_1 = \sqrt{w(z)}, \ g_2 = z g_1. \end{aligned}$$

1.3 Design for a general nonlinear model

Suppose some response variable y has probability model $p(y|\underline{x}, \underline{\theta})$ where $\underline{x}$ is a given explanatory or design variable in the design space $\mathcal{X}$ $[\underline{x} \in \mathcal{X}]$ and $\underline{\theta}$ is a k-vector of unknown parameters $[\underline{\theta} \in \mathcal{R}^k]$. Suppose further that it depends on $\underline{\theta}$ only through its expectation function,

$$\mathbb{E}(y|\underline{x}, \underline{\theta}) = \eta(\underline{x}, \underline{\theta})$$

where η is a known non-linear function of parameters $\underline{\theta}$ and $\underline{x}$. Also let

$$\mathbb{V}(y|\underline{x}, \underline{\theta}) = a(\underline{x}, \underline{\theta})$$

where $a(\underline{x}, \underline{\theta})$ is a known conditional variance function. All observations are assumed to be conditionally independent.

Suppose that N observations are taken and consist of n_i observations taken at $\underline{x}_i$ $i = 1, \cdots, m$. This is the exact design

$$\begin{pmatrix} \underline{x}_1 & \underline{x}_2 & \cdots \underline{x}_m \\ n_1 & n_2 & \cdots n_m \end{pmatrix} \quad . \tag{1.4}$$

Asymptotic Covariance Matrix : Assuming we estimate θ by maximum likelihood and assuming the standard asymptotic results: Let $\hat{\theta}_{ML}$ denote the Maximum Likelihood (ML) estimator of $\underline{\theta}$. Then from standard asymptotic theory $\hat{\underline{\theta}}_{ML}$ is asymptotically unbiased, efficient (and normally distributed) i.e.

$$\begin{aligned} \mathbb{E}(\hat{\underline{\theta}}_{ML}) &\approx \underline{\theta}, \qquad N \quad large \\ Cov(\hat{\underline{\theta}}_{ML}) &\approx \left[\sum_{i=1}^{m} n_i I(\underline{x}_i, \underline{\theta}) \right]^{-1} \end{aligned}$$

where $I(\underline{x}_i, \underline{\theta})$ is the Fisher Information Matrix for $\underline{\theta}$ for a single observation at the point $\underline{x}_i$ and is given by

$$I(\underline{x}, \underline{\theta}) = \frac{1}{a(\underline{x}, \underline{\theta})} \eta_\theta(\underline{x}, \underline{\theta}) \eta_\theta(\underline{x}, \underline{\theta})^T$$

where η_θ denotes the vector of partial derivatives

$\eta_\theta(\underline{x}, \underline{\theta}) = [\partial\eta/\partial\theta_1, \cdots, \partial\eta/\partial\theta_k]^T$ and $a(\underline{x}, \underline{\theta}) = \mathbb{V}(\boldsymbol{y}|\underline{x})$.

Under the above design

$$Cov(\hat{\underline{\theta}}) = [\Sigma_{i=1}^m n_i I(\underline{x}_i, \underline{\theta})]^{-1}$$

Approximate design -Continuous design

Equivalently, let $p_i = \frac{n_i}{N}$ so that p_i is the proportion of observations taken at $\underline{x}_i$.

Then for large N

$$\begin{aligned} Cov(\hat{\underline{\theta}}) &= [N\Sigma_{i=1}^m p_i I(\underline{x}_i, \underline{\theta})]^{-1} \\ &\propto [\Sigma_{i=1}^m p_i I(\underline{x}_i, \underline{\theta})]^{-1} = [M(\underline{\theta}, p)]^{-1}. \end{aligned}$$

These weights (proportions) define an approximate design. If the design has trials at m distinct points in $\mathcal{X}$, taking a proportion p_i of observations at $\underline{x}_i$ $i = 1, \cdots, m$ we denote it by

$$p = (\underline{x}_1, \ldots, \underline{x}_m; p_1, \ldots, p_m).$$

Clearly $\Sigma p_i = 1$, $p_i \geq 0$. In general we may consider continuous design measures $\xi(\cdot)$ satisfying measure $\int_{\mathcal{X}} \xi(d\underline{x}) = 1$ and $\xi(\underline{x}) \geq 0$ for which the per observation information matrix is

$$M(\underline{\theta}, \xi) = \int I(\underline{x}_i, \underline{\theta}) d\xi(\underline{x}).$$

1.4 Design Criteria

These must now be functions of $M(\underline{\theta}, \xi)$. In particular the D-optimal criterion is

$$log\{det[M(\underline{\theta}, \xi)]\} = -log\{det[M(\underline{\theta}, \xi)]^{-1}\}.$$

Again this has a confidence ellipsoid interpretation.

Suppose that $\hat{\underline{\theta}}$ is the M.L estimator of $\underline{\theta}$ obtained from data arising under a design ξ chosen on the provisional assumption that $\underline{\theta} = \tilde{\underline{\theta}}$. Then log likelihood confidence regions for $\underline{\theta}$ can be closely approximated by ellipsoids of the form

$$\{\underline{\theta} : (\underline{\theta} - \hat{\underline{\theta}})^T M(\xi, \tilde{\underline{\theta}})(\underline{\theta} - \hat{\underline{\theta}}) \leq \text{constant}\}.$$

The volume of the above ellipsoid is proportional to $\{det[M(\xi, \tilde{\underline{\theta}})]\}^{-1/2}$. So maximising $\log det[M(\xi, \tilde{\underline{\theta}})]$ would be equivalent to minimising the volume of confidence ellipsoids for $\underline{\theta}$ of the above form. *That is, we are making our confidence regions, in some sense, as small as possible.*

To make the Information matrix small we should choose p_i optimally. In practice, observations will be taken at a finite subset of points. We focus on the proportion p_i of observations taken at $\underline{x}_i$ for good estimation. The objective is to choose ξ to maximize $det(M(\xi, \theta))$ which is to minimize the volume of a confidence ellipsoid for $\underline{\theta}$.

1.5 Optimality Conditions

We need conditions for identifying optimality. An important result in this connection is the General Equivalence Theorem. It can be viewed as an extension of the result that the derivatives are zero at an unconstrained maximum (or minimum) of a function.

The derivative of $\Phi(\cdot)$ at $M(\xi_1)$ in the direction of $M(\xi_2)$, see Whittle (1973) is

$$F_\Phi(M(\xi_1), M(\xi_2)) = \lim_{\alpha \to 0^+} \frac{1}{\alpha}\left[\Phi\{(1-\alpha)M(\xi_1) + \alpha M(\xi_2)\} - \Phi\{M(\xi_1)\}\right].$$

The General Equivalence Theorem states the equivalence of the following three conditions on ξ^* :

If $\Phi(M)$ is strictly concave on the set of symmetric positive definitive matrices then:

- the design ξ^* maximises $\Phi\{M(\xi)\}$,

- the design ξ^* minimizes the maximum over $\underline{x} \in \mathcal{X}$ of $F_\Phi(M(\xi^*), I(\underline{x}, \underline{\theta}))$ (=the minimum of $F_\Phi(M(\xi^*), I(\underline{x}, \underline{\theta}) \geq 0$),
- the derivative $F_\Phi(M(\xi^*), I(\underline{x}, \underline{\theta}))$ achieves its maximum of zero at the support points of the design $\xi^*(\underline{x})$. i.e. $F_\Phi(M(\xi^*), I(\underline{x}, \underline{\theta})) = 0$ if $\xi^*(\underline{x}) > 0$.

In summary

$$F_\Phi(M(\xi^*), I(\underline{x}, \underline{\theta})) \begin{cases} = 0 & if \quad \xi^*(\underline{x}) > 0 \\ \leq 0 & if \quad \xi^*(\underline{x}) = 0 \end{cases}$$

For $\Phi = \log det(M)$ which is strictly concave

$$\begin{aligned} F_\Phi(M(\xi_1), M(\xi_2)) &= tr(M^{-1}(\xi_1)(M(\xi_2) - M(\xi_1))) \\ &= tr(M^{-1}(\xi_1)M(\xi_2)) - k \end{aligned}$$

where $k = tr(M^{-1}(\xi_1)M(\xi_1))$,

$$\begin{aligned} F_\Phi(M(\xi_1), I(\underline{x}, \underline{\theta})) &= tr(M^{-1}(\xi_1)I(\underline{x}, \underline{\theta})) - k \\ &= tr\{(M^{-1}(\xi_1))\frac{\eta_{\underline{\theta}}\eta_{\underline{\theta}}^T}{a(\underline{x}, \underline{\theta})}\} - k \\ &= \frac{\eta_{\underline{\theta}}^T M^{-1}(\xi_1)\eta_{\underline{\theta}}}{a(\underline{x}, \underline{\theta})} - k \end{aligned}$$

So ξ^* is D-optimal if only if

$$\begin{aligned} \frac{1}{a(\underline{x}, \underline{\theta})}\eta_{\underline{\theta}}^T M^{-1}(\xi^*)\eta_{\underline{\theta}} &\leq k \qquad \forall \quad \eta_{\underline{\theta}} \\ &= k \qquad \xi^*(\underline{x}) > 0 \ . \end{aligned}$$

Note this defines an ellipsoid centred on the origin containing the set $\{\eta_{\underline{\theta}} = \eta(\underline{x}, \underline{\theta}) : \underline{x} \in \mathcal{X}\}$ with the support points of ξ^* on the boundary. Silvey (1972) conjectured that this was the smallest such ellipsoid and Sibson (1972) proved the conjecture to be true.

Chapter 2

Weighted Regression Model Construction of D-optimal Design : The Case of the Two Parameter Model.

2.1 Model under consideration

We consider a binary regression model in which the observed variable u depends on a design variable $x \in \mathcal{X} = [c, d] \subset \mathcal{R}$. u can take only two possible values, according as some event of interest occurs $u = 1$ or does not $u = 0$. We may write the probabilities of the two outcomes as follows:

$$\mathbf{Pr}(u = 0|x) = 1 - \pi(x) \qquad \mathbf{Pr}(u = 1|x) = \pi(x)$$

Namely, $u \sim Bi(1, \pi(x))$. We assume $\pi(x) = F(\alpha + \beta x)$, where $F(\cdot)$ is a chosen cumulative distribution function. So

$$\begin{aligned} \mathbb{E}(u|x) &= \pi(x) = F(\alpha + \beta x) \\ \mathbb{V}(u|x) &= \pi(x)[1 - \pi(x)] \end{aligned}$$

Crucially the dependence of π on x occurs only through a nonlinear function of the linear combination

$$z = \alpha + \beta x \tag{2.1}$$

for unknown parameters α, β. This is an example of a generalized linear model.

2.1.1 Design For Binary Regression

We now apply the design theory on Chapter 1 to our binary regression model. It is convenient to adopt the parameter dependent linear transformation $z = \alpha + \beta x$. For the above model the information matrix can be written as follows

$$I(x, \underline{\theta}) = \frac{f^2(z)}{F(z)[1 - F(z)]} \begin{pmatrix} 1 \\ x \end{pmatrix} (1\ , x)$$

where $f(z) = F'(z)$ and

$$\begin{aligned} \eta &= \pi(x) \\ &= F(\alpha + \beta x), \qquad z = \alpha + \beta x \\ &= F(z) \end{aligned}$$

and

$$\begin{aligned} a(\underline{x}, \underline{\theta}) &= \mathbb{V}(u|x) \\ &= \pi(x)[1 - \pi(x)] \\ &= F(\alpha + \beta x)[1 - F(\alpha + \beta x)] \\ &= F(z)[1 - F(z)] \end{aligned}$$

and

$$\begin{aligned}\eta_\theta &= \left[\frac{\partial F(z)}{\partial z}\frac{\partial z}{\partial \alpha}, \frac{\partial F(z)}{\partial z}\frac{\partial z}{\partial \beta}\right]^T \\ &= \left[f(z), f(z)x\right]^T \\ &= f(z)\begin{pmatrix}1\\x\end{pmatrix} .\end{aligned}$$

Now let the vector

$$\begin{aligned}\underline{v} &= \frac{1}{\sqrt{\mathbb{V}(u|x)}}\left[\frac{\partial F(z)}{\partial z}\frac{\partial z}{\partial \alpha}, \frac{\partial F(z)}{\partial z}\frac{\partial z}{\partial \beta}\right]^T \\ &= \frac{f(z)}{\sqrt{F(z)[1-F(z)]}}\begin{pmatrix}1\\x\end{pmatrix} .\end{aligned}$$

Further, given $z = \alpha + \beta x$, then $z \in [a, b]$, (a,b determined by c,d) and

$$\begin{aligned}\begin{pmatrix}1\\z\end{pmatrix} &= \begin{pmatrix}1 & 0\\ \alpha & \beta\end{pmatrix}\begin{pmatrix}1\\x\end{pmatrix} \\ &= \mathcal{B}\begin{pmatrix}1\\x\end{pmatrix}\end{aligned}$$

Hence if

$$\underline{g}(z) = \mathcal{B}\underline{v},$$

then

$$\underline{g}(z) = \frac{f(z)}{\sqrt{F(z)[1-F(z)]}}\begin{pmatrix}1\\z\end{pmatrix}$$

and

$$\underline{v} = \mathcal{B}^{-1}\underline{g}(z) .$$

D-optimality is invariant under non-singular linear transformations of the design space. So as did Ford, Torsney and Wu (1992) we consider the D-optimal linear design problem with design vectors

$$\underline{g} = \sqrt{w(z)}(1\ z)^T \qquad\qquad z \in [a, b]$$

where $w(z) = \frac{f^2(z)}{F(z)[1-F(z)]}$. This corresponds to a weighted linear regression design problem with weight function $w(z)$.

Therefore these nonlinear design problems transform to linear design problems for weighted linear regression in z with weight function $w(z) = \frac{f^2(z)}{F(z)(1-F(z))}$, where $f(z) = F'(z)$ is the density of $F(\cdot)$.

Table 2.1 lists examples of this kind of weight function (binary regression weight functions) in two groups: Group I and Group II. Two other groups (III and IV) which we will consider are also listed. Firstly, we consider finding D-optimal designs for Group I (Table 2.2) and Group III (Table 2.3), and then investigate Group II (Table 2.8) and Group IV (Table 2.11) separately.

In Table 2.2 we list details of the binary weight functions in Group I, namely the Logistic, the Skewed Logistic, the Generalized Binary, the Complementary log-log and the Probit; details are **the pdf, the cdf, explicit formulae for the weight functions and the support points of the two parameter case (global) D-optimal design.**

In Table 2.8 we give the same information for the two special binary weight functions of Group II, namely the Double Reciprocal and the Double Exponential weight functions.

Table 2.3 records the explicit formulae for the weight functions of Group III, namely the beta, the gamma and the normal **density functions**, together with the corresponding **support points of the global D-optimal designs.**

Finally, Table 2.11, shows the corresponding information for the weight functions of Group IV.

2.2 Characterisation of Optimal Designs

Let ξ^* be a design measure on $[a, b]$. ξ^* is D-optimal iff

$$\begin{aligned} \mathcal{V}(z) = \underline{g}^T(z)M^{-1}(\xi^*)\underline{g}(z) &\leq 2 \qquad \xi^*(z) = 0 \\ &= 2 \qquad \xi^*(z) > 0. \end{aligned}$$

$\mathcal{V}(z)$ is known as the variance function. This defines Silvey's minimal ellipse.

It is useful to introduce the following set:

$$G = G(\mathcal{Z}) = \{\underline{g}(z) : \begin{pmatrix} g_1 \\ g_2 \end{pmatrix}, \quad g_1 = \sqrt{w(z)} \qquad g_2 = zg_1, z \in \mathcal{Z}\}. \quad (2.2)$$

We call this an induced design space as did Ford *et al.* (1992). An alternative and probably better name would **design locus** as orginally as used by Box and Lucas (1959). In this two dimensional case, Silvey's geometrical characterization can provide us with some insights into the support points of a D-optimal design or at least their number. The support points are the points of contact between $G(\mathcal{Z})$, the design space, and the smallest ellipse ($SE(G)$) centred on the origin which contains $G(\mathcal{Z})$. The idea was first conjectured by Silvey (1972), and proved by Sibson (1972), both of these being contributions to discussion of Wynn (1972). Pictures of $G(\mathcal{Z})$ are important.

Our objective is to find D-optimal designs for all possible interval subsets $\mathcal{Z} = [a, b]$ of $\mathcal{Z}_w$, where $\mathcal{Z}_w$ is the widest possible design space.

- **Case 1** : $\mathcal{Z} = \mathcal{Z}_w$

  We consider $\mathcal{Z} = \mathcal{Z}_w$ initially for all of the above weight functions Beta, Gamma, Normal and Binary. The induced design space G is **a closed convex curve** in $\mathcal{R}^2$ for the widest choice of $\mathcal{Z}$ $(= \mathcal{Z}_w)$. For these cases in Figure 2.1 and Figure 2.2 it seems likely that the minimal central ellipsoid containing $G(\mathcal{Z}_w)$ can only touch it twice, in which case the D-optimal design has two support points and must be the best two point design.

Consider a design with 2 support points u, v. For this to be D-optimal on this support the weights must be $1/2, 1/2$. Denote this design by ξ. Then the **best** two point design on $\mathcal{Z}_w$ must maximise

$$det(M(\xi))$$

with respect to u, v. Let a^*, b^* be the optimal values of u, v. These are listed in Table 2.2 and Table 2.3. All of these weight functions $w(\cdot)$ have similar properties. In particular they are typically unimodal with one maximal Turning Point at z_{max} ($w'(z_{max})) = 0$, and $w'(z) > 0$ if $z < z_{max}$ and $w'(z) < 0$ if $z > z_{max}$) [see Figure 2.3 and Figure2.4].

In general support points must be found numerically. However for some of the weight functions (Beta, Gamma, Normal) there is an explicit solution for the support points, see Table 2.3 [Karlin and Studden (1966) Fedorov (1972), Torsney and Musrati (1993)].

In most cases this best 2 point design is the D-optimal design. The Equivalence Theorem is satisfied in all cases except for the double exponential and the double reciprocal .

- **Case 2 - 5** :

  We now consider those weight functions for which the above D-optimal design is a two-point design.

- **Case 2**: $\mathcal{Z} = [a, b]$ $a \leq a^*$ $b \geq b^*$.

  The D-optimal design is the same as above since $\mathcal{Z}$ contains the above support points.

We now repeat the conjectures of Ford, Torsney, Wu (1992) Ford *et al.* (1992) for three other cases and extend these to the non-binary weight functions listed in the Table 2.3.

We conjecture that the D-optimal designs ξ have the two support points listed in the following three cases. (We assume $a < b$.)

- **Case 3** : $a \geq a^*$, $b \geq b^*$

  $Supp(\xi) = \{a, \, min\{b, \, b^*(a)\}\}$, $(b^*(a) \, > \, b^*)$

- **Case 4** : $a \leq a^*$ $b \leq b^*$

  $Supp(\xi) = \{max\{a, \, a^*(b)\}, b\}$, $(a^*(b) \, < \, a^*)$

- **Case 5** : $a \geq a^*$ $b \leq b^*$

  $Supp(\xi) = \{a, \, b\}$

Here $b^*(a)$ maximises $det(M(\xi))$ with respect to d (**over** $d \geq a$) where ξ is the design

$$\xi \;=\; \binom{a \;\; d}{1/2 \; 1/2} \;,$$

and $a^*(b)$ maximises $det(M(\xi))$ with respect to c (**over** $c \, \leq \, b$) where ξ is the design

$$\xi \;=\; \binom{c \;\; b}{1/2 \; 1/2} \;.$$

2.2.1 Justification of the Conjecture

The D-optimal design must satisfy the Equivalence Theorem. According to the theorem (Silvey, 1980), a design $\xi(\cdot)$ is D-optimal iff

$$\begin{aligned} w(z)(1 \; z)M^{-1}(\xi) \begin{pmatrix} 1 \\ z \end{pmatrix} \;&\leq\; 2 \quad \forall \quad z \in \mathcal{Z} \qquad (2.3) \\ &=\; 2 \quad \text{if} \quad \xi(z) > 0. \qquad (2.4) \end{aligned}$$

This is true iff

$$\begin{aligned} v(z) \;=\; \frac{1}{2}Q(z) - \frac{1}{w(z)} \;&\leq\; 0 \quad \forall \quad z \in \mathcal{Z} \\ &=\; 0 \quad \text{if} \quad \xi(z) > 0, \end{aligned}$$

where $Q(z) = (1 \ z)M^{-1}(\xi)\begin{pmatrix}1\\z\end{pmatrix}$ is a quadratic function. **So for an optimal design we wish to see** $v(z) \leq 0 \ \forall z \in \mathcal{Z}$. To explore the shape of $v(z)$ we analyze its derivatives. The derivative of $v(z)$ can be written as

$$v'(z) \quad = \quad L(z) - H(z) \ , \tag{2.5}$$

where $H(z) = \frac{-w'(z)}{[w(z)]^2}$ and $L(z)$ is an increasing linear function of z because the coefficient of z is the second diagonal element of the design Matrix $M(\xi)$ which is positive definite. In fact $L(z) = (2E(w(Z))z - 2E(Zw(Z)))/Det(M(\xi))$ where Z is a random variable with probability measure ξ since

$$M(\xi) \quad = \quad \begin{pmatrix} \mathbb{E}(w(Z)) & \mathbb{E}(Zw(Z)) \\ \mathbb{E}(Zw(Z)) & \mathbb{E}(Z^2w(Z)) \end{pmatrix}.$$

The intercept will be negative if $E(Zw(Z))$ is positive and vice versa. The consequence is $v'(z) = 0$ iff $L(z) = H(z)$. That is, $v'(z) = 0$ when the line $L(z)$ crosses $H(z)$.

A question of interest is : **"How many times can an increasing line** $L(z)$ **cross the function** $H(z)$ **?"** Plots of $H(z)$ are given in Figure 2.5 and Figure 2.6 for various weight functions $w(\cdot)$. These plots (appear to) have similar shapes and properties. In particular let $\mathcal{Z}_w = [A, B]$. Then $H(A) = -\infty$, $H(B) = +\infty$ and $H(z)$ is concave increasing up to some point and thereafter is convex increasing. Also $H'(A) = \infty$, $H'(B) = \infty$, while the second derivative of $H(z)$ has **one change of sign** for all the weight functions considered. This was observed empirically in most cases. Only a few of them like the logistic and the normal weight functions offer an $H(z)$ function whose change of sign can be seen analytically.

Given such an $H(z)$, an upward sloping line $L(z)$ can cross it, over the whole range of z, either one or three times. This depends on **the slope** of the line. This means that the derivative of $v(z)$ can have at most 3 zeros in $(-\infty, \infty)$. Further such a line must initially lie above $H(z)$. So if there is only one **Turning**

Point it is a maximal one, or if there are three the first is a maximal one and hence so is the third while the second is a minimal TP. So $v(z)$ has only one minimum turning point (TP) and at most two maximum TP's. Hence given three solutions to $v'(z) = 0$ the middle one must be a minimum turning point. (The line crosses first from above, then from below, then from above, then from below the curve [Figure 2.7].)

Consequently, this implies that there are two support points, because three support points would need two minimum TPs. As a result of this, all the above weight functions have two support points with optimal weights $\frac{1}{2}$. We list the $H(z)$, $H'(z)$ and $H''(z)$ functions for the Binary weight functions and those of Group III, Table 2.4 and Table 2.5, respectively.

We note that a some upword sloping lines may only cross $H(z)$ once from above in which case $v(z)$ would only have one maximal TP while others might be tangential either to the concave or convex section of $H(z)$ in which case $v(z)$ has one maximal TP and one point of inflexion. In either this means that a horizantal line can only cross $v(z)$ twice. More over $v(z)$ lies above any such line between c and d where these are the values of z at which crossing takes place. This cannot be the case if $v(z)$ arises under a design which is D-optimal on an interval say $[c, d]$. We must have $v(z) \leq 2$ on $[c, d]$.

Hence the lines arising under designs ξ must cross the $H(z)$ three times.

2.2.2 Determination of support points binary regression case

We now need to establish what the support points are of these D-optimal two-point designs. We consider the arbitrary design

$$\xi = \binom{z_1 \;\; z_2}{1/2 \; 1/2} .$$

Then

$$det(M(\xi)) = (1/2)^2 (z_2 - z_1)^2 w(z_1)w(z_2) \qquad z_1 < z_2.$$

We take the **log** function of the ***determinant*** which is a concave function of $M(\cdot)$.

$$\ln[detM(\xi)] = -2\ln 2 + 2\ln(z_2 - z_1) + \ln w(z_1) + \ln w(z_2)$$

We note that we will be interested in the derivatives of this function with respect to z_1 and/or z_2. To find the best two-point design on $\mathcal{Z}_w$ we need to maximise $\ln[det(M(\xi))]$ w.r.t. z_1 and z_2; or if we wish to find the best two point design subject to z_1 (or z_2) being a support point we need to maximise $\ln[det(M(\xi))]$ w.r.t. z_2 (or z_1). So we consider derivatives w.r.t. z_1 and z_2. A rearrangement of the first order stationary conditions introduces a function $h(z)$.

$$\begin{aligned}\frac{\partial[\ln detM(\xi)]}{\partial z_1} &= \frac{-2}{z_2 - z_1} + \frac{w'(z_1)}{w(z_1)} \\ &= \frac{w'(z_1)}{w(z_1)(z_2 - z_1)}\left[\frac{-2w(z_1)}{w'(z_1)} + z_2 - z_1\right] \quad if \quad w'(z_1) \neq 0 \\ &= \frac{w'(z_1)}{w(z_1)(z_2 - z_1)}\left[z_2 - h(z_1)\right]\end{aligned}$$

where $\quad h(z_1) = z_1 + \dfrac{2w(z_1)}{w'(z_1)}$. So $\quad \dfrac{\partial[\ln detM(\xi)]}{\partial z_1} \propto w'(z_1)\left[z_2 - h(z_1)\right]$.

Furthermore,

$$\frac{\partial[\ln detM(\xi)]}{\partial z_1} = 0 \iff \frac{w'(z_1)}{w(z_1)(z_2 - z_1)}\left[z_2 - h(z_1)\right] = 0$$

$$i.e. \Rightarrow z_2 = h(z_1) \quad \left[given \quad w'(z_1) \neq 0\right].$$

Note 2.1. *If* $w'(z_1) = 0$, $\frac{\partial[\ln detM(\xi)]}{\partial z_1} = \frac{-2}{z_2 - z_1} \neq 0$. *So,* z_{max} *is not a solution of* $\frac{\partial[\ln detM(\xi)]}{\partial z_1} = 0$.

Similarly,

$$\begin{aligned}
\frac{\partial[\ln detM(\xi)]}{\partial z_2} &= \frac{w'(z_2)}{w(z_2)} + \frac{2}{z_2 - z_1} \\
&= \frac{w'(z_2)}{w(z_2)(z_2 - z_1)}\left[z_2 - z_1 + \frac{2w(z_2)}{w'(z_2)}\right] \quad (if w'(z_2) \neq 0) \\
&= \frac{w'(z_2)}{w(z_2)(z_2 - z_1)}\left[h(z_2) - z_1\right]
\end{aligned}$$

where $h(z_2) = z_2 + \frac{2w(z_2)}{w'(z_2)}$. So $\frac{\partial[\ln detM(\xi)]}{\partial z_2} \propto w'(z_1)\left[h(z_2) - z_1\right]$.

Further,

$$\begin{aligned}
\frac{\partial[\ln detM(\xi)]}{\partial z_2} = 0 \quad &\iff \quad \frac{w'(z_2)}{w(z_2)(z_2 - z_1)}\left[h(z_2) - z_1\right] = 0 \\
i.e. \quad &\Rightarrow \quad z_1 = h(z_2) \quad \left[given \quad w'(z_2) \neq 0\right].
\end{aligned}$$

Note 2.2. *If* $w'(z_2) = 0$, $\frac{\partial[\ln detM(\xi)]}{\partial z_2} = \frac{2}{z_2 - z_1} \neq 0$. *So,* z_{max} *is not a solution of* $\frac{\partial[\ln detM(\xi)]}{\partial z_2} = 0$.

We can be interested in solving one or both of the equations

$$z_1 = h(z_2) \tag{2.6}$$

$$h(z_1) = z_2 \tag{2.7}$$

$$h(z) = z + \frac{2w(z)}{w'(z)} \tag{2.8}$$

As we can see it is useful to study $h(z)$ since the solutions to the above equations clearly depend on the nature of $h(z)$. Plots of $h(z)$ are shown in Figure 2.8 and Figure 2.9 for choices of $w(z)$ which are unimodal and stationary at their maximum, say z_{max}. These reveal examples of $h(z)$ which are increasing both over $z \leq z_{max}$ and over $z \geq z_{max}$ with a vertical asymptote at z_{max} . This proves useful to us.

Let's now consider the single equation in z :

$$h(z) \; = \; c.$$

An implication of the plots is that there is one solution to this equation say $z_L^*(c)$ in the range $z \leq z_{max}$ and one, say $z_U^*(c)$, in the range $z \geq z_{max}$. Moreover since $w'(z_L^*(c)) > 0$ and $w'(z_U^*(c)) < 0$ we have $z_L^*(c) < c < z_U^*(c)$.

In equations (2.6) and (2.7) we have two versions of equation (2.8). Their joint solution with $z_1 < z_2$, must be $z_1^* = a^*$, $z_2^* = b^*$, $a^* < b^*$, a^* and b^* being the support points of the optimal two-point design on $\mathcal{Z}_w$ as defined in the conjectures above.

Note 2.3. *This means that*

$$h(a^*) \; = \; b^*, \qquad h(b^*) \; = \; a^* \qquad \text{and} \quad z_1^* = z_L^*(z_2^*), \quad z_2^* = z_U^*(z_1^*).$$

We can now consider checking these conjectures against the Equivalence Theorem. Consider an arbitrary two point design say

$$\xi \; = \; \binom{z_1 \;\; z_2}{1/2 \; 1/2} \; .$$

The corresponding design matrix is

$$M(\xi) \; = \; 1/2 \begin{pmatrix} [w(z_1) + w(z_2)] & [z_1\, w(z_1) + z_2\, w(z_2)] \\ [z_1\, w(z_1) + z_2\, w(z_2)] & [z_1^2 w(z_1) + z_2^2 w(z_2)] \end{pmatrix},$$

and the determinant of $M(\xi)$ is

$$det(M(\xi)) \; = \; (1/2)^2\, w(z_1) w(z_2) (z_2 - z_1)^2.$$

The inverse of the above design matrix is therefore

$$M^{-1}(\xi) \; = \; \frac{2}{w(z_1)w(z_2)(z_2 - z_1)^2} \begin{pmatrix} [z_1^2 w(z_1) + z_2^2 w(z_2)] & -[z_1\, w(z_1) + z_2\, w(z_2)] \\ -[z_1\, w(z_1) + z_2\, w(z_2)] & [w(z_1) + w(z_2)] \end{pmatrix}.$$

If the above design is to be D-optimal on a set of values of z, say $\mathcal{Z}$, then we must have

$$v(z) \leq 0 \qquad \forall z \in \mathcal{Z}$$

where $v(z) = \frac{1}{2}Q(z) - \frac{1}{w(z)}$. In fact $v(z)$ must be maximised at z_1, z_2 as $v(z_1) = v(z_2) = 0$.

It is of interest to consider the derivative of $v(z)$ at z_1, z_2 . Recall that

$$v'(z) = L(z) - H(z),$$

where $L(z) = \frac{1}{2}Q'(z)$ and $H(z) = \frac{-w'(z)}{[w(z)^2]}$.

$$Q(z) = (1\ z)M^{-1}(\xi)\ \binom{1}{z} .$$

Here

$$Q(z) = \frac{2}{w(z_1)w(z_2)(z_2 - z_1)^2} \left([z_1^2 w(z_1) + z_2^2 w(z_2)] - 2[z_1 w(z_1) + z_2 w(z_2)]z + [w(z_1) + w(z_2)]z^2 \right).$$

And

$$\begin{aligned} L(z) &= \frac{1}{2}\, Q'(z) \\ &= \frac{1}{w(z_1)w(z_2)(z_2 - z_1)^2} \left(-2[z_1 w(z_1) + z_2 w(z_2)] + 2[w(z_1) + w(z_2)]z \right). \end{aligned}$$

Therefore $L(z_1)$ and $L(z_2)$ can be written as follows:

$$\begin{aligned} L(z_1) &= \frac{1}{w(z_1)w(z_2)(z_2 - z_1)^2}(2z_1[w(z_1) + w(z_2)] - 2[z_1 w(z_1) + z_2 w(z_2)]) \\ &= \frac{2}{w(z_1)w(z_2)(z_2 - z_1)^2}[(z_1 - z_2)w(z_2)] \\ &= \frac{-2}{w(z_1)(z_2 - z_1)} \\ L(z_2) &= \frac{1}{w(z_1)w(z_2)(z_2 - z_1)^2}(2z_2[w(z_1) + w(z_2)] - 2[z_1 w(z_1) + z_2 w(z_2)]) \\ &= \frac{2}{w(z_2)(z_2 - z_1)} \end{aligned}$$

Now with

$$\begin{aligned} v'(z) &= L(z) - H(z) \\ &= L(z) + \frac{w'(z)}{[w(z)^2]}, \end{aligned}$$

we have

$$\begin{aligned} v'(z_1) &= L(z_1) - H(z_1) \\ &= L(z_1) + \frac{w'(z_1)}{[w(z_1)^2]} \\ &= \frac{-2}{w(z_1)(z_2 - z_1)} + \frac{w'(z_1)}{[w(z_1)^2]} \\ &= \frac{1}{w(z_1)}\left[\frac{-2}{(z_2 - z_1)} + \frac{w'(z_1)}{w(z_1)}\right] \\ &= \frac{w'(z_1)}{[w(z_1)]^2(z_2 - z_1)}\left[\frac{-2w(z_1)}{w'(z_1)} + (z_2 - z_1)\right] \\ &= \frac{w'(z_1)}{[w(z_1)]^2(z_2 - z_1)}\left[z_2 - [z_1 + \frac{2w(z_1)}{w'(z_1)}\right] \\ &= \frac{w'(z_1)}{[w(z_1)]^2(z_2 - z_1)}\left[z_2 - h(z_1)\right]. \end{aligned}$$

Similarly,

$$\begin{aligned} v'(z_2) &= L(z_2) - H(z_2) \\ &= L(z_2) + \frac{w'(z_2)}{[w(z_2)^2]} \\ &= \frac{2}{w(z_2)(z_2 - z_1)} + \frac{w'(z_2)}{[w(z_2)^2]} \\ &= \frac{w'(z_2)}{[w(z_2)]^2(z_2 - z_1)}\left[\frac{2w(z_2)}{w'(z_2)} + (z_2 - z_1)\right] \\ &= \frac{w'(z_2)}{[w(z_2)]^2(z_2 - z_1)}\left[h(z_2) - z_1\right]. \end{aligned}$$

So,

$$\begin{aligned} v'(z_1) &\propto w'(z_1)[z_2 - h(z_1)] \\ v'(z_2) &\propto w'(z_2)[h(z_2) - z_1]. \end{aligned}$$

2.2.3 Some properties of $v(z)$

We also enumerate some properties of $v(z)$ on the assumption that for any design $v(z)$ is continous and has two maximal and one minimal TP over $\mathcal{Z}_w$. This will be the case if $H(z)$ has the properties mentioned above as demonstrated for several examples.

Denote the TP's by TP_L, TP_M, TP_U representing the Lower, the Middle and the Upper Turning Points respectively so that $TP_L < TP_M < TP_U$ and
$TP_L < z_{max} < TP_U$.

Some simple properties of $v(z)$ will therefore be:

(i) $v(z_1) = v(z_2) = 0$

(ii) It is possible that $v(TP_L) = v(TP_U)$ but $v(TP_M) \neq v(TP_L)$, $v(TP_U)$.

(iii) $v'(z) > 0$ for $z < TP_L$ and $TP_M < z < TP_U$
$v'(z) < 0$ for $z > TP_U$ and $TP_L < z < TP_M$

(iv) If $v(z^\times) = v(TP_U)$ then either $z^\times < TP_L$ with $v'(z^\times) > 0$ or $TP_L < z^\times < TP_M$ with $v'(z^\times) < 0$; in the last case $v(z) \leq v(TP_U)$ over $z \geq z^\times$.

(v) If $v(z^\times) = v(TP_L)$ then either $z^\times > TP_U$ with $v'(z^\times) < 0$ or $TP_M < z^\times < TP_U$ with $v'(z^\times) > 0$; in the last case $v(z) \leq v(TP_L)$ over $z \leq z^\times$.

(vi) If $z_1 \geq TP_L$ and $v'(z_2) > 0$ then $v'(z_1) < 0$ i.e. $TP_L < z_1 < TP_M$;
If $z_2 \leq TP_U$ and $v'(z_1) < 0$ then $v'(z_2) > 0$ i.e. $TP_M < z_2 < TP_U$.

(vii) Suppose $v'(z_1) \leq 0$ and $v'(z_2) \geq 0$. Then $TP_L \leq z_1 < TP_M < z_2 \leq TP_U$ and $v(z) < 0$, $\forall\ z \in [z_1, z_2]$.

(viii) Suppose $v'(z^\times) < 0$ and $v(z^\times) = v(TP)$ where $v'(TP) = 0$.
Then $TP = TP_U$.

Suppose $v'(z^{\times}) > 0$ and $v(z^{\times}) = v(TP)$ where $v'(TP) = 0$.

Then $TP = TP_L$.

2.2.4 Confirmation of D-Optimality

We now consider taking z_1, z_2 to be the support points of the conjectured optimal designs of the various cases of $\mathcal{Z} = [a, b]$ above. Our primary objective is to establish that $v(z) \leq 0$ on $\mathcal{Z}$. The above properties confirm that this will be true if $v'(z_1) \leq 0$ and $v'(z_2) \geq 0$.

First we establish a preliminary result.

Theorem 2.1. *There can only be one solution satisfying* $z_1 < z_2$ *to*

$$\begin{aligned} h(z_1) &= z_2 \\ h(z_2) &= z_1 \end{aligned}$$

Hence there can only be one solution to the equations

$$\begin{aligned} v'(z_i) &= 0 \qquad i = 1, 2 \\ \text{and also} \qquad \frac{\partial[\ln det M(\xi)]}{\partial z_i} &= 0 \qquad i = 1, 2. \end{aligned}$$

Proof. Suppose there are more than two pairs of solutions (z_1, z_2). Then the $v-$function for each solution satisfies

$$\begin{aligned} v(z_1) &= v(z_2) = 0 \\ v'(z_1) &= v'(z_2) = 0. \end{aligned}$$

Hence z_1, z_2 are TP's of $v(\cdot)$ with a comman value of zero. Since $v(z)$ only has 3 TP's these must be two maximal ones. So, $z_1 = TP_L$, $z_2 = TP_U$ and $v(z) \leq 0 \; \forall \; z$. Hence the design $\binom{z_1 \;\; z_2}{1/2 \; 1/2}$ is D-optimal and so then is any convex combination of them. Moreover they share a common design matrix and hence a common

$v(z)$ which therefore must be zero at all support points of these designs. Given $v(z) \leq 0$ they are all Maximal TP's. This conflicts with the assumption that $v(z)$ has only 2 max and 1 min TP. □

We now establish results confirming that $v'(z_1) \leq 0$, $v'(z_2) \geq 0$ as required. Crucially we assume that $h(z)$ is increasing in z over $z \leq z_{max}$ and over $z \geq z_{max}$ where z_{max} is the point where the weight function $w(z)$ reaches a maximum.
(Note: Equal weights guarantee $v(z_1) = v(z_2) = 0$.)

1-

- $z_1 = a > z_{max}$, $z_2 = b > a$. We show that $v'(a) < 0$.

$$v'(a) = \frac{w'(a)}{[w(a)]^2(b-a)}[b - h(a)]$$

Now since $a > z_{max}$, $w'(a) < 0$. So $v'(a) < 0$ is true if $[b - h(a)] > 0$.

$$b - h(a) = (b-a) - \frac{2w(a)}{w'(a)}$$

The right side of equation is always positive, because $a < b$ and $w'(a) < 0$. Therefore $v'(z_1) < 0$.

- $z_2 = b < z_{max}$, $z_1 = a < b$. We show $v'(b) > 0$.

$$v'(b) = \frac{w'(b)}{[w(b)]^2(b-a)}[h(b) - a]$$

Now $w'(b) > 0$. So $v'(b) > 0$ is true if $[h(b) - a] > 0$.

$$h(b) - a = (b-a) + \frac{2w(b)}{w'(b)}$$

$[h(b) - a]$ is always positive,because $a < b$ and $w'(b) > 0$. Therefore $v'(z_2) > 0$.

2- $a^* < a < z_{max} < b < b^*$ $z_1 = a$, $z_2 = b$

Because $a < z_{max}$, and $b > z_{max}$, $w'(a) > 0$ and $w'(b) < 0$.

- Since $h(z)$ is increasing over $(-\infty, z_{max}]$, $h(a) > h(a^*)$ and since $b < b^*$ then $[b - h(a)] < [b^* - h(a^*)] = 0$. Therefore $v'(z_1) < 0$.

- Since $h(z)$ is increasing over $[z_{max}, \infty)$, $h(b) < h(b^*)$ and then $[h(b) - a] < [h(b^*) - a^*] = 0$. Therefore $v'(z_2) > 0$.

  Hence the two-point design

$$\xi = \binom{a \quad b}{1/2\ 1/2}$$

  is D-optimal for all $\mathcal{Z} = [a, b]$ where $a^* < a < z_{max} < b < b^*$.

3- $z_1 = a^*$, $z_2 = b^*$ $\quad a^* < b^*$

$$b^* = h(a^*) \tag{2.9}$$

$$h(b^*) = a^* \tag{2.10}$$

From Theorem (2.1), a^*, b^* is the only possible solution ($a^* < b^*$) to equations (2.9) and (2.10). So $v'(a^*) = 0$ $v'(b^*) = 0$. Thus a^*, b^* identify 2 max TP's of $v(z)$. Moreover they are TP's at which $v(z)$ has a common value of zero since $v(z_1) = v(z_2) = 0$. From property (ii) of $v(z)$ the only possibility is that they are the two maximal TP's of $v(z)$. i.e. $z_1 = a^* = TP_L$, $z_2 = b^* = TP_U$. Hence the Equivalence Theorem is satisfied. Moreover since $v(z) \leq 0 \ \forall \ z \in \mathcal{Z}_w$ then $v(z) \leq 0 \ \forall \ z \in \mathcal{Z} = [a, b]$, where $[a, b] \subseteq [a^*, b^*]$. Hence the two-point design

$$\xi = \binom{a^* \quad b^*}{1/2\ 1/2} .$$

is D-optimal for all $\mathcal{Z} = [a, b]$ where $a \leq a^*$ and $b \geq b^*$.

4- $a < a^*$, $b > b^*$ $z_1 = a^*$, $z_2 = b^*$

Same design as for 3.

5- $a^* \leq a \leq z_{max} \leq b = b^*(a)$ $\quad (b^*(a) > b^*)$ $\quad z_1 = a$, $z_2 = b^*(a)$

Clearly $v'(z_2) = 0$. First $w'(a) > 0$. We want $v'(z_1) = v'(a)$ negative. So we need to investigate the derivative of $v'(a)$. Here

$$\begin{aligned} v'(a) &= \frac{w'(a)}{[w(a)]^2(b^*(a)-a)}[b^*(a)-h(a)] \\ &= \frac{w'(a)}{[w(a)]^2(b^*(a)-a)}[h_R^{-1}(a)-h_L(a)] \end{aligned}$$

where $h_R^{-1}(z)$ is the inverse function of $h(z)$ for $z \geq z_{max}$ and $h_L(z) = h(z)$ for $z \leq z_{max}$. Consider $d_{RL}(a) = h_R^{-1}(a) - h_L(a)$, and recall that $a < z_{max}$. We know from Theorem (1.1), $a = a^*$ is the only solution to $d_{RL} = 0$. So the functions $h_R^{-1}(a)$ and $h_L(a)$ cross only once at a^* . Further since $h_L(z_{max}) = \infty$ while $h_R^{-1}(z_{max}) < \infty$. Then $h_R^{-1}(a) < h_L(a)$ at $a = z_{max}$. Hence this inequality is true for $a^* < a < z_{max}$. Therefore $v'(z_1) = v'(a) < 0$. Thus the two point design

$$\xi = \begin{pmatrix} a & b^*(a) \\ 1/2 & 1/2 \end{pmatrix}$$

is D-optimal for all $\mathcal{Z} = [a,b]$, where $a^* < a < z_{max}$ and $b = b^*(a)$. Since $v'(z_2) = 0$ for $z_2 = b^*(a)$. This design is also D-optimal for $b > b^*(a)$.

6- $a^* \leq a \leq z_{max}, \quad b^* < b < b^*(a) \quad z_1 = a, \quad z_2 = b$

First $w'(a) > 0$. Secondly $[b-h(a)] < b^*(a)-h(a) = h_R^{-1}(a)-h_L(a) < 0$ by above. So $v'(a) < 0$. Also we need to show $v'(b) > 0$. Because of $b \geq z_{max}$ $w'(b) < 0$. We assumed $b < b^*(a)$. If $h(\cdot)$ is an increasing function, $h(b) < h(b^*(a)) = a$. Hence $h(b) - a < [h(b^*(a)) - a] = 0$. Thus $v'(z_2) = v'(b) > 0$ and the two point design

$$\xi = \begin{pmatrix} a & b \\ 1/2 & 1/2 \end{pmatrix}$$

is D-optimal for all $\mathcal{Z} = [a,b]$ where $a^* \leq a \leq z_{max}$ and $b < b^*(a)$.

7- $a = a^*(b)\ (\leq a^*) \leq z_{max} \leq b \leq b^* \quad z_1 = a^*(b),\ z_2 = b$

(This is the complementary case of 5.)

Now $v'(a) = 0$. $w'(b) < 0$ and $v'(b) > 0$ is true if

$h_L^{-1}(b) > h_R(b) \quad \forall b \in [z_{max}, b^*]$. So we need to investigate the derivative of $v'(b)$.

$$\begin{aligned} v'(b) &= \frac{w'(b)}{[w(b)]^2(b - a^*(b))}[h(b) - a^*(b)] \\ &= \frac{w'(b)}{[w(b)]^2(b - a^*(b))}[h_R(b) - h_L^{-1}(b)] \end{aligned}$$

where $h_L^{-1}(b)$ is the inverse function of $h(z)$ for $z \leq z_{max}$ and $h_R(z) = h(z)$ for $z \geq z_{max}$. Consider $d_{LR}(b) = h_R(b) - h_L^{-1}(b)$ and recall $b > z_{max}$. We know from item 3 that $b = b^*$ is only solution to $d_{LR}(b) = 0$. So the functions $h_R(b)$ and $h_L^{-1}(b)$ cross only once at b^* and since $h_R(z_{max}) = -\infty$ while $h_L(z_{max}) > -\infty$. Then $h_L^{-1}(b) > h_R(b)$ or $h_R < h_L^{-1}(b)$ at $\forall b \in [z_{max}, b^*]$. Therefore this inequality is true for $z_{max} < b < b^*$. And therefore $v'(z_2) = v'(b) > 0$. Thus the two point design

$$\xi = \begin{pmatrix} a^*(b) & b \\ 1/2 & 1/2 \end{pmatrix}$$

is D-optimal for all $\mathcal{Z} = [a, b]$ where $a = a^*(b)$ and $z_{max} < b < b^*$. Since $v'(z_1) = 0$ for $z_1 = a^*(b)$. This is also D-optimal for $a > a^*(b)$.

8- $a^*(b) < a < z_{max} \leq b \leq b^* \quad z_1 = a \qquad z_2 = b$

This is the complementary of case 6.

Now $v'(a) = v'(b) = 0$. First consider $w'(b) < 0$, $v'(b)$ positive. So we need to investigate the derivative $v'(b)$:

$$v'(b) = \frac{w'(b)}{[w(b)]^2(b - a)}[h(b) - a]$$

where $h(b) - a < h(b) - a^*(b) < 0$ because of $b \geq z_{max}$. Therefore $v'(b) > 0$. Secondly $w'(a) > 0$ and $v'(a)$ negative.

$$v'(a) = \frac{w'(a)}{[w(b)]^2(b - a)^2}[b - h(a)]$$

where

$$h(a) > h(a^*(b)) = b$$
$$\Rightarrow -h(a) < -h(a^*(b))$$
$$\Rightarrow b - h(a) < b - h(a^*(b)) = b - b = 0$$

Therefore $v'(a) < 0$. Thus the two point design

$$\xi = \binom{a \quad b}{1/2 \; 1/2}$$

is D-optimal.

2.2.5 Some Conclusions

These results confirm that :

$$\begin{aligned}
Supp(\xi^*) &= \{a^*, b^*\} \qquad a < a^*, b > b^* \\
Supp(\xi^*) &= \{max\{a,\ a^*(b)\}, b\} \qquad a < a^*, b < b^* \\
Supp(\xi^*) &= \{a, min\{b,\ b^*(a)\}\} \qquad a > a^*, b > b^* \\
Supp(\xi^*) &= \{a, b\} \qquad a > a^*, b < b^*
\end{aligned}$$

So the equivalence theorem is satisfied by our conjectured optimal designs for all possible design intervals $[a, b]$ if the function

$$h(z) \quad = \quad z + \frac{2w(z)}{w'(z)}$$

is increasing over $z \leq z_{max}$ and over $z \geq z_{max}$. We have noted that this appears to be true for a range of $w(z)$. Plots of $h(z)$ functions are given in Figure 2.8 and Figure 2.9. Interestingly these properties also guarantee that $G(\mathcal{Z}_w)$ is a closed convex set, as Wu (1988) reports. He established them analytically for a number of our binary regression weight functions. In some cases he established the stronger result that the ratio $w(z)/w'(z)$ is increasing over $z \leq z_{max}$ and over $z \geq z_{max}$ (for the logistic, complementary log-log and skewed logistic binary weight functions). This implies that $w(z)$ is log-concave. For the other cases he proved analytically that $h(z)$ is an increasing function (Probit, Double exponential, Double reciprocal). **We report and extend these results.** Plots of $w(z)/w'(z)$ are given in Figure (2.10) and Figure (2.11). We summarize some aspects of the functions $h(z)$ and $w(z)/w'(z)$ for some binary regression weight functions and some non-binary weight functions in Table (2.6) and Table (2.7). One of the most obvious remarks is that the functions $h(z)$ and $w(z)/w'(z)$ have **almost the same shape** and are **increasing** over $(-\infty, z_{max})$ and (z_{max}, ∞) with a vertical asymptote at z_{max} for the weight functions listed. This confirms and extends Wu's (1988) findings for some of them. Thus we have established the final condition needed to satisfy the Equivalence Theorem.

Recently, Sebastiani and Settimi (1997) established the truth of the Ford *et al.* (1992) conjecture for Logistic Regression using exactly our approach. Independently of Wu they established the necessary property of $h(z)$.

In effect we have established the following theorem:

Theorem 2.2. *Assume that $w(z)$ is continuous, differentiable and unimodal and suppose $Z_w = [A, B]$. Let $w(z)$ be a weight function and*

$$
\begin{aligned}
H(z) &= \frac{-w'(z)}{[w(z)]^2} \\
h(z) &= z + \frac{2w(z)}{w'(z)} \\
v(z) &= w(z)(1\, z)M^{-1}(\xi)(1\, z)^T.
\end{aligned}
$$

If $H(z)$ is continous with $H(A) = -\infty$, $H(B) = \infty$, differentiable with $H'(A) = H'(B) = \infty$, AND is first concave increasing then convex increasing the function $v(z)$ [or the variance function] can have at most 3 TP's two maximal ones & one minimal one. In consequence a D-optimal design on any $\mathcal{Z} = [a, b]$ has 2 support points.

Further if $h(z)$ is increasing over $z \leq z_{max}$ and over $z \geq z_{max}$ these support points are as follows :

i - $\mathcal{Z} = \mathcal{Z}_w$ $Supp\{\xi\} = \{a^*, b^*\}$

a^* b^* maximise $det(M(\xi)) \propto (b - a)^2 w(a)w(b)$ and $b^* = h(a^*)$, $a^* = h(b^*)$.

ii - $\mathcal{Z} = (A, b)$ and $b \leq b^*$ $Supp\{\xi\} = \{max\{a, a^*(b)\}, b\}$

$a^*(b)$ solves $h(a) = b$.

iii - $\mathcal{Z} = (a, B)$ and $a \geq a^*$ $Supp\{\xi\} = \{a, min\{b, b^*(a)\}\}$

$b^*(a)$ solves $a = h(b)$.

iv - $\mathcal{Z} = [a, b]$ and $a \geq a^*$ $b \leq b^*$ $Supp\{\xi\} = \{a, b\}$.

This result provides confirmation for the Ford *et al.* (1992) conjecture.

Before closing this chapter we also consider the weight functions in Group II and IV.

Double Reciprocal & Double Exponential Binary Weight Functions

Our objective is to find D-optimal designs for possible interval subsets of $\mathcal{Z}_w = \mathcal{R}$ for two symmetric binary weight functions : the Double Reciprocal & Double Exponential weight functions which are presented in Table 2.8 .

- **Case 1 :** $\mathcal{Z} = \mathcal{Z}_w = \mathcal{R}$

  The first striking remark is that the Double Reciprocal and Double Exponential weight functions are unimodal and both functions reach their maximum value at $z = z_{max} = 0$ at which point both are non-differentiable. So these are not stationary values [Figure 2.12]. For these two weight functions, the induced Design Space $G(\mathcal{Z})$ is again a closed convex curve in $\mathcal{R}^2$ for the $\mathcal{Z} = \mathcal{Z}_w = (-\infty, \infty)$ [Figure 2.13] . However it has a sharp vertex at $z = 0$. For these cases it seems likely that the minimal central ellipsoid containing $G(\mathcal{Z}_w)$ will touch more than twice: at $z = 0$ and also, given the symmetry of $w(z)$ about zero, at two other points symmetric about zero, in which case the D-optimal design has three support points. This impression is confirmed by the plots of $H(z)$ for these cases. $H(z)$ is discontinuous at $z = 0$. An upward sloping line can cross $H(z)$ four times. We discuss $H(z)$ in more detail below. The distribution of the weights must be symmetric too, that is the support is of the form $\{-z^*, 0, z^*\}$ with optimal weights $(\hat{p}, 1 - 2\hat{p}, \hat{p})$ Musrati (1992). z^* and $\hat{p}$ maximize the determinant of the information matrix $\mathbf{det}(M(\xi))$ with respect to ξ where ξ is the design

t	$-z$	0	z
$\xi(t)$	p	1 - 2p	p .

When there are three support points there is an explicit solution for the optimal weights as first reported by Torsney and Musrati (1993). If these points are z_1, z_2, z_3, the respective optimal weights p_1, p_2, p_3 are given by

$$p_i \; = \; D_i/(D_1 + D_2 + D_3), \qquad i \; = \; 1, 2, 3 \qquad (2.11)$$

where

$$D_i \; = \; D_{jk}/(D_{ij} + D_{ik} + D_{jk}), (i,j,k) = (1,2,3), (2,1,3)(3,1,2)\,,$$
$$D_{ij} \; = \; w(z_i)w(z_j)(z_i \; - \; z_j)^2, \qquad (i,j) \; = \; (1,2), (1,3), (2,3)\,.$$

The support points and the optimal weights for both models are as follows (Torsney,Musrati(1993)):

Name	Support Points	Optimal Weights
Double Reciprocal	$-\sqrt{2}, 0, \sqrt{2}$	$0.2617, 0.4766, 02617$
Double Exponential	$-1.5936, 0, 1.5936$	$0.2819, 0.4362, 0.2819$

Now we consider the variance function under these optimal designs. The implication is that its maximum occurs at three local maxima (at $\pm z^*$ and 0) while it has two local minima (at $\pm z^{\oplus}$ and $z^{\oplus} \; < \; z^*$, for some $z^{\oplus}$). All are stationary values except the local maximum at zero. This is indeed the case so that the necessary and sufficent condition of the equivalence theorem is satisfied. These designs are globally D-optimal.

- **Case 2 :** $\mathcal{Z} = [a, b] = \mathcal{Z} \subseteq [0, \infty)$ **or** $\mathcal{Z} \subseteq (-\infty, 0]$

  For these cases results similar to the other binary weight functions hold. Namely ξ^* has two support points with equal weights;

$$if \qquad a \geq 0 \qquad Supp(\xi^*) \; = \; \{a, \, min\{b^*(a), b\}\}$$
$$if \qquad b \leq 0 \qquad Supp(\xi^*) \; = \{max\{a, a^*(b)\}, b\}\,.$$

Justification of the Conjecture

1- Let's first consider the function $H(z)$. We have $H(z) = 6sz^2 + 10|z| + 4s$ and $H(z) = 2se^{|z|}$ for the Double Reciprocal and the Double exponential weight functions respectively where s is sign of z. [see Table 2.9 and Figure 2.14]. In both cases $H(z)$ is positive for positive z, negative for negative z and discontinous at $z = 0$. However we only need to consider its behaviour in $[0, \infty)$ and $(-\infty, 0]$ separately. As Table 2.9 shows, $H'(z) = 12|z| + 10$, $H''(z) = 12s$ for the Double Reciprocal and $H'(z) = 2e^{|z|}$ $H''(z) = 2se^{|z|}$ for the Double Exponential weight functions . In both cases, it is clear that $H'(z)$ is positive for all z and $H''(z)$ is negative for negative z and positive for positive z. Hence $H(z)$ is concave increasing from $-\infty$ over $(-\infty, 0]$ and $H(z)$ convex increasing over $[0, \infty)$ to $+\infty$ [see plots in Figure 2.14]. An upward sloping line with a negative intercept must cross $H(z)$ twice in $[0, \infty)$ while one with a positive intercept must cross $H(z)$ twice in $(-\infty, 0]$. Note that under a design on a subset of $[0, \infty)$ the line $L(z)$ has a negative intercept since $E(Zw(Z)) > 0$ and vice versa for a design with a support on $(-\infty, 0]$. Thus $v(z)$ can only have two TP's in these intervals. In the case of $z > 0$ an upward sloping line with a negative intercept crossing twice must cross from below then from above. So the first TP is a minimum, the second a maximum. The converse holds for $z < 0$. So $v(z)$ has only one minimum TP and at most one maximum TP in $[a, b]$. As a result of this there are two support points in which case the optimal weights are $\frac{1}{2}, \frac{1}{2}$.

2 - As before the function of $h(z)$ is increasing in both regions. As Table (2.10) shows $h'(z) = 1 - \frac{[2|z|(3|z|+4)+3]}{[3|z|(3|z|+4)+4]}$ for the double reciprocal weight function and $h'(z) = 1 - e^{-|z|}$ for the double exponential weight function (Wu 1988). It is clear that the ratio $\frac{[2|z|(3|z|+4)+3]}{[3|z|(3|z|+4)+4]} < 1$, and that $e^{-|z|} < 1$. This shows that the derivative $h'(z)$ is positive for both weight functions. Interestingly $w(z)/w'(z)$ is not increasing. Plots of $w(z)/w'(z)$ and of $h(z)$

are given in Figures (2.15) and (2.16), respectively. These characteristics are summarized in Table (2.10). Since $h(z)$ is increasing the support points are as given above.

- **Case 3 :** $\mathcal{Z} = [a, b]$ $(a < 0,\ b > 0)$

  Torsney & Musrati (1993) and Musrati(1992) showed that in this case sometimes there are two support points with optimal equal weights $(\frac{1}{2}, \frac{1}{2})$, sometimes three support points (including $z = 0$) with optimal weights given by equation (1.9); and in each of these cases sometimes neither, or one or both endpoints a, b are support points.

Group IV : Two Non-Binary Weight Functions

Now we are going to apply our theorem to the weight functions listed in Table (2.11), namely $w(z) = e^z$ and $w(z) = z^t$. Ford *et al.* (1992) derived the D-optimal designs for all $\mathcal{Z} = [a, b] \subseteq \mathcal{Z}_w$ for the above functions. However it is of interest to see that the above approach also works here.

If we look at plots of these weight functions, they have a shape that is different from the shape of the previously studied weight functions [Figure 2.17]. Also $G(\mathcal{Z})$ is no longer bounded for all $\mathcal{Z}$, [Figure 2.18]. Moreover, unlike before, there now exists a one to one relationship between the components of $\underline{g}(z)$, namely $g_1 = \sqrt{w(z)}$ and $g_2 = z\sqrt{w(z)}$, which can be derived explicitly. These are

$$\begin{aligned} g_2 &= 2g_1 \ln g_1 \qquad for \qquad w(z) = e^z \\ g_2 &= g_1^{\frac{(2+t)}{t}} \qquad for \qquad w(z) = z^t. \end{aligned}$$

We consider these weight functions in turn:

a - $w(z) = e^z$

From the definition of $\underline{g}(z)$, we have

$$g_1 = \sqrt{w(z)} = e^{z/2} \qquad \Rightarrow z = 2\ln g_1$$

Hence,

$$\begin{aligned} g_2 &= z\sqrt{w(z)} \\ &= 2g_1 \ln g_1. \end{aligned}$$

If we compute the first and second derivatives of g_2 with respect to g_1, we get the following:

$$\begin{aligned} \frac{\partial g_2}{\partial g_1} &= 2(\ln g_1 + 1) \\ \frac{\partial^2 g_2}{\partial g_1^2} &= \frac{2}{g_1}. \end{aligned}$$

Since $g_1 > 0$, the second derivative $\dfrac{\partial^2 g_2}{\partial g_1^2} > 0$ which establishes the convexity of g_2 as a function of g_1.

b - $w(z) = z^t$

From the definition of $\underline{g}(z)$ once again, we can write the following:

$$g_1 = z^{t/2} \qquad \Rightarrow z = g_1^{2/t}$$

Hence,

$$g_2 = zg_1 = g_1^{2/t} g_1 = g_1^{(2+t)/t}.$$

Further

$$\frac{\partial g_2}{\partial g_1} = (\frac{2+t}{t}) g_1^{\frac{2+t}{t}-1} = (\frac{2+t}{t}) g_1^{2/t}$$

Now $g_1^{2/t} > 0$, which means that the sign of $\dfrac{\partial g_2}{\partial g_1}$ depends only on the sign of $\frac{2+t}{t}$.

Further

$$\frac{\partial^2 g_2}{\partial g_1^2} = \frac{2(2+t)}{t^2} g_1^{\frac{(2-t)}{t}}, \qquad \frac{2}{t^2} g_1^{\frac{(2-t)}{t}} > 0.$$

Hence $sign(\dfrac{\partial^2 g_2}{\partial g_1^2}) = sign(2+t)$.

We now consider three distinct ranges of values of t.

- If $t < 0$ then $2+t < 0$ $\frac{\partial g_2}{\partial g_1} > 0$ and $\frac{\partial^2 g_2}{\partial g_1^2} < 0$. The weight function $w(z) = z^t$ is concave increasing in the interval $t < -2$.
- If $-2 < t < 0$, $\frac{\partial g_2}{\partial g_1} < 0$ and $\frac{\partial^2 g_2}{\partial g_1^2} > 0$. The weight function $w(z) = z^t$ is therefore convex decreasing in the interval $-2 < t < 0$.
- If $t > 0$ $\frac{\partial g_2}{\partial g_1} > 0$ and $\frac{\partial^2 g_2}{\partial g_1^2} > 0$. The weight function $w(z) = z^t$ is convex increasing in the interval $t > 0$.

Therefore

- $G(\mathcal{Z})$ is convex increasing for $w(z) = \exp(z)$ $\mathcal{Z} \subset \mathcal{R}$
- $G(\mathcal{Z})$ is convex increasing for $w(z) = z^t$, $t > 0$ $\mathcal{Z} \subset \mathcal{R}^+$
- $G(\mathcal{Z})$ is concave increasing for $w(z) = z^t$, $t < 0$ $\mathcal{Z} \subset \mathcal{R}^+$
- $G(\mathcal{Z})$ is convex decreasing for $w(z) = z^t$, $-2 < t < 0$ $\mathcal{Z} \subset \mathcal{R}^+$

The boundedness of $G(\mathcal{Z})$ requires the following conditions:

w(z)	$\mathcal{Z} = [a, b]$
z^t, $-2 \geq t \geq 0$	$a > 0$, $b < \infty$
z^t, $t < -2$	$a > 0$, $b \leq \infty$
z^t, $t > 0$	$a \geq 0$, $b < \infty$
$\exp(z)$	$a \geq -\infty$, $b < \infty$

We now show that D-optimal designs on any $\mathcal{Z} = [a, b]$ which guarantees that $G(\mathcal{Z})$ is bounded have similar structure to those of our non-binary weight functions.

We consider again the function $H(z)$, [Figure 2.19].

- For $w(z) = \exp(z)$, $H(z) = H''(z) = -e^{-z} < 0$, $H'(z) = e^{-z}$ (see Table (2.12)). This means that $H(z)$ is concave increasing from $-\infty$ up to 0 with an infinite derivative at $z = -\infty$.

- For $w(z) = z^t$, $H(z) = -tz^{-t-1}$, $H'(z) = t(t+1)z^{-(t+2)}$, and $H''(z) = -t(t+1)(t+2)z^{-(t+3)}$. So for $t > 0$ $H(z)$ is concave increasing from $-\infty$ to zero with an infinite derivative at $z = 0$ (see Table (2.12)) while for $t < -2$, $H(z)$ is convex increasing from zero to ∞ with a zero derivative at $z = 0$.

- For $w(z) = z^t$, $-2 < t < 0$, $H(z)$ is convex decreasing.

  For these weight functions, $H(z)$, $H'(z)$ and $H''(z)$ functions are presented in Table (2.12).

We can then argue that an upward sloping line can cross $H(z)$ at most twice in the cases $w(z) = e^z$, z^t, $t > 0$; z^t, $t < -2$. Hence $v(z)$ has at most two TP's, one a maximal TP, one a minimal TP. Thus there can be only two support points on any $\mathcal{Z} = [a, b]$.

In the case $w(z) = z^t$, $-2 < t < 0$ $(a > 0, b < \infty)$ an upward sloping line crosses $H(z)$ only once from below. So $v(z)$ has one TP, a minimal TP. The implication is again that there can be only two support points. These must be the endpoints a and b. In fact the plot of $G(\mathcal{Z})$ shows a convex decreasing curve. The minimal central ellipse containing $G(\mathcal{Z})$ can only touch it at its endpoints.

Determination of support points

We consider the functions $h(z)$ and $w(z)/w'(z)$.

- For $w(z) = \exp(z)$, $h(z) = z + 2$ which means that $h(z)$ is an increasing linear function from $-\infty$ to ∞.

- For $w(z) = z^t$, $h(z) = z + \frac{2z}{t} = (\frac{t+2}{t})z$, which is linear increasing if $t > 0$ or $t < -2$. If $-2 < t < 0$, $h(z)$ is decreasing.

$w(z)/w'(z)$ and $h(z)$ are plotted in Figure (2.20) and Figure (2.21), respectively and they are summarized in Table (2.13).

Implications for support points of the D-optimal design on $\mathcal{Z} = [a, b]$ are

$w(z)$	$\mathcal{Z} = [a, b]$	$Supp(\xi^*)$
z^t, $-2 \leq t \leq 0$	$a > 0, b < \infty$	$\{a, b\}$
z^t, $t < -2$	$a > 0, b \leq \infty$	$\{a, min\{b, b^*(a)\}\}$
z^t, $t > 0$	$a \geq 0, b < \infty$	$\{\{max\{a, a^*(b)\}, b\}$
$\exp(z)$	$a \geq -\infty, b < \infty$	$\{max\{a, a^*(b)\}, b\}$

In fact since $h(z)$ is linear in z there are explicit solutions for $a^*(b)$, $b^*(a)$, i.e. for the solutions to the equations

$$\begin{aligned} h(a) &= b \\ a &= h(b) \,. \end{aligned}$$

These are $a^*(b) = b - 2$ for $w(z) = \exp(z)$ while for $w(z) = z^t$, $a^*(b) = tb/(t+2)$ (if $t > 0$) and $b^*(a) = ta/(t+2)$ (if $t < -2$). These values are reported by Ford *et al.* (1992).

1 - Binary Weight Functions : Group I $w(z) = f^2(z)/F(z)[1-F(z)]$ $(F(\cdot) = c.d.f., \ f(\cdot) = F'(\cdot))$				
Name	$f(z)$	$F(z)$	$w(\cdot)$	$Z_w{}^a = [a,b]$
LOGISTIC	$\frac{e^{-z}}{(1+e^{-z})^2}$	$\frac{1}{1+e^{-z}}$	$\frac{e^z}{(1+e^z)^2}$	$-\infty \ \infty$
SKEWED LOGISTIC	$m[F_1(z)]^{m-1}f_1(z)$	$(1+e^{-z})^{-m}$	$\left[\frac{me^{-z}}{1+e^{-z}}[(1+e^{-z})^m - 1]^{-\frac{1}{2}}\right]^2$	$-\infty \ \infty$
GENERALIZED	$(e^z)(\lambda e^z+1)^{-\frac{1}{\lambda}-1}$	$1-[\lambda e^z+1]^{\frac{-1}{\lambda}}$	$\frac{e^{2z}}{(\lambda e^z+1)^2[(\lambda e^z+1)^{\frac{1}{\lambda}}-1]}$	$-\infty \ \infty$
COMPLEMENTARY LOG-LOG	$e^{(z-e^z)}$	$1-e^{-e^z}$	$e^{2z}(e^{e^z}-1)^{-1}$	$-\infty \ \infty$
PROBIT	$\frac{1}{\sqrt{2\pi}}e^{\frac{-z^2}{2}}$	$\int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}}\, e^{\frac{-t^2}{2}}\, dt$	$w(\cdot)$	$-\infty \ \infty$
2 - Binary Weight Functions : Group II $w(z) = f^2(z)/F(z)[1-F(z)]$ $(F(\cdot) = c.d.f., \ f(\cdot) = F'(\cdot), \ s\, is\, sign\, of\, z)$				
DOUBLE RECIPROCAL[b]	$\frac{1}{2}(1+\lvert z\rvert)^{-2}$	$\frac{(1+s)}{2} - \frac{s}{2}(1+\lvert z\rvert)^{-1}$	$\frac{(1+\lvert z\rvert)^{-2}}{2\lvert z\rvert+1}$	$-\infty \ \infty$
DOUBLE EXPONENTIAL[b]	$\frac{1}{2}e^{(-\lvert z\rvert)}$	$\frac{1+s}{2} - \frac{s}{2}e^{(-\lvert z\rvert)}$	$(2e^{\lvert z\rvert}-1)^{-1}$	$-\infty \ \infty$
3 - Non-Binary weight functions : Group III				
BETA	– – –	– – –	$(1-z)^{\alpha+1}(1+z)^{\beta+1}$ $\alpha, \beta > -1$	$(-1,1)$
GAMMA	– – –	– – –	$z^{\gamma+1}\, e^{-z}$ $\gamma > -1$	$(0, \infty)$
NORMAL	– – –	– – –	e^{-z^2}	$(-\infty, \infty)$
4 - Non Binary weight functions : Group IV				
	– – –	– – –	$z^t, \ -2 \le t \le 0$	$a > 0, \ b < \infty$
	– – –	– – –	$z^t, \ t < -2$	$a > 0, \ b \le \infty$
	– – –	– – –	$z^t, \ t > 0$	$a \ge 0, \ b < \infty$
	– – –	– – –	$exp(z)$	$a \ge -\infty, \ b < \infty$

Table 2.1: Weight Functions under consideration

[a] Z_w is widest possible design space
[b] s is sign of z

Name	$f(z)$	$F(z)$	$w(\cdot)$	z_1	z_2
LOGISTIC	$\frac{e^{-z}}{(1+e^{-z})^2}$	$\frac{1}{1+e^{-z}}$	$\frac{e^{z}}{(1+e^{z})^2}$	-1.543	1.543
SKEWED LOGISTIC	$m[F_1(z)]^{m-1}f_1(z)$	$(1+e^{-z})^{-m}$	$\left[\frac{me^{-z}}{1+e^{-z}}[(1+e^{-z})^m-1]^{-\frac{1}{2}}\right]^2$	$--$	$--$
m=1/3	$\cdots$	$\cdots$	$\cdots$	-4.409	0.552
m=2/3	$\cdots$	$\cdots$	$\cdots$	-2.284	1.191
m=3/2	$\cdots$	$\cdots$	$\cdots$	-0.939	1.898
m=3	$\cdots$	$\cdots$	$\cdots$	-0.060	2.525
GENERALIZED	$(e^z)(\lambda e^z+1)^{-\frac{1}{\lambda}-1}$	$1-[\lambda e^z+1]^{\frac{-1}{\lambda}}$	$\frac{e^{2z}}{(\lambda e^z+1)^2[(\lambda e^z+1)^{\frac{1}{\lambda}}-1]}$		
COMPLEMANTERY LOG-LOG	$e^{(z-e^z)}$	$1-e^{-e^z}$	$e^{2z}(e^{e^z}-1)^{-1}$	-1.338	0.980
PROBIT	$\frac{1}{\sqrt{2\pi}}e^{\frac{-z^2}{2}}$	$\int_{-\infty}^{z}\frac{1}{\sqrt{2\pi}}e^{\frac{-t^2}{2}}\,dt$	$\frac{\frac{1}{2\pi}e^{-z^2}}{\Phi(z)[1-\Phi(z)]}$	-1.138	1.138

Table 2.2: Group I : Binary weight functions ($w(z)=\frac{f^2(z)}{F(z)[1-F(z)]}$, $F(\cdot)=c.d.f.$ and $f(\cdot)=F'(\cdot)$), and support points z_1, z_2 of (global) D-optimal designs on widest possible design interval $Z_w=(-\infty,\infty)$.

NAME	$w(\cdot)$	Z_w[a]		$Support Points$
Beta	$(1-z)^{\alpha+1}(1+z)^{\beta+1}$	$(-1,1)$	$\alpha,\, \beta > -1$	$z_i = \frac{(\beta-\alpha)(\alpha+\beta+3)\pm 2\sqrt{(\alpha+2)(\beta+2)(\alpha+\beta+3)}}{(\alpha+\beta+3)(\alpha+\beta+4)},\ i=1,2$
Gamma	$z^{\gamma+1}\, e^{-z}$	$(0,\, \infty)$	$\gamma > -1$	$z_i = (\gamma+2) \pm \sqrt{(\gamma+2)},\ i=1,2$
Normal	e^{-z^2}	$(-\infty,\, \infty)$	$--$	$z_1 = \frac{-1}{\sqrt{2}},\ z_2 = \frac{1}{\sqrt{2}}$

Table 2.3: Group III : Non-Binary weight functions and support points z_1, z_2 of (global) D-optimal designs on Z_w.

[a] Z_w is widest possible design space

Name	$w(\cdot)$	$H(z) = -\frac{w'(z)}{[w(z)]^2}$	$H'(z)$	$H''(z)$
LOGISTIC	$\frac{e^z}{(1+e^z)^2}$	$e^z - e^{-z}$	$e^z + e^{-z}$	$e^z - e^{-z}$
SKEWED LOGISTIC	$\frac{m^2 e^{-2z}}{(1+e^{-z})^2[(1+e^{-z})^m - 1]}$	$\frac{(1+e^{-z})\{me^{-z}(1+e^{-z})^m - 2[(1+e^{-z})^m - 1]\}}{m^2 e^{-2z}}$	$---$	$---$
GENERALIZED	$\frac{e^{2z}}{(\lambda e^z+1)^2[(\lambda e^z+1)^{\frac{1}{\lambda}} - 1]}$	$2e^{-2z}(\lambda e^z + 1)\{(\lambda e^z + 1)^{\frac{2}{\lambda}-1} + \lambda e^z - e^z(\lambda e^z + 1)^{\frac{2}{\lambda}-2} - (\lambda e^z + 1)\}$	$---$	$---$
COMPLEMENTARY LOG-LOG	$e^{2z}(e^{e^z} - 1)^{-1}$	$\frac{2-2e^{e^z}}{e^{2z}} + \frac{e^{e^z}}{e^z}$	$\frac{e^{e^z}[4+e^{2z}-3e^z]-4}{e^{2z}}$	$\frac{e^{e^z}[7e^z+e^{3z}-3e^{2z}-8]+8}{e^{2z}}$
PROBIT	$\frac{\frac{1}{2\pi}e^{-z^2}}{\Phi(z)[1-\Phi(z)]}$	$---$	$---$	$---$

Table 2.4: Group I : Binary Weight Functions $w(z) = \frac{f^2(z)}{F(z)[1-F(z)]}$ and $H(z)$ $H'(z)$ $H''(z)$ functions on $Z_w = (-\infty, \infty)$

NAME	$w(\cdot)$	Z_w [a]		$H(z) = -\frac{w'(z)}{[w(z)]^2}$	$H'(z)$	$H''(z)$
BETA	$(1-z)^{\alpha+1}(1+z)^{\beta+1}$	$(-1,1)$	$\alpha, \beta > -1$	$\frac{(\alpha+1)(1-z)^{-1}-(\beta+1)(1+z)^{-1}}{[(1-z)^{(\alpha+1)}(1+z)^{(\beta+1)}]}$	$\frac{A_1{}^b}{A^c}$	$\frac{B_1{}^d+B_2{}^e+B_3{}^f+B_4{}^g}{B^h}$
BETA	$(1-z^2)^{\alpha+1}$	$(-1,1)$	$\alpha > -1$	$2z(\alpha+1)(1-z^2)^{-\alpha-2}$	$2(\alpha+1)(1-z^2)^{-\alpha-2}[1+\frac{2z^2(\alpha+2)}{(1-z^2)}]$	$4z(\alpha+1)(\alpha+2)(1-z^2)^{-\alpha-3}[3+\frac{2z^2(\alpha+3)}{(1-z^2)}]$
GAMMA	$z^{\gamma-1}e^{-z}$	$(0,\infty)$	$\gamma > 1$	$z^{-\gamma-2}e^z[z-(\gamma-1)]$	$z^{-\gamma-1}e^z\{[z-(\gamma-1)]^2+(\gamma-1)\}$	$z^{-\gamma-2}e^z\{[z-(\gamma-1)]^3-3(\gamma-1)(z-(\gamma-1))-2(\gamma-1)\}$
NORMAL	$e^{\frac{-z^2}{2}}$	$(-\infty,\infty)$	--	$ze^{\frac{z^2}{2}}$	$e^{\frac{z^2}{2}}(1+z^2)$	$e^{\frac{z^2}{2}}(3z+z^3)$

Table 2.5: Group III : Non-Binary weight Functions, $H(z)$, $H'(z)$ and $H''(z)$ functions

[a] Z_w is widest possible design space

[b] $A_1 = (1-z)^{-2}[(\alpha+1)+(\alpha+1)^2]+(1+z)^{-2}[(\beta+1)(\beta+1)^2]-2(\alpha+1)(\beta+1)(1-z)^{-1}(1+z)^{-1}$

[c] $A = (1-z)^{(\alpha+1)}(1+z)^{(\beta+1)}$

[d] $B_1 = (1-z)^{-3}[2(\alpha+1)+(\alpha+1)^2-(\alpha+1)^3]+(1+z)^{-3}[(\beta+1)^3-(\beta+1)^2-2(\beta+1)]$

[e] $B_2 = 2(1-z)^{-1}(1+z)^{-2}[(\alpha+1)(\beta+1)-(\alpha+1)(\beta+1)^2]$

[f] $B_3 = 2(1-z)^{-2}(1+z)^{-1}[(\alpha+1)^2(\beta+1)-(\alpha+1)(\beta+1)]$

[g] $B_4 = (1-z)^{-1}(1+z)^{-1}[(\beta+1)[(\alpha+1)+(\alpha+1)^2](1-z)^{-1}-(\alpha+1)[(\beta+1)+(\beta+1)^2](1+z)^{-1}]$

[h] $B = [(1-z)^{(\alpha+1)}(1+z)^{(\beta+1)}]^2$

Name	Logistic	Skewed Logistic	Generalized	Complementary log log	Probit
$w(\cdot)$	$\frac{e^z}{(1+e^z)^2}$	$\left[\frac{me^{-z}}{1+e^{-z}}[(1+e^{-z})^m-1]^{-\frac{1}{2}}\right]^2$	$\frac{e^{2z}}{(\lambda e^z+1)^2[(\lambda e^z+1)^{\frac{1}{\lambda}}-1]}$	$e^{2z}(e^{e^z}-1)^{-1}$	$\frac{\frac{1}{2\pi}e^{-z^2}}{\Phi(z)[1-\Phi(z)]}$
$\frac{w(z)}{w'(z)}$	$\frac{(1+e^z)}{(1-e^z)}$	$\frac{(1+e^{-z})[(1+e^{-z})^m-1]}{me^{-z}(1+e^{-z})^m-2[(1+e^{-z})^m-1]}$	$\frac{(\lambda e^z+1)^2-(\lambda e^z+1)^{\frac{2}{\lambda}}}{2\{(\lambda e^2+1)[(\lambda e^2+1)-\lambda e^z]+(\lambda e^z+1)^{\frac{2}{\lambda}}[e^z(\lambda e^z+1)^{-1}-1]\}}$	$\frac{(e^{e^z}-1)}{2(e^{e^z}-1)-e^z e^{e^z}}$	– – –
	$\uparrow^*$	$\uparrow^*$	$\uparrow^\dagger$	$\uparrow^*$	$\uparrow^\dagger$
$h(z)=z+\frac{2w(z)}{w'(z)}$	$z+\frac{2(1+e^z)}{(1-e^z)}$	$z+\frac{2(1+e^{-z})[(1+e^{-z})-1]}{me^{-z}(1+e^{-z})^m-2[(1+e^{-z})^m-1]}$	$z+\frac{[1-(\lambda e^z+1)]^{\frac{2}{\lambda}-2}}{[1-(\lambda e^z+1)^{\frac{2}{\lambda}-2}-\lambda e^{-z}(\lambda e^z+1)^{-1}+e^z(\lambda e^z+1)^{\frac{2}{\lambda}-3}]}$	$z+\frac{2(e^{e^z}-1)}{2(e^{e^z}-1)-e^z e^{e^z}}$	– – –
	$\uparrow^\dagger$	$\uparrow^\dagger$	$\uparrow^\dagger$	$\uparrow^\dagger$	$\uparrow^*$
$h'(z)=3-\frac{2w(z)w''(z)}{[w'(z)]^2}$	$1+\frac{4e^z}{(1-e^z)^2}$	$3-\frac{(1+e^{-z})[(1+e^{-z})-1]A^a}{me^{-z}(1+e^{-z})^m-2[(1+e^{-z})^m-1]^2}$		$1+\frac{e^z e^{e^z}(e^{e^z}-1)^{-1}\{e^z e^{e^z}(e^{e^z}-1)^{-1}-e^z-1\}}{\{2-e^z e^{e^z}(e^{e^z}-1)^{-1}\}^2}$	– – –

Table 2.6: Group I : Binary Weight Functions $(w(z)=\frac{f^2(z)}{F(z)[1-F(z)]})$ $\frac{w(z)}{w'(z)}$, $h(z)$ and $h'(z)$ functions on $Z_w=(-\infty,\infty)$

$^aA = \{(1+e^{-z})^m[2-3me^{-z}]+(1+e^{-z})^{m-1}[3me^{-2z}-m^2e^{-2z}-2me^{-z}]+2m^2e^{-2z}(1+e^{-z})^{2m-1}[(1+e^{-z})^m-1]^{-1}-6e^{-z}(1+e^{-z})^{-1}[(1+e^{-z})^m-1]\}$

$\uparrow^\dagger$ implies empirical evidence of nondecreasing property.

$\uparrow^*$ Wu(1988) Wu (1988) proved these cases.

NAME	$w(\cdot)$	Z_w [a]		$\frac{w(z)}{w'(z)}$	$h(z) = z + \frac{2w(z)}{w'(z)}$	$h'(z) = 3 - \frac{2w(z)w''(z)}{[w'(z)]^2}$
BETA	$(1-z)^{\alpha+1}(1+z)^{\beta+1}$	$(-1, 1)$	$\alpha, \beta > -1$	$\frac{(1-z^2)}{(\beta-\alpha)-(\beta+\alpha+2)z}$	$z + \frac{2(1-z^2)}{(\beta-\alpha)-(\beta+\alpha+2)z}$	$1 + \frac{2(\beta+\alpha+2)z^2-4z(\beta-\alpha)+2(\beta+\alpha+2)}{[(\beta-\alpha)-(\beta+\alpha+2)z]^2}$
BETA	$(1-z^2)^{\alpha+1}$	$(-1, 1)$	$\alpha > -1$	$\frac{(1-z^2)}{-2z(\alpha+1)}$	$z - \frac{(1-z^2)}{z(\alpha+1)}$	$1 + \frac{2}{(1-z^2)(\alpha+1)} + \frac{(1-z^2)}{z^2(\alpha+1)}$
GAMMA	$z^{\gamma-1} e^{-z}$	$(0, \infty)$	$\gamma > 1$	$\frac{1}{[(\gamma-1)z^{-1}-1]}$	$z + \frac{2}{[(\gamma-1)z^{-1}-1]}$	$1 + \frac{(\gamma-1)}{z^2} + \frac{1}{[(\gamma-1)z^{-1}-1]}$
NORMAL	$e^{\frac{-z^2}{2}}$	$(-\infty, \infty)$	$--$	$\frac{1}{-z}$	$z - \frac{2}{z}$	$1 + \frac{2}{z^2}$

Table 2.7: Group III : Non-Binary weight Functions, $\frac{w(z)}{w'(z)}$, $h(z)$ and $h'(z)$ functions

[a] Z_w widest possible design interval

Name	$f(z)$	$F(z)$	$w(\cdot)$	z_1	z_2
DOUBLE RECIPROCAL	$\frac{1}{2}(1+\|z\|)^{-2}$	$\frac{(1+s)}{2} - \frac{s}{2}(1+\|z\|)^{-1}$	$\frac{(1+\|z\|)^{-2}}{2\|z\|+1}$	−0.390	0.390
DOUBLE EXPONENTIAL	$\frac{1}{2}e^{(-\|z\|)}$	$\frac{1+s}{2} - \frac{s}{2}e^{(-\|z\|)}$	$(2e^{\|z\|}-1)^{-1}$	−0.768	0.768

Table 2.8: Group II : Binary Weight Functions $w(z) = \frac{f^2(z)}{F(z)[1-F(z)]}$ ($F(\cdot) = c.d.f.$, $f(\cdot) = F'(\cdot)$, s is sign of z) and Support points z_1, z_2 of two-point D-Optimal designs on widest possible design space $Z_w = (-\infty, \infty)$.

Name	$w(\cdot)$	$H(z) = \frac{-w'(z)}{[w(z)]^2}$	$H'(z)$	$H''(z)$
DOUBLE RECIPROCAL	$\frac{(1+\|z\|)^{-2}}{s\|z\|+1}$	$6sz^2 + 10\|z\| + 4s$	$12\|z\| + 10$	$12s$
DOUBLE EXPONENTIAL	$(2e^{\|z\|} - 1)^{-1}$	$2se^{\|z\|}$	$2e^{\|z\|}$	$2se^{\|z\|}$

Table 2.9: Group II : Binary Weight Functions $w(z) = \frac{f^2(z)}{F(z)[1-F(z)]}$ $(F(\cdot) = c.d.f.,$ $f(\cdot) = F'(\cdot)$, s is sign of z) and $H(z)$, $H'(z)$, $H''(z)$ functions on $Z_w = (-\infty, \infty)$

Name	$w(\cdot)$	$\frac{w(z)}{w'(z)}$		$h(z) = z + \frac{2w(z)}{w'(z)}$		$h'(z) = 3 - \frac{2w(z)w''(z)}{[w'(z)]^2}$
DOUBLE RECIPROCAL	$\frac{(1+\|z\|)^{-2}}{2\|z\|+1}$	$-\frac{1}{2}\{s + \frac{2z^2-1}{3z+2s}\}$	$\downarrow^{\dagger}$	$z - \frac{2z^2+3\|z\|+1}{3z+2s}$	$\uparrow^{*}$	$1 - \frac{[2\|z\|(3\|z\|+4)+3]}{[3\|z\|(3\|z\|+4)+4]}$
DOUBLE EXPONENTIAL	$(2e^{\|z\|} - 1)^{-1}$	$\frac{2e^{\|z\|}-1}{2se^{\|z\|}}$	$\downarrow^{\dagger}$	$z - s(2 - e^{-\|z\|})$	$\uparrow^{*}$	$1 - e^{-\|z\|}$

Table 2.10: Group II : Binary Weight Functions $[w(z) = \frac{f^2(z)}{F(z)[1-F(z)]})$ $(F(\cdot) = c.d.f.,\ \ f(\cdot) = F'(\cdot)$, s is sign of $z]$ $\frac{w(z)}{w'(z)}$, $h(z)$ and $h'(z)$ functions on the $Z_w = (-\infty, \infty)$

$\downarrow^{\dagger}$ implies emprical evidence of decreasing property.
$\uparrow^{*}$ Wu (1988) Wu (1988) proved these cases.

$w(\cdot)$	$Z = [a, b]$	z_1	z_2
z^t, $-2 \leq t \leq 0$	$a > 0$, $b < \infty$	a	b
z^t, $t < -2$	$a > 0$, $b \leq \infty$	a	$min\{b,\, ta/(t+2)\}$
z^t, $t > 0$	$a \geq 0$, $b < \infty$	$max\{a,\, tb/(t+2)\}$	b
$exp(z)$	$a \geq -\infty$, $b < \infty$	$max\{a,\, b-2\}$	b

Table 2.11: Group IV : Non-Binary weight functions and support points z_1, z_2 of D-optimal design on $Z = [a, b]$. (See Ford,Torsney & Wu (1992))

$w(\cdot)$	$Z = [a, b]$	$H(z) = -\frac{w'(z)}{[w(z)]^2}$		$H'(z)$	$H''(z)$	
$z^t, -2 \leq t \leq 0$	$a > 0, b < \infty$	$-tz^{-t-1}$	$\downarrow$	$t(t+1)z^{-(t+2)}$	$-t(t+1)(t+2)z^{-(t+3)}$	$\downarrow$
$z^t, t < -2$	$a > 0, b \leq \infty$	$-tz^{-t-1}$	$\uparrow$	$t(t+1)z^{-(t+2)}$	$-t(t+1)(t+2)z^{-(t+3)}$	$\uparrow$
$z^t, t > 0$	$a \geq 0, b < \infty$	$-tz^{-t-1}$	$\uparrow$	$t(t+1)z^{-(t+2)}$	$-t(t+1)(t+2)z^{-(t+3)}$	$\uparrow$
e^z	$a \geq -\infty, b < \infty$	$-e^{-z}$	$\uparrow$	e^{-z}	$-e^{-z}$	$\uparrow$

Table 2.12: Group IV : Non-Binary Weight Functions, $H(z)$ $H'(z)$ and $H''(z)$

$w(\cdot)$	$Z = [a, b]$	$\frac{w(z)}{w'(z)}$		$h(z) = z + \frac{2w(z)}{w'(z)}$		$h'(z)$	
$z^t, -2 \leq t \leq 0$	$a > 0, b < \infty$	$\frac{z}{t}$	$\downarrow$	$z + \frac{2z}{t}$	$\downarrow$	$1 + \frac{2}{t}$	$\downarrow$
$z^t, t < -2$	$a > 0, b \leq \infty$	$\frac{z}{t}$	$\downarrow$	$z + \frac{2z}{t}$	$\uparrow$	$1 + \frac{2}{t}$	$\uparrow$
$z^t, t > 0$	$a \geq 0, b < \infty$	$\frac{z}{t}$	$\uparrow$	$z + \frac{2z}{t}$	$\uparrow$	$1 + \frac{2}{t}$	$\uparrow$
e^z	$a \geq -\infty, b < \infty$	1	$\uparrow$	$z + 1$	$\uparrow$	1	$\uparrow$

Table 2.13: Group IV : Non-Binary weight Functions, $\frac{w(z)}{w'(z)}$, $h(z)$ and $h'(z)$ functions

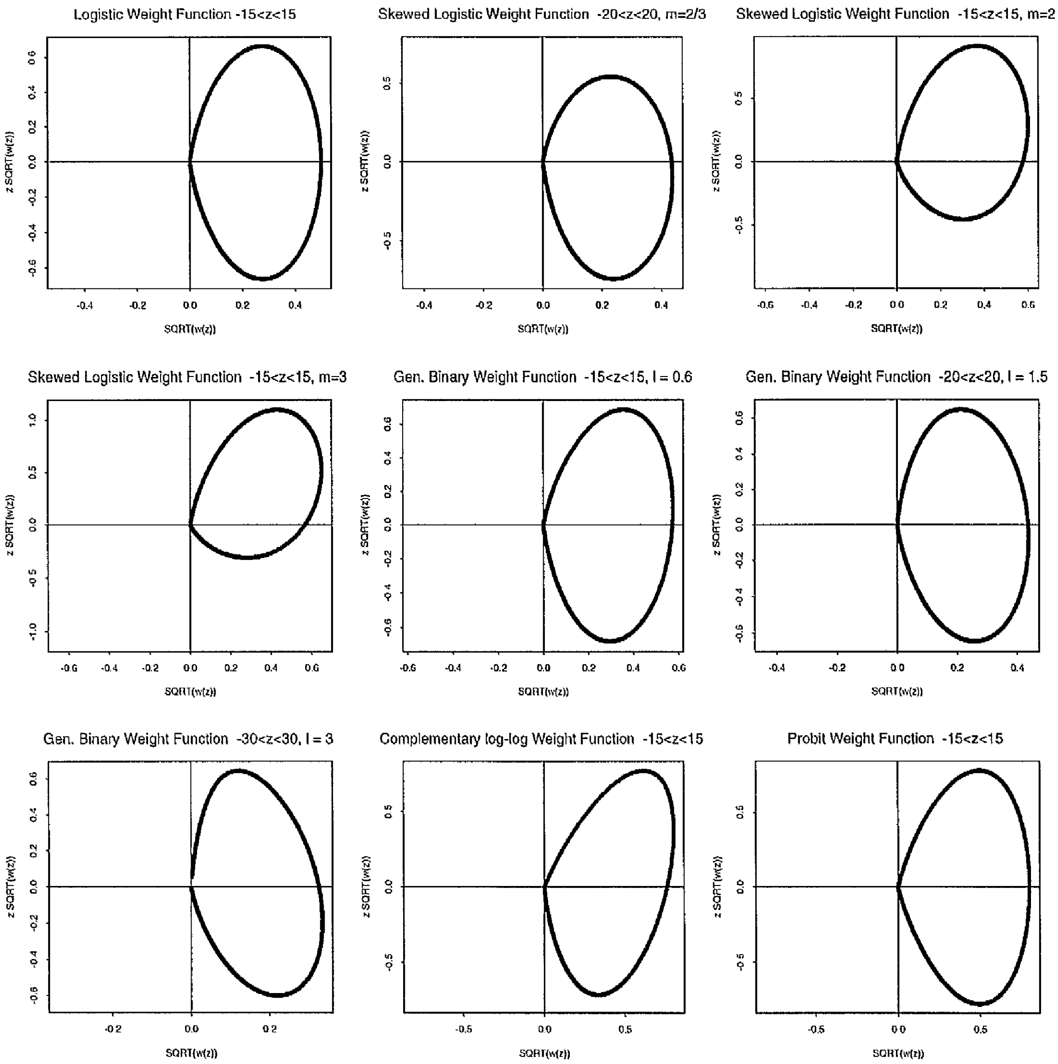


Figure 2.1: Group : Plot of $G(\mathcal{Z})$ for Binary Weight Functions $\mathcal{Z}_w = (-\infty, \infty)$, (Note: l represents λ).

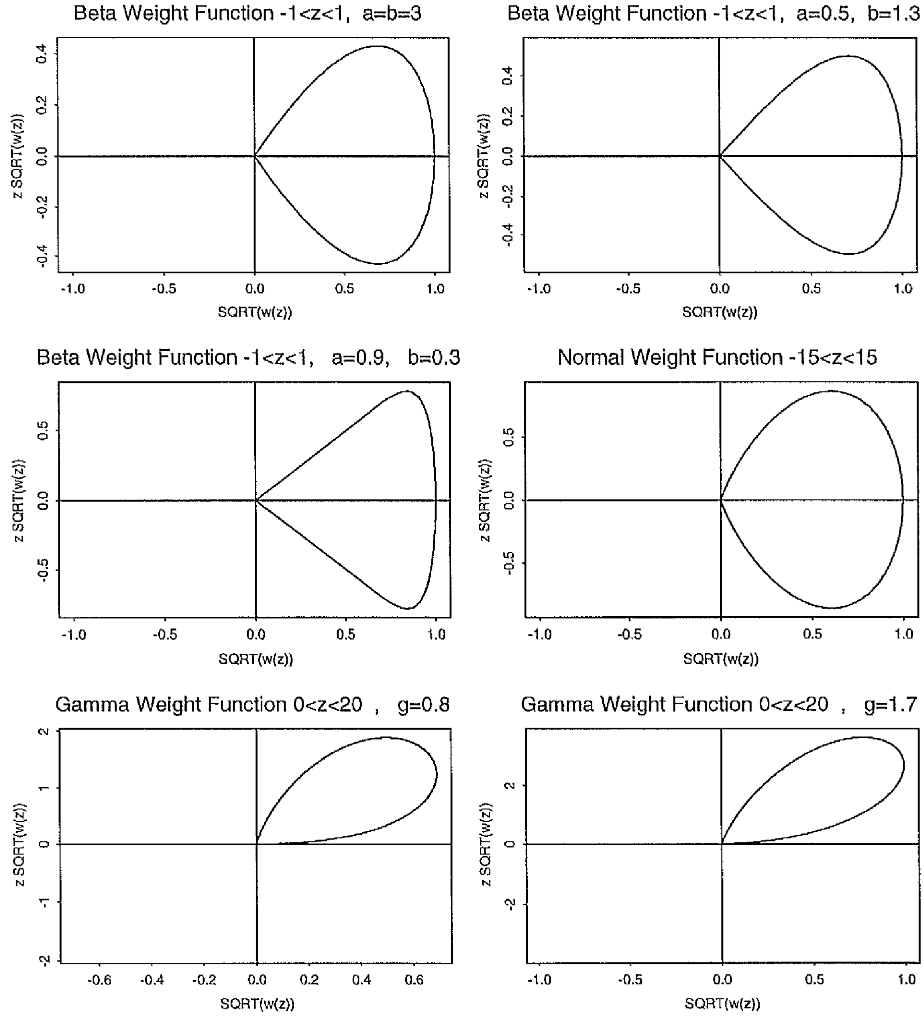


Figure 2.2: Group III : Plot of $G(\mathcal{Z})$ for Non-Binary Weight Functions on $\mathcal{Z}_w$, (Note: l represents λ).

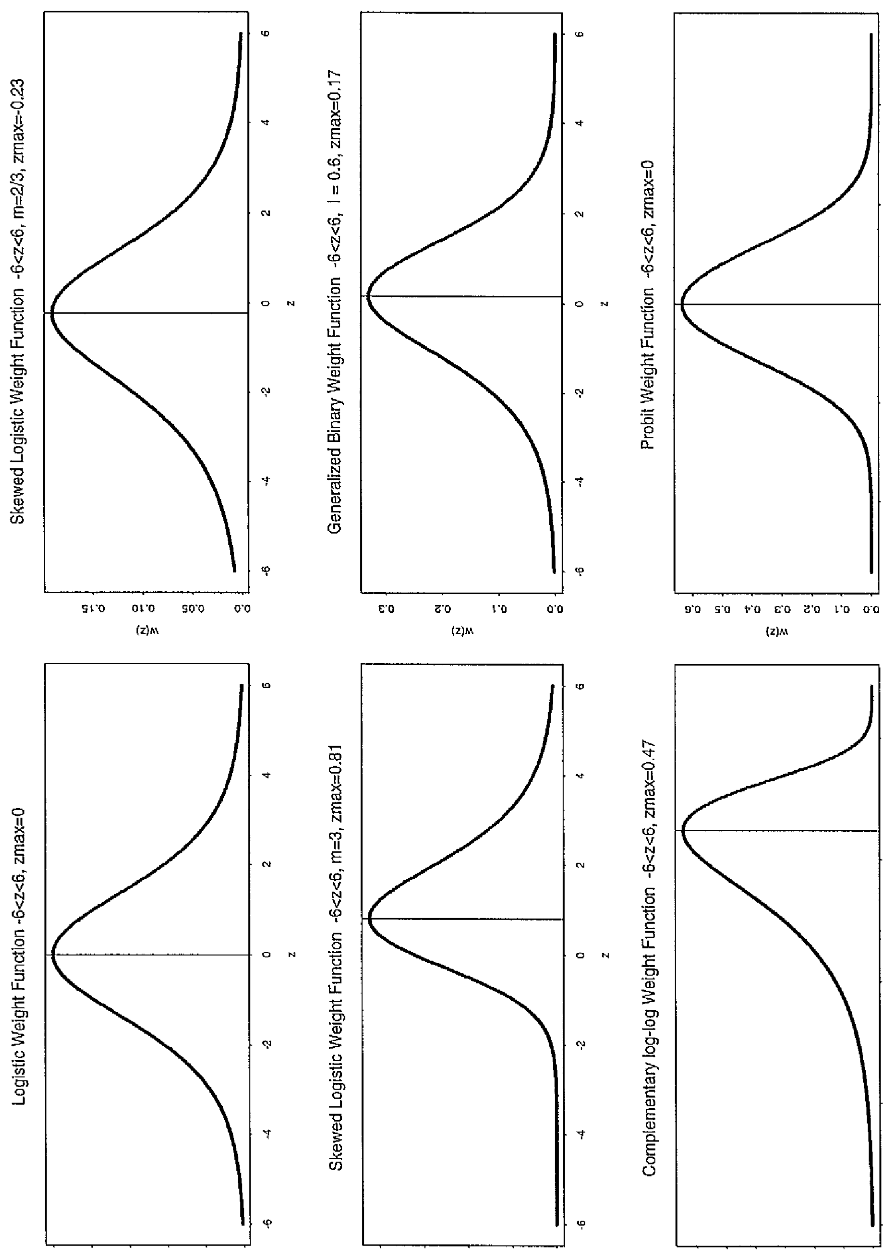


Figure 2.3: Group I : Plot of Binary Weight Functions ($w(z)$) on $\mathcal{Z}_w = (-\infty, \infty)$, (Note: l represents λ).

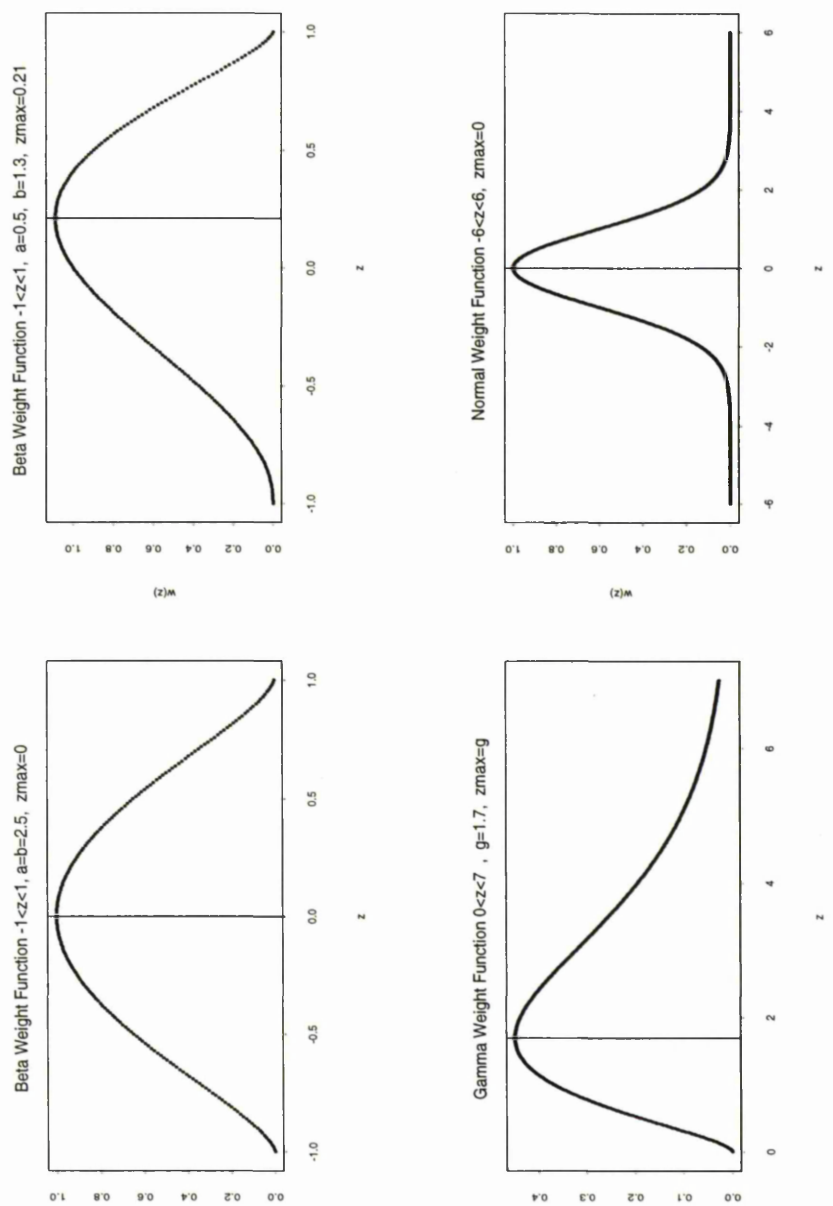


Figure 2.4: Group III : Plot of Non-Binary Weight Functions $(w(z))$ $\mathcal{Z}_w = (-\infty, \infty)$

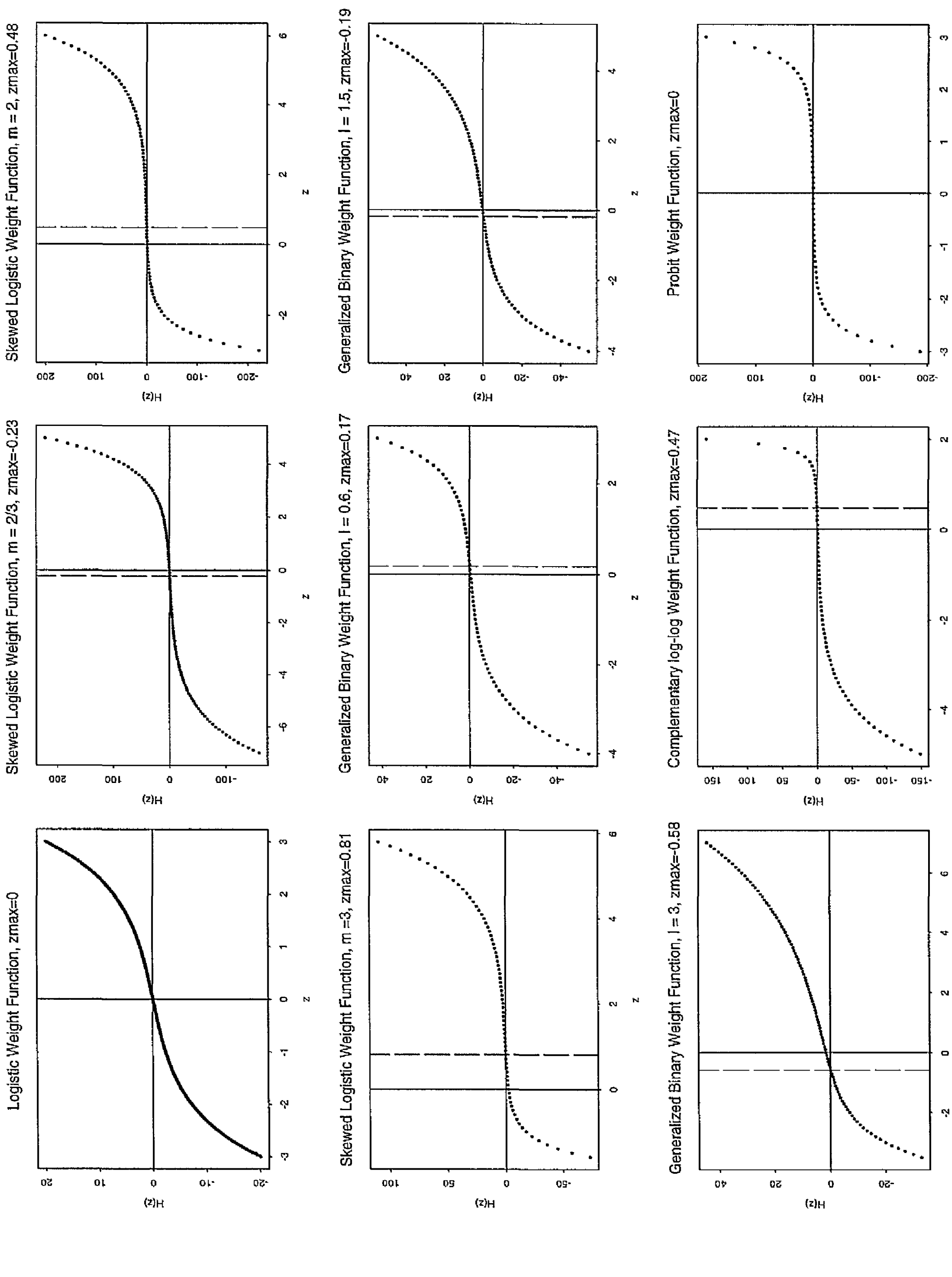


Figure 2.5: Group I : Plots of $H(z)$ for Binary Weight Functions (Note: l represents λ)

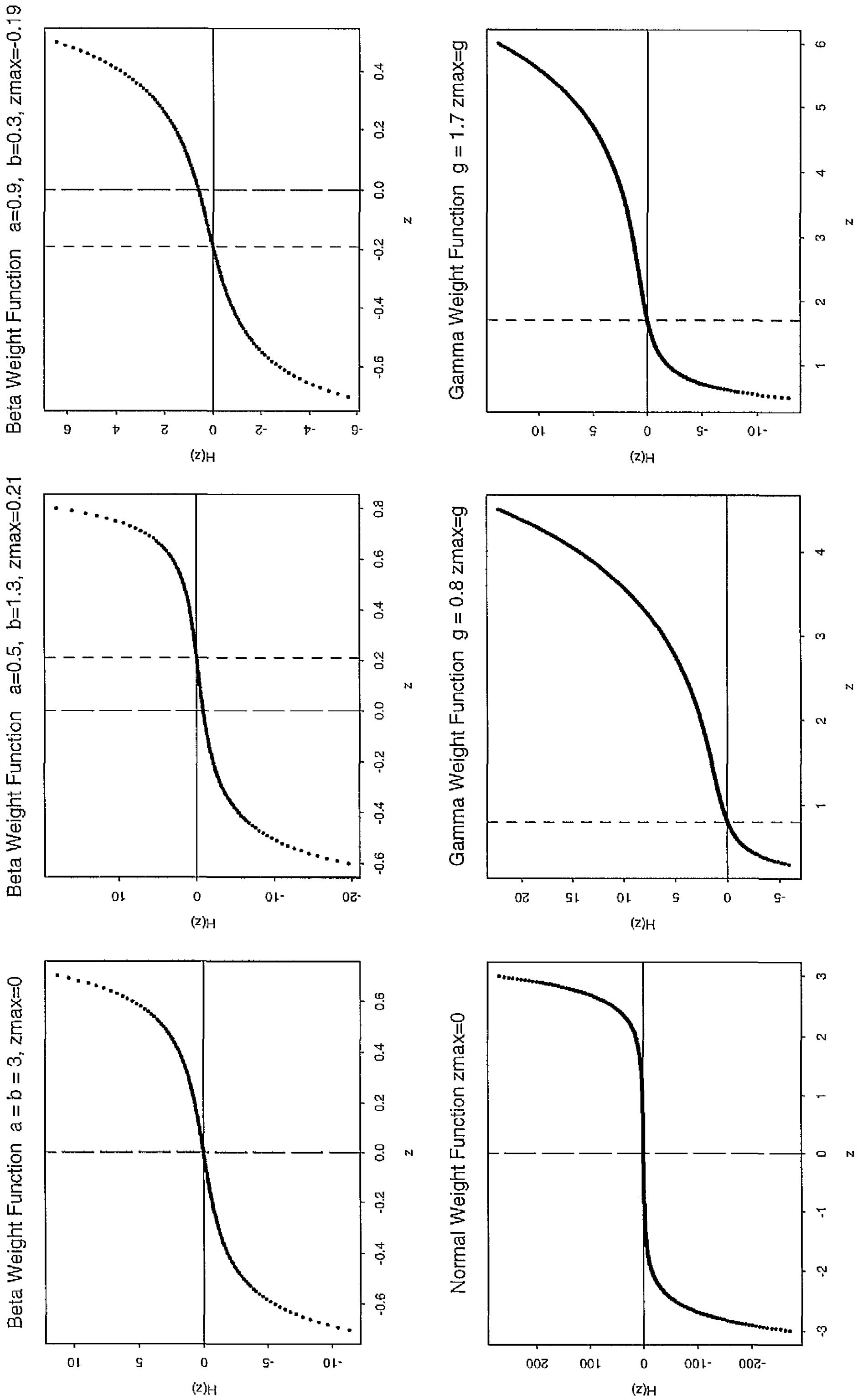


Figure 2.6: Group III : Plot of $H(z)$ for Non-Binary Weight Functions.

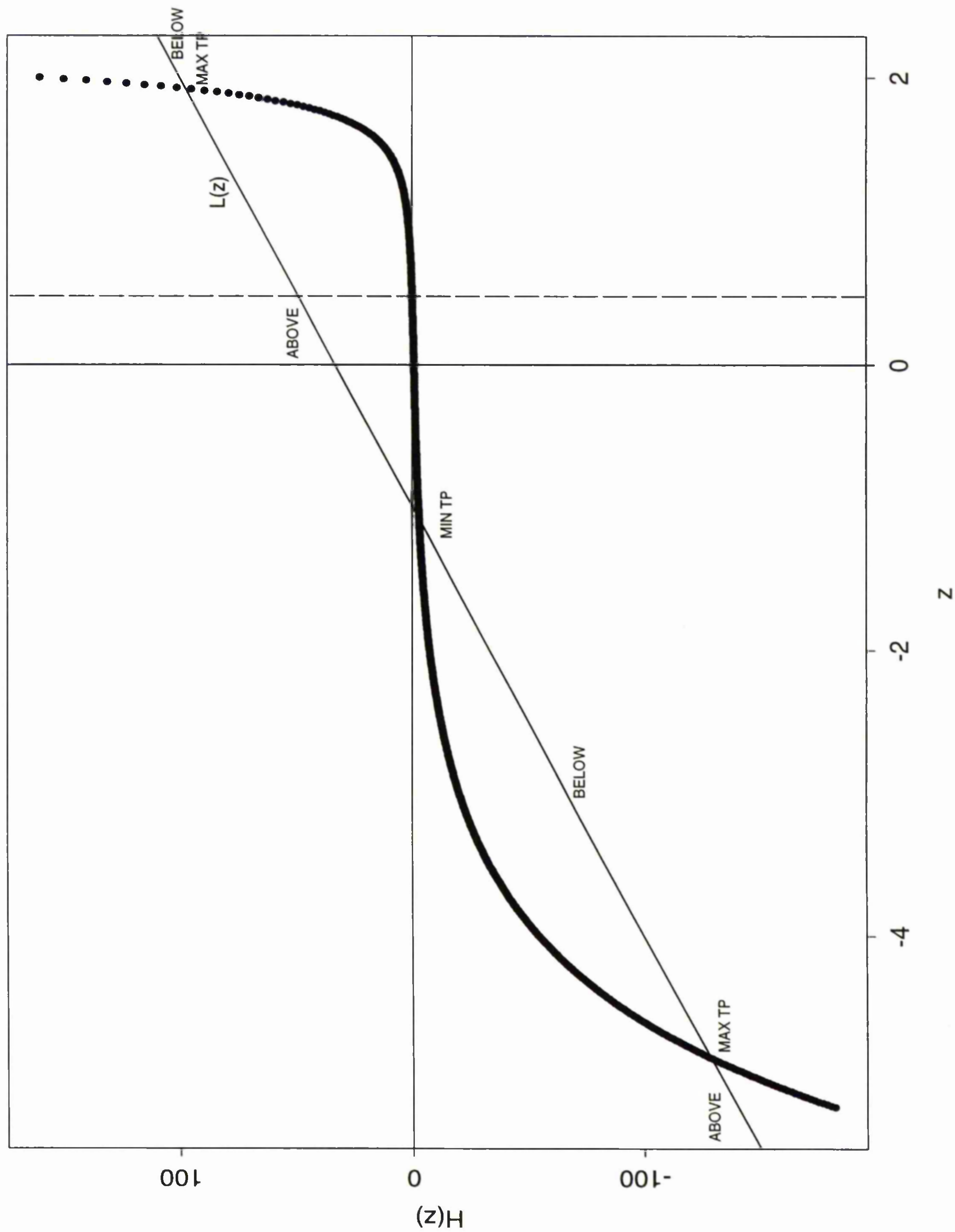


Figure 2.7: Plot of $H(z)$ and $L(z)$ for Complementary log-log Weight Function, $z_{max} = 0.47$.

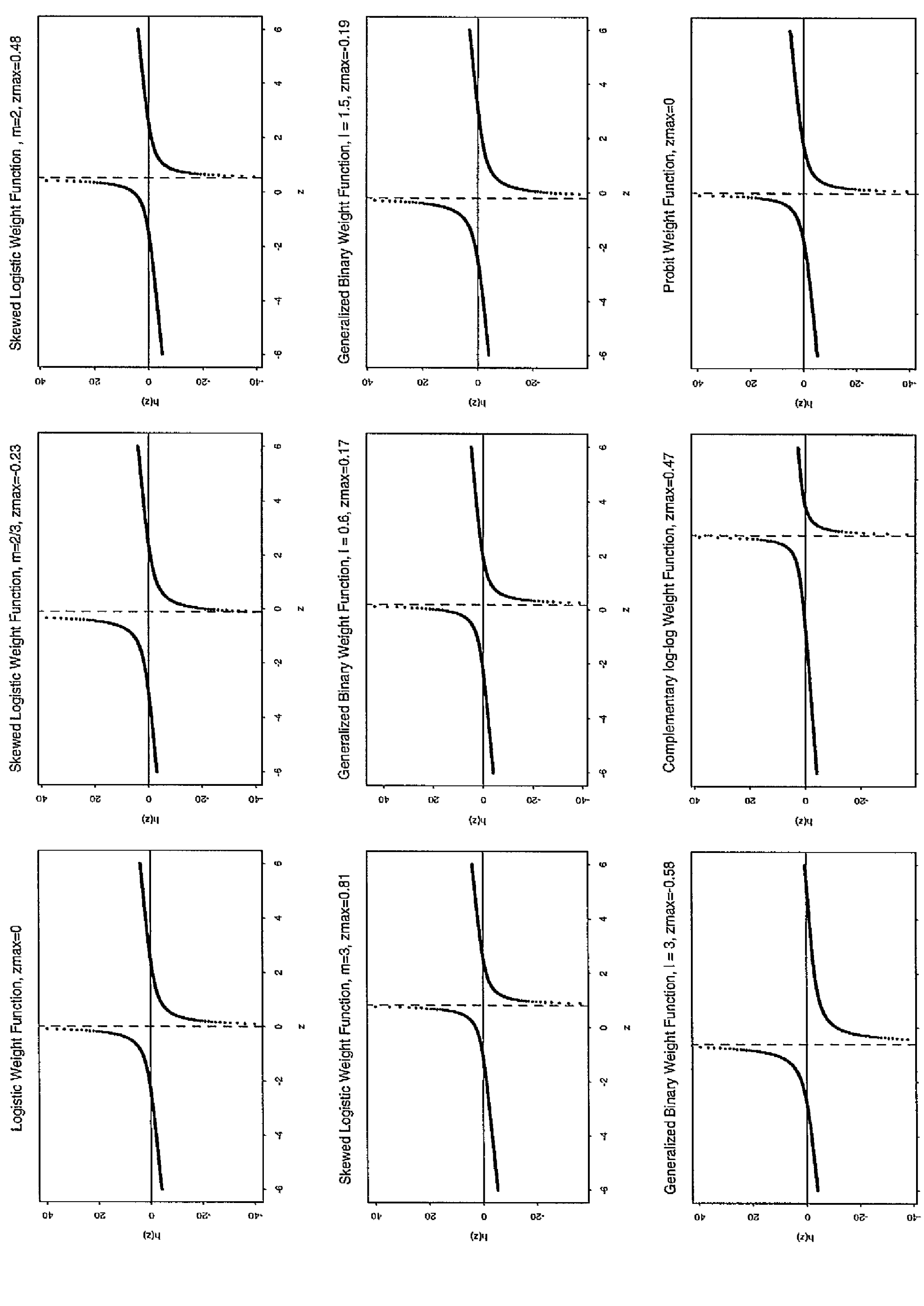


Figure 2.8: Group I : Plot of $h(z)$ for Binary Weight Functions, (Note: l represents λ).

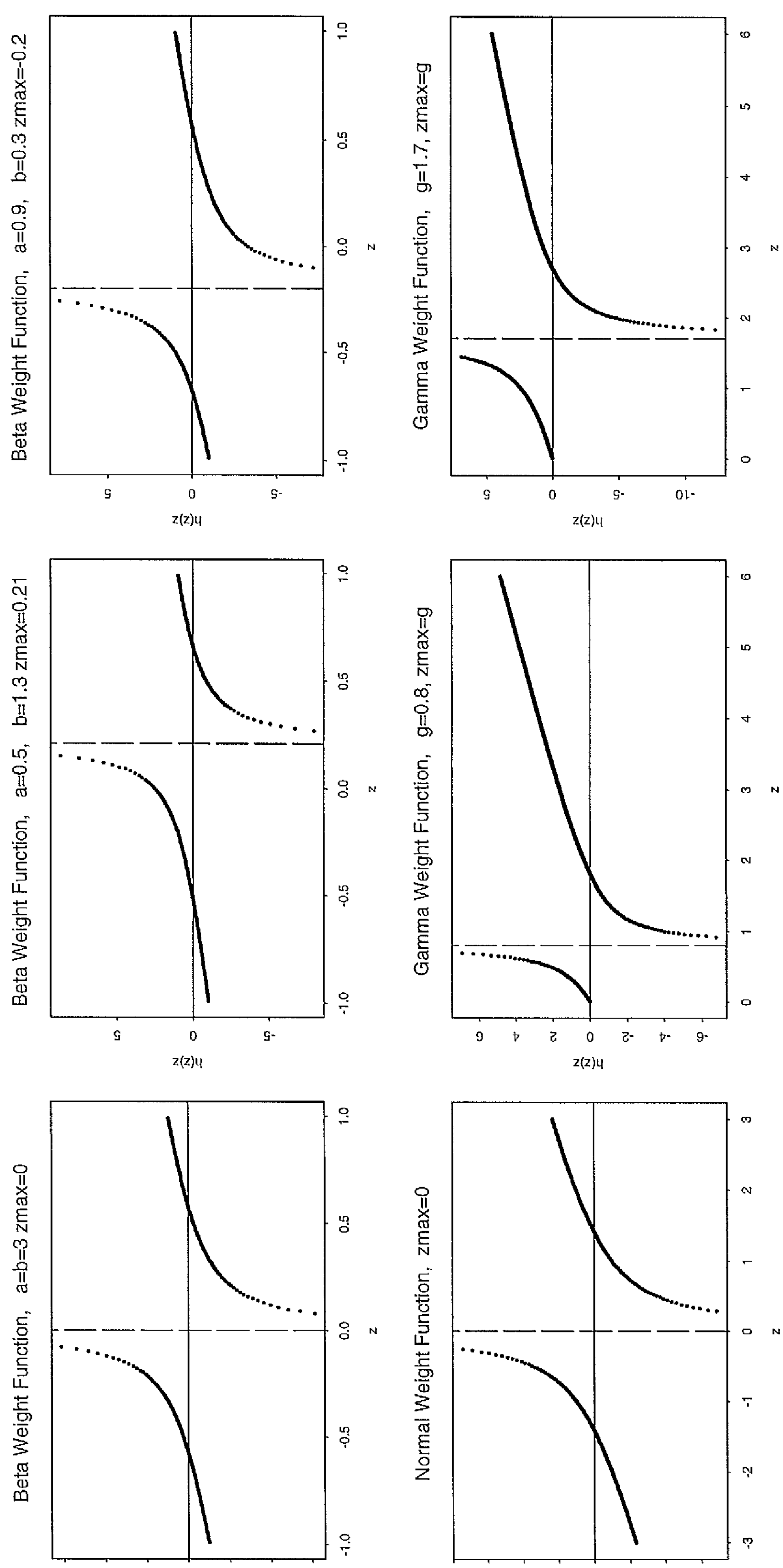


Figure 2.9: Group III : Plot of $h(z)$ for Non-Binary Weight Functions.

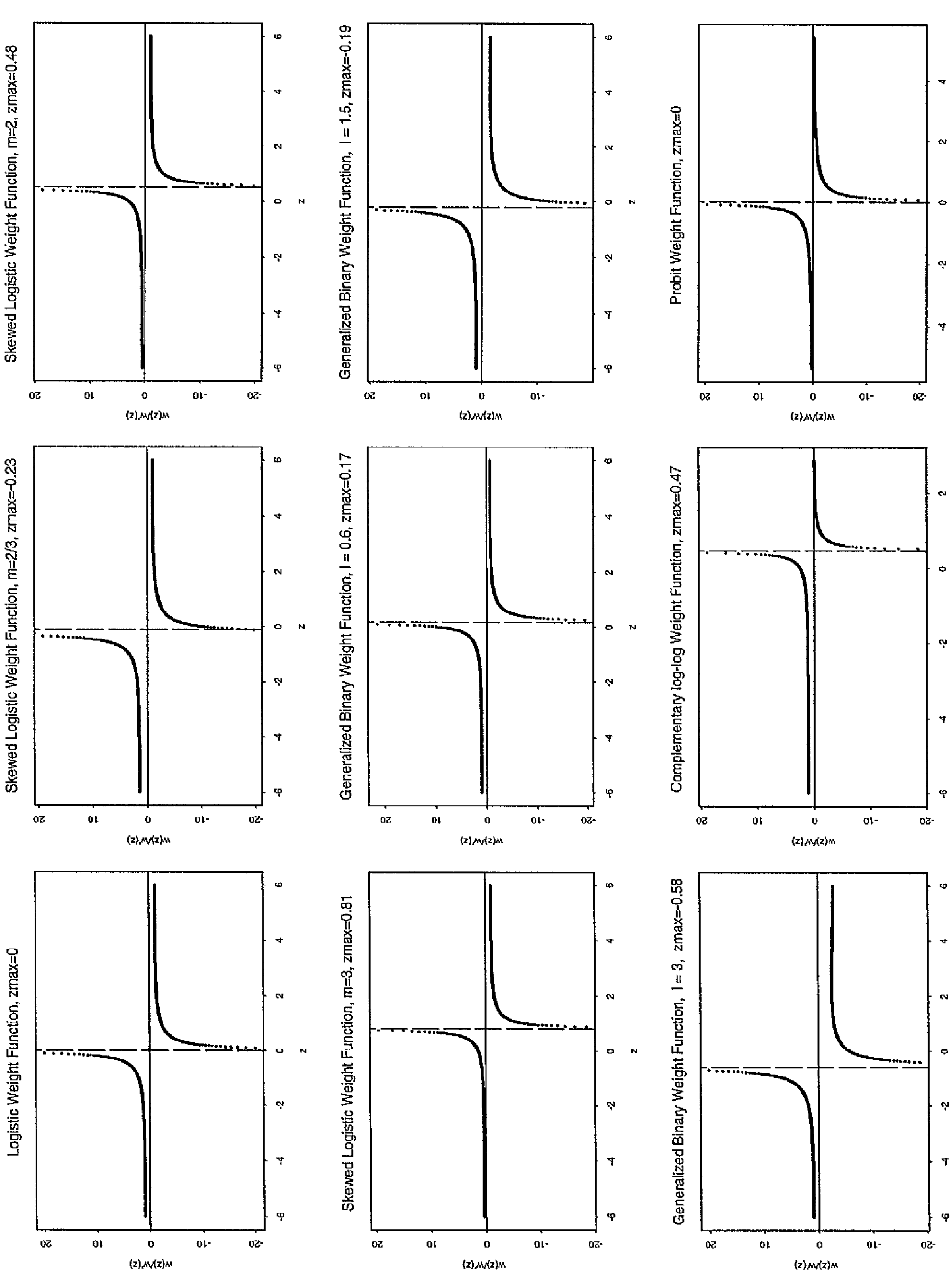


Figure 2.10: Group I : Plot of $w(z)/w'(z)$ for Binary Weight Functions, (Note: l represents λ).

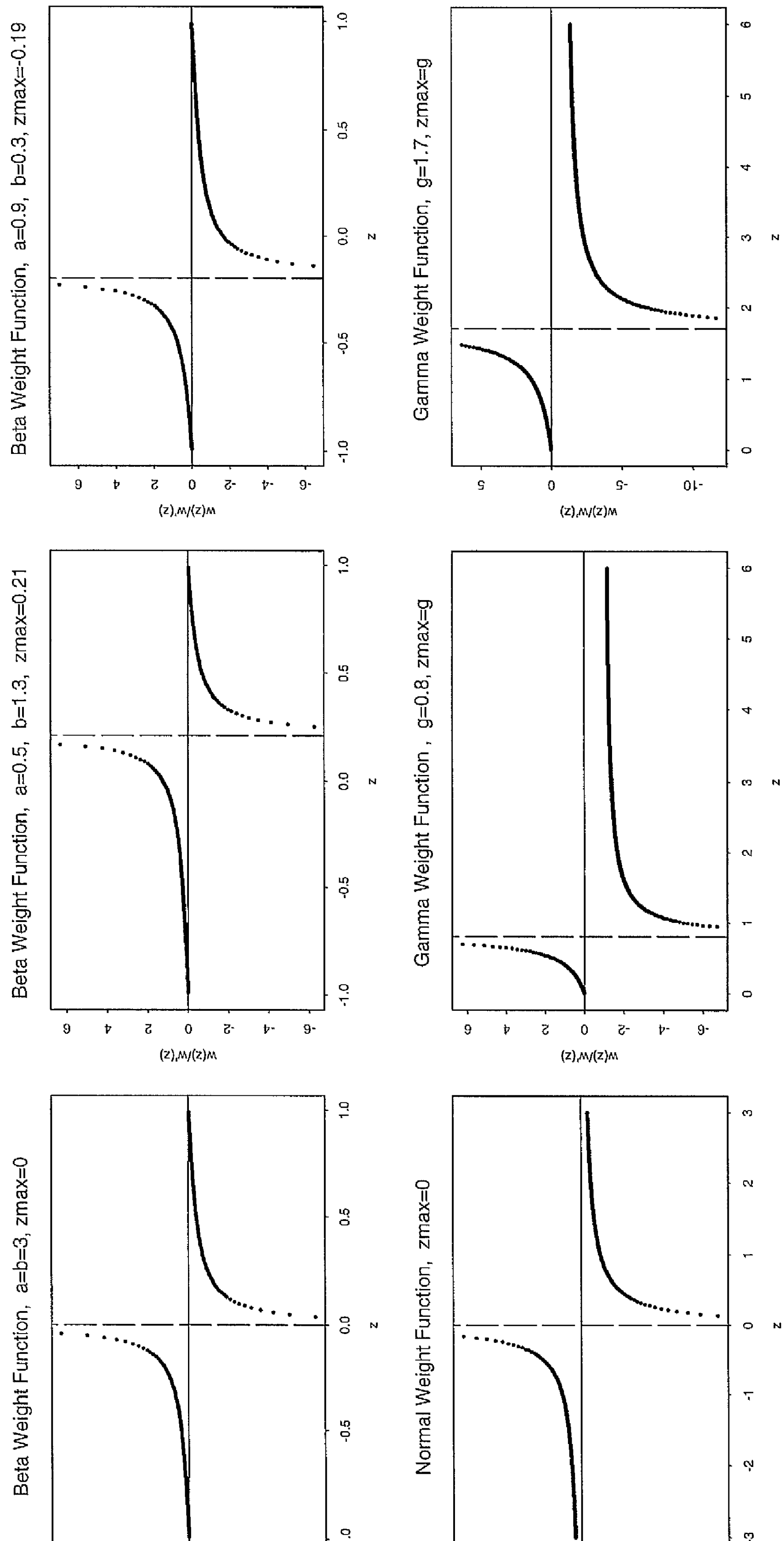


Figure 2.11: Group III : Plot of $w(z)/w'(z)$ for Non-Binary Weight Functions.

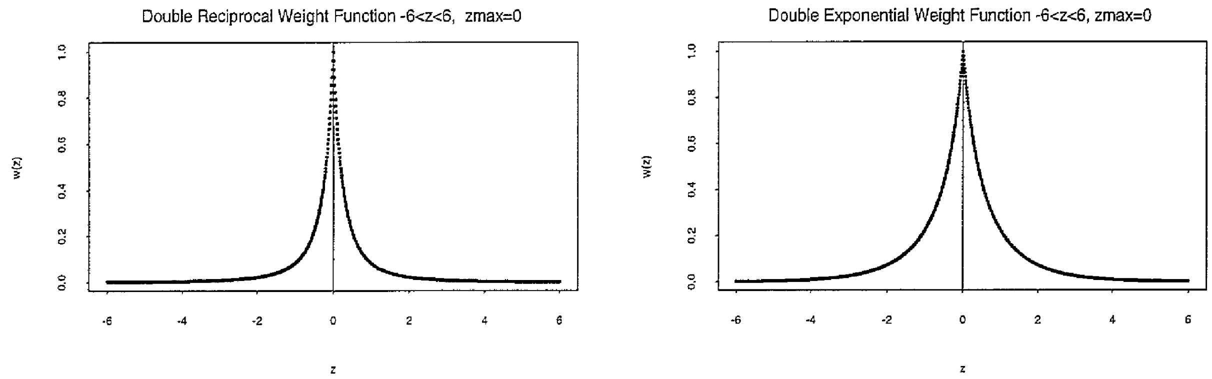


Figure 2.12: Group II : Plot of Double Reciprocal & Double Exponential Binary Weight Functions $(w(z))$ $\mathcal{Z}_w = (-\infty, \infty)$.

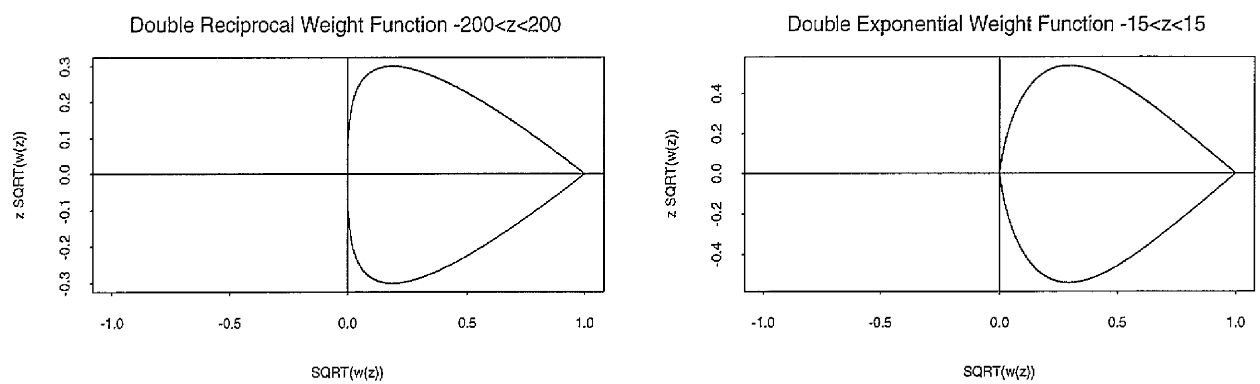


Figure 2.13: Group II :Plot of $G(\mathcal{Z})$ for Double Reciprocal & Double Exponential Weight Functions $(w(z))$ $\mathcal{Z}_w = (-\infty, \infty)$.

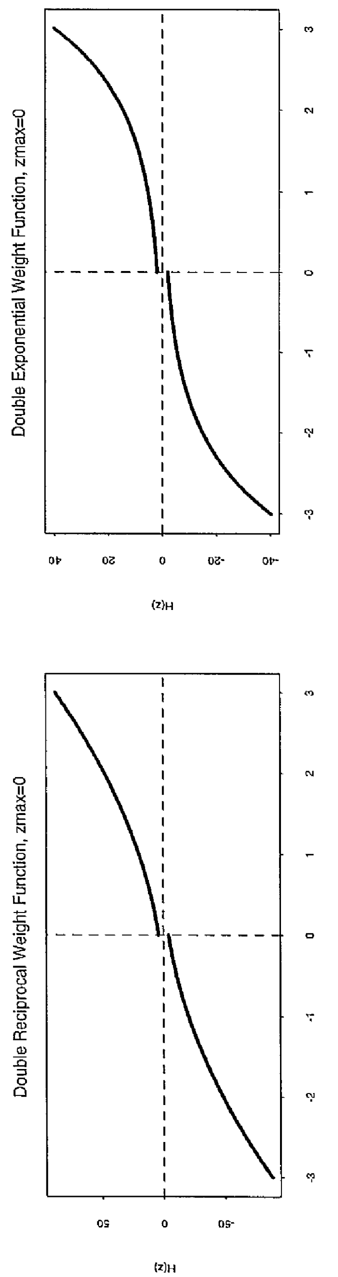


Figure 2.14: Group II : Plot of $H(z)$ for Double Reciprocal & Double Exponential Binary Weight Functions.

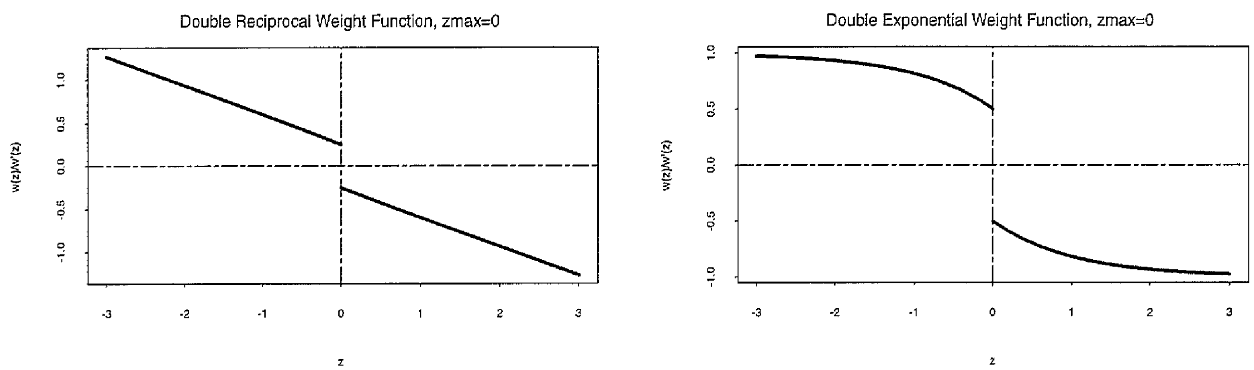


Figure 2.15: Group II : Plot of $w(z)/w'(z)$ for Double Reciprocal & Double Exponential Weight Functions $\mathcal{Z}_w = (-\infty, \infty)$.

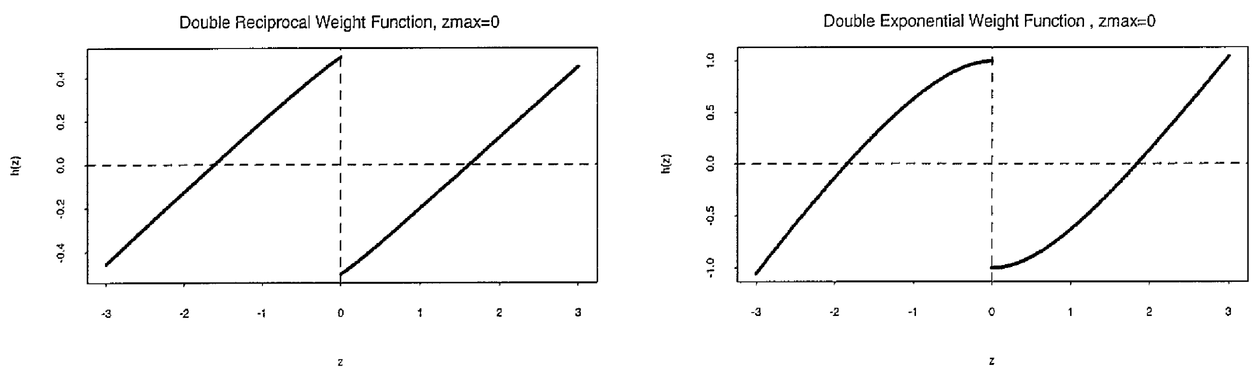


Figure 2.16: Group II : Plot of $h(z)$ for Double Reciprocal & Double Exponential Binary Weight Functions $\mathcal{Z}_w = (-\infty, \infty)$.

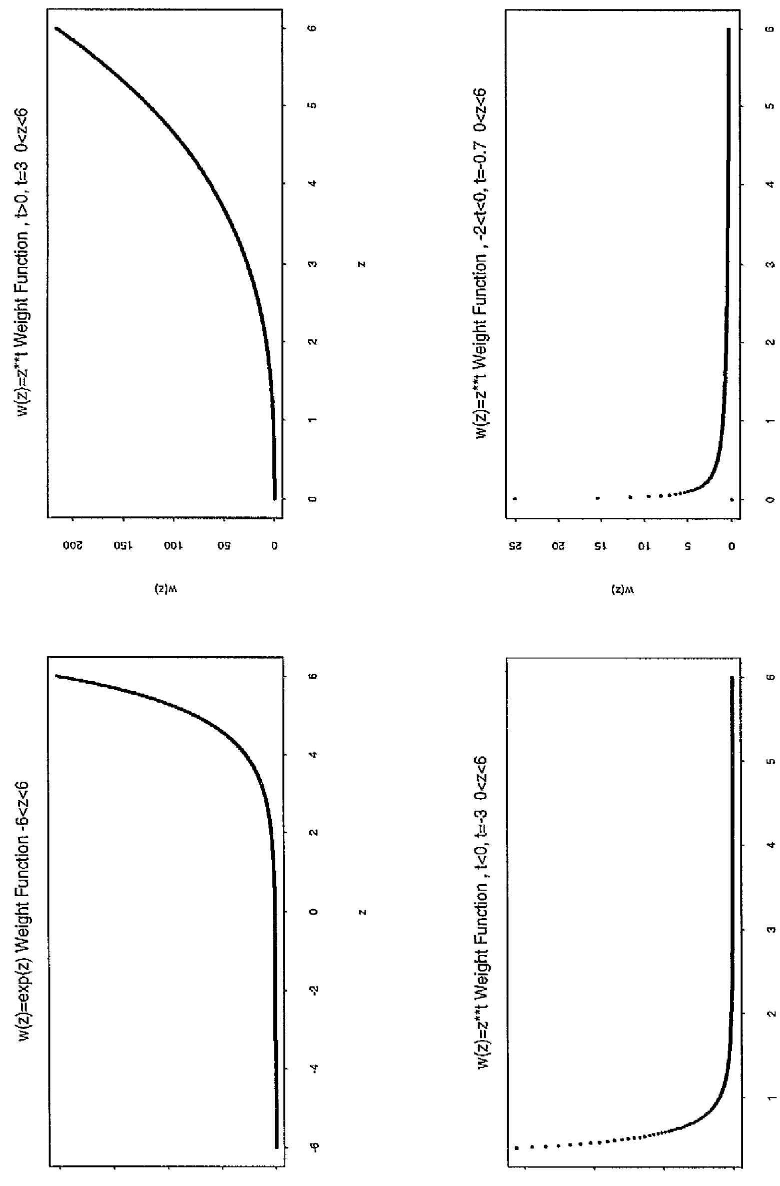


Figure 2.17: Group IV : Plot of Some Non-Binary Weight Functions $(w(z))$ $\mathcal{Z}_w = (-\infty, \infty)$.

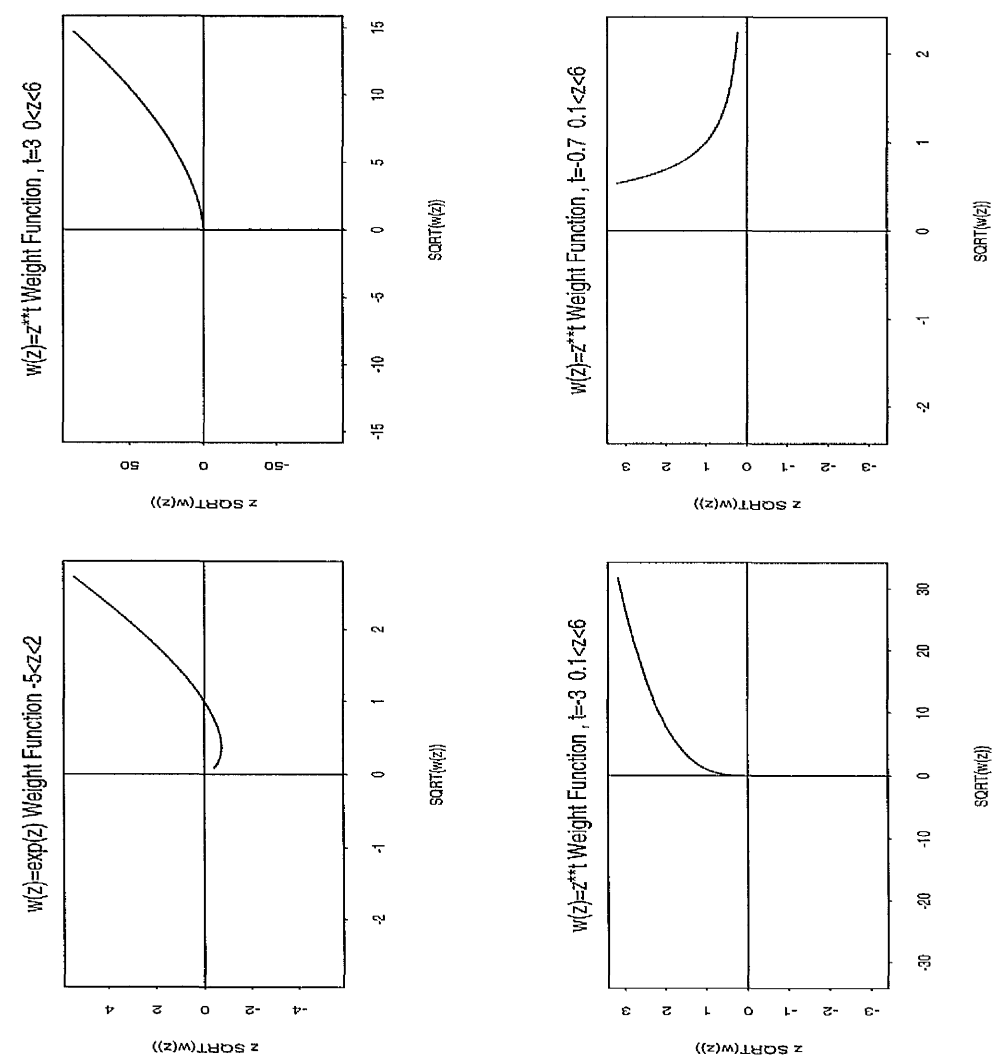


Figure 2.18: Group IV : Plot of $G(\mathcal{Z})$ for Some Non-Binary Weight Functions on $\mathcal{Z}$.

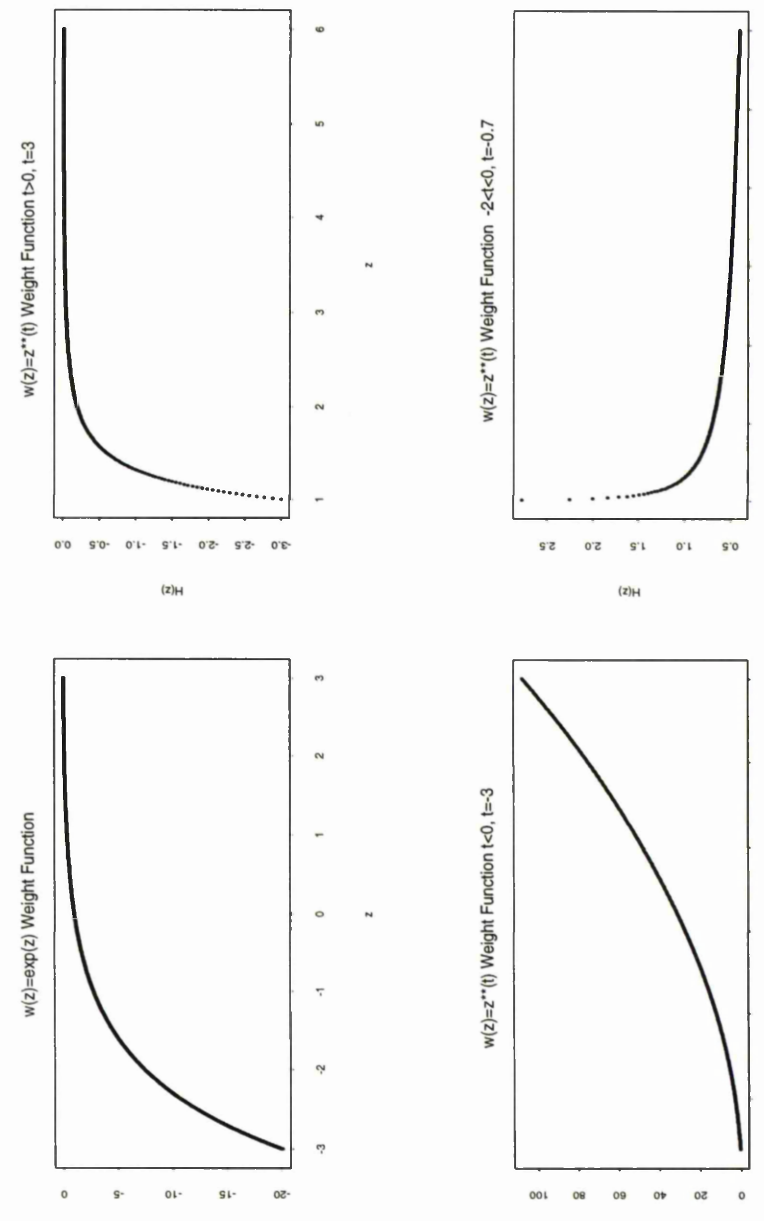


Figure 2.19: Group IV : Plot of $H(z)$ for Non-Binary Weight Functions.

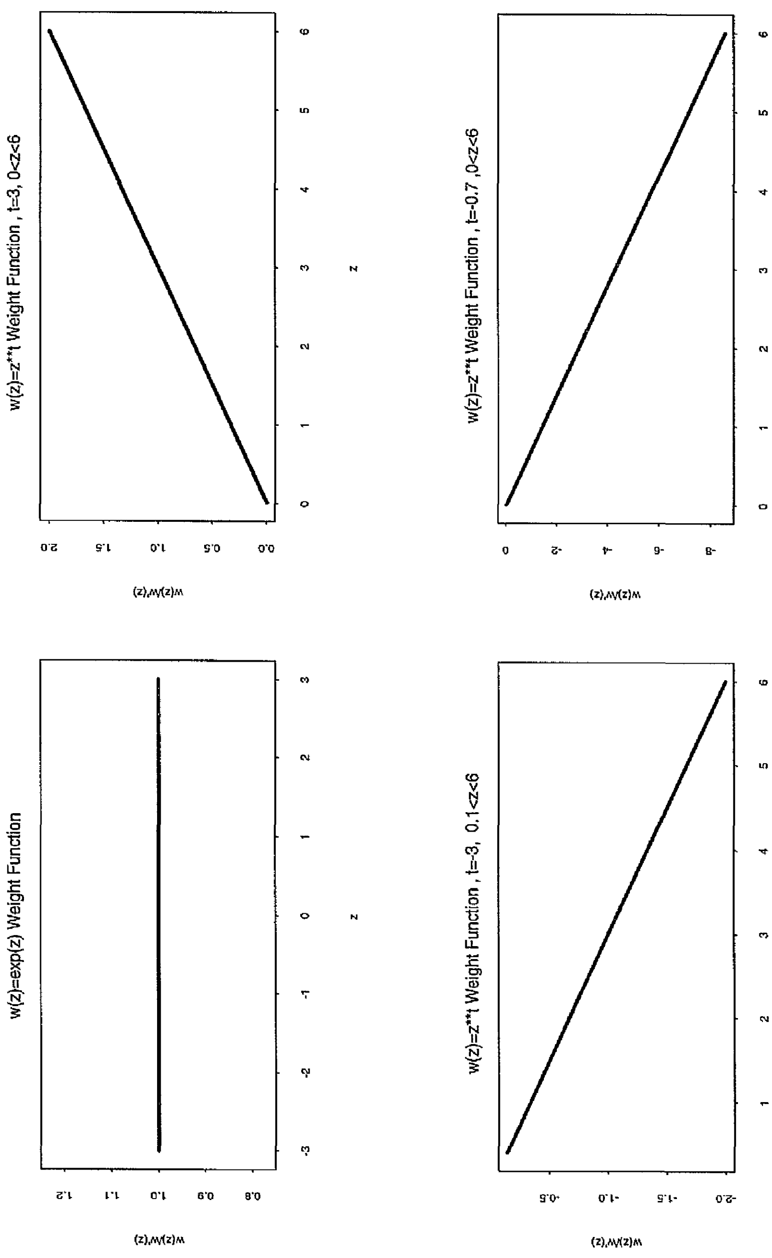


Figure 2.20: Group IV : Plot of $w(z)/w'(z)$ for Non-Binary Weight Functions.

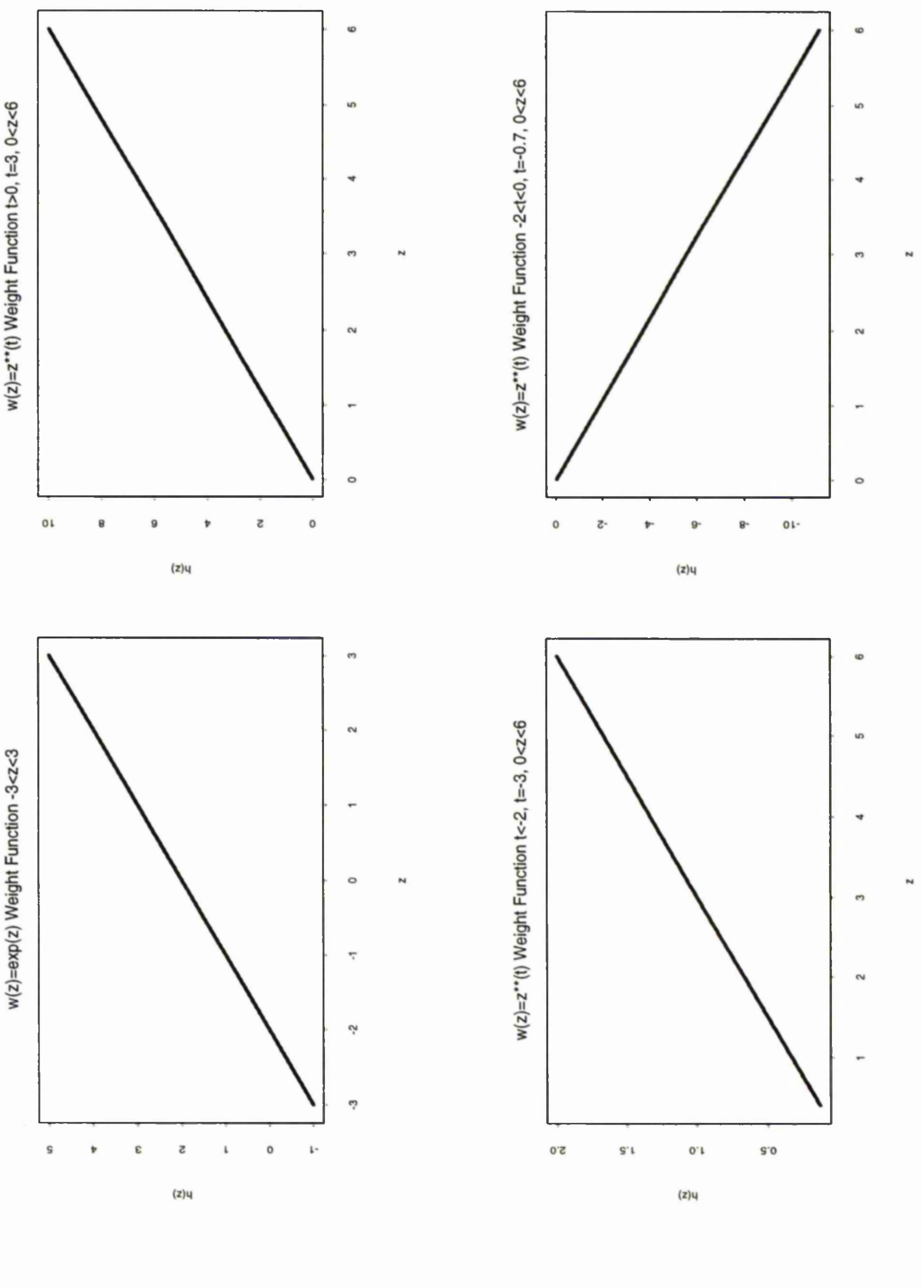


Figure 2.21: Group IV : Plot of $h(z)$ for Non-Binary Weight Functions.

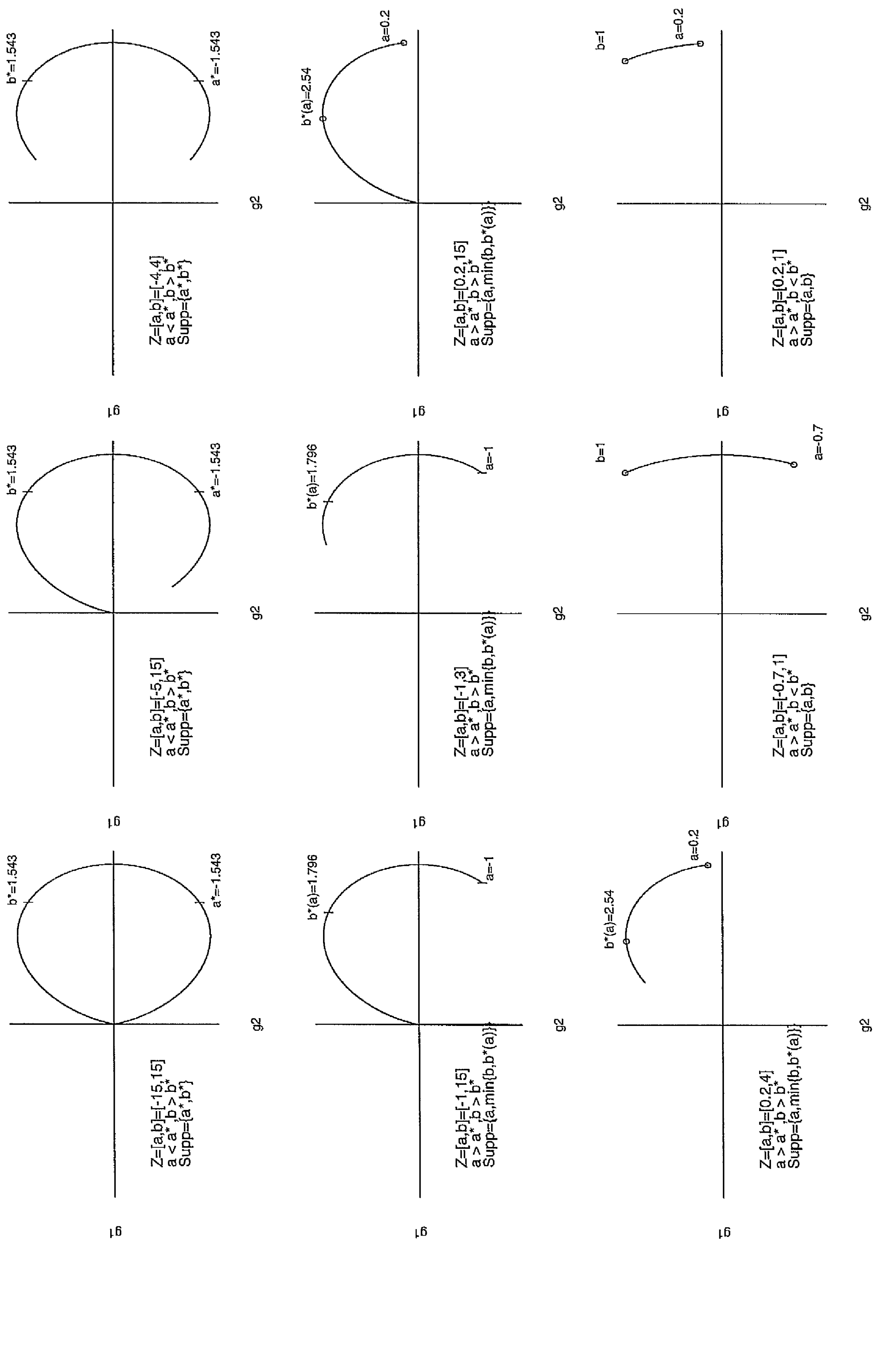


Figure 2.22: Various Plots of $G(\mathcal{Z})$ for Logistic Weight Functions.

Chapter 3

Weighted Regression Model Construction of D-optimal Design : The Case of the Three Parameter Model

3.1 Model under consideration

We consider a binary regression model in which an observed value Y, depends on a vector $\underline{x}$ of 2 design variables $\underline{x} = (x_1, x_2)$ which are selected from a design space $\mathcal{X} \in \mathcal{R}^2$. The outcome Y is binary, i.e., response or non-response, with probabilities

$$\mathbf{Pr}(Y = 0|\underline{x}) = 1 - \pi(\underline{x}) \qquad \mathbf{Pr}(Y = 1|\underline{x}) = \pi(\underline{x}).$$

Thus, $Y \sim Bi(1, \pi(\underline{x}))$. We investigate the relationship between the response probability $\pi(\underline{x})$ and the explanatory or design variables $\underline{x} = (x_1, x_2)$. We assume that $\pi(\underline{x}) = F(\alpha + \beta_1 x_1 + \beta_2 x_2)$, where $F(\cdot)$ is a cumulative distribution, so this is a GLM under which the dependence of

π on $\underline{x} = (x_1, x_2)$ is through the linear function

$$z_1 = \alpha + \beta_1 x_1 + \beta_2 x_2$$

for unknown parameters α, β_1, β_2 . So

$$\begin{aligned} E(Y|\underline{x}) &= \pi(\underline{x}) = F(\alpha + \beta_1 x_1 + \beta_2 x_2) = F(z_1) \\ V(Y|\underline{x}) &= \pi(\underline{x})[1 - \pi(\underline{x})]. \end{aligned}$$

3.2 Design for three parameter Binary regression

We now apply the theory of section (2.2) - (2.4) of Chapter 2 to this problem. The material is similar to that of section (2.4.1). For the above model the information matrix can be written as follows

$$I(x, \underline{\theta}) = \frac{f^2(z_1)}{F(z_1)[1 - F(z_1)]} \begin{pmatrix} 1 \\ x_1 \\ x_2 \end{pmatrix} (1\ , x_1\ , x_2),$$

where $f(z_1) = F'(z_1)$,

$$\begin{aligned} \eta &= \pi(\underline{x}) \\ &= F(\alpha + \beta_1 x_1 + \beta_2 x_2), \qquad z_1 = \alpha + \beta_1 x_1 + \beta_2 x_2 \\ &= F(z_1) \end{aligned}$$

and

$$\begin{aligned} a(\underline{x}, \underline{\theta}) &= \mathbb{V}(Y|\underline{x}) \\ &= \pi(\underline{x})[1 - \pi(\underline{x})] \\ &= F(\alpha + \beta_1 x_1 + \beta_2 x_2)[1 - F(\alpha + \beta_1 x_1 + \beta_2 x_2)] \\ &= F(z_1)[1 - F(z_1)]. \end{aligned}$$

Also

$$\begin{aligned}
\eta_\theta &= \left[\frac{\partial F(z_1)}{\partial z_1}\frac{\partial z_1}{\partial \alpha}, \frac{\partial F(z_1)}{\partial z_1}\frac{\partial z_1}{\partial \beta_1}\frac{\partial F(z_1)}{\partial z_1}\frac{\partial z_1}{\partial \beta_2}\right]^T \\
&= \left[f(z_1), f(z_1)x_1, f(z_1)x_2\right]^T \\
&= f(z_1)\begin{pmatrix} 1 \\ x_1 \\ x_2 \end{pmatrix}.
\end{aligned}$$

Now define the vector

$$\begin{aligned}
\underline{v} &= \frac{1}{\sqrt{V(Y|x)}}\left[\frac{\partial F(z_1)}{\partial z_1}\frac{\partial z_1}{\partial \alpha}, \frac{\partial F(z_1)}{\partial z_1}\frac{\partial z_1}{\partial \beta_1}\frac{\partial F(z_1)}{\partial z_1}\frac{\partial z_1}{\partial \beta_2}\right]^T \\
&= \frac{f(z_1)}{\sqrt{F(z_1)[1-F(z_1)]}}\begin{pmatrix} 1 \\ x_1 \\ x_2 \end{pmatrix}.
\end{aligned}$$

Clearly, z_1 plays a similar role as $z = \alpha + \beta x$ in the two parameter case. It is again convenient to exploit this linear transformation. In the two parameter case we had

$$\begin{aligned}
\binom{1}{z} &= \begin{pmatrix} 1 & 0 \\ \alpha & \beta \end{pmatrix}\binom{1}{x} \\
&= \mathcal{B}\binom{1}{x}.
\end{aligned}$$

We now consider the transformation

$$\begin{aligned}
\begin{pmatrix} 1 \\ z_1 \\ z_2 \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 0 \\ \alpha & \beta_1 & \beta_2 \\ a & b & c \end{pmatrix}\begin{pmatrix} 1 \\ x_1 \\ x_2 \end{pmatrix} \\
&= \mathcal{B}\begin{pmatrix} 1 \\ x_1 \\ x_2 \end{pmatrix}
\end{aligned}$$

where a, b, c are arbitrary constants to be chosen by the experimenter. They define an extra variable z_2. We have transformed to two new design variables z_1, z_2. Their design space will be the image of $\mathcal{X}$ under the transformation. Denote this by $\mathcal{Z}$. Hence

$$\begin{aligned} \underline{g}(\underline{z}) &= \mathcal{B}\underline{v} \\ &= \frac{f(z_1)}{\sqrt{F(z_1)[1-F(z_1)]}} \begin{pmatrix} 1 \\ z_1 \\ z_2 \end{pmatrix} \end{aligned}$$

and

$$\underline{v} = \mathcal{B}^{-1}\underline{g}(\underline{z}) \; .$$

D-optimality is invariant under non-singular linear transformations of the design space. So, as did Ford *et al.* (1992), we consider the D-optimal linear design problem with design vectors

$$\underline{g} = \sqrt{w(z_1)}(1,\ z_1,\ z_2)^T \qquad\qquad (z_1,\ z_2)^T \in \mathcal{Z},$$

where $w(z_1) = \frac{f^2(z_1)}{F(z_1)[1-F(z_1)]}$, which corresponds to a weighted linear regression design problem with weight function $w(z_1)$.

Therefore these nonlinear design problems transform into linear design problems for weighted linear regression in z_1 and z_2 with weight function $w(z_1) = \frac{f^2(z_1)}{F(z_1)(1-F(z_1))}$, where $f(z_1) = F'(z_1)$, the density of $F(\cdot)$. A geometrical approach to the construction of D- optimal designs is useful. A crucial role is played in this by the induced design space

$$G = G(\mathcal{Z}) = \{g_z = (g_1, g_2, g_3)^T : g_1 = \sqrt{w(z_1)}, g_2 = z_1\sqrt{w(z_1)}, g_3 = z_2\sqrt{w(z_1)},\ z \in \mathcal{Z}\}.$$

3.2.1 Characterization of the Optimal Design

Let ξ^* be a design measure on $\mathcal{Z}$. ξ^* is D-optimal iff

$$\begin{aligned} \underline{g}^T(z_1, z_2)\, M^{-1}(\xi^*)\, \underline{g}\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} &\leq 3 \qquad \xi^*(\underline{z}) = 0 \\ &= 3 \qquad \xi^*(\underline{z}) > 0. \end{aligned}$$

D-optimal designs have as support points the points of contact between G and the smallest ellipsoid centred on the origin containing G (Silvey,1980).

Clearly G must be bounded . This will be the case if $\mathcal{X}$ is bounded. However as seen in the 2-parameter case g_1 and g_2 are bounded $\forall\ z_1$ for the weight functions considered. But clearly g_3, and therefore G, will be unbounded if z_2 is unbounded. So bounds are needed on z_2 . Due to the invariance of the D- criterion to linear transformations, without lost of generality we assume $-1 \leq z_2 \leq 1$. This implies $\mathcal{Z} = \mathcal{Z}_w$:

$$\mathcal{Z}_w = \{(z_1, z_2) : -\infty < z_1 < \infty \qquad -1 \leq z_2 \leq 1\}.$$

This is an analogue of $\mathcal{Z}_w$ in the two parameter case. It is the 'largest' possible $\mathcal{Z}$ we can consider.

We first consider optimum designs for this space and later consider optimal designs for certain subsets of it.

- **Case 1 :** $\mathcal{Z} = \mathcal{Z}_w = \{(z_1, z_2) : -\infty < z_1 < \infty, \quad -1 \leq z_2 \leq 1)$.

  We consider $\mathcal{Z} = \mathcal{Z}_w$ initially for the Beta, Gamma, Normal and Binary weight functions. Plots of $G(\mathcal{Z}_w)$ for these weight functions are given in Figure (3.1), Figure (3.2) and Figure (3.3). It is immediately clear that any ellipsoid centred on the origin containing G can only touch G on the upper and lower ridges. Since the support points of the D-optimal design are the points of contact between G and the smallest such ellipsoid we conclude that D-optimal support points lie on these ridges and hence have $z_2 = \pm 1$.

Further, G is symmetric about $g_3 = 0$ ($z_2 = 0$). This leads to the conjecture that D-optimal supports are such that if observations are taken at a particular value of z_1, then these are shared equally between $z_2 = \pm 1$. (Sitter and Torsney, 1995a, 1995b)

- **Case 2 :** $\mathcal{Z} = \{(z_1, z_2) : a \leq z_1 \leq b, \ -1 \leq z_2 \leq 1\}$ $(= Z \subset Z_w)$. We now consider the case $z_1 \in$ [a,b] so that

$$G = G_{ab} = \{\underline{g} \in \mathcal{R}^3 : \underline{g} = \sqrt{w(z_1)}\,(1, z_1, z_2)^T, \ a \leq z_1 \leq b, \ -1 \leq z_2 \leq 1\}.$$

This is the case of a subset of $G(\mathcal{Z}_w)$ which is a 'vertical' (in g_3-direction) portion of $G(\mathcal{Z}_w)$. We will consider other subsets later. Again we can argue that support points can only be on the ridges of G and we can restrict attention to weights equally distributed between $z_2 = \pm 1$, since G is still symmetric about $g_3 = 0$ ($z_2 = 0$). Thus we can restrict attention to the simplified designs considered for G_w with the proviso that the z_1-values must lie in [a,b] .

The next point is : **"how many support points are there?"**. It is well established that (by Caratheodory's theorem) if there are $\boldsymbol{k}$ parameters, a D-optimal design has at least $\boldsymbol{k}$ and at most $\boldsymbol{k(k+1)/2}$ support points. Since there are $\boldsymbol{k = 3}$ parameters, there are at least 3 and at most 6 of them.

Given the above argument that observations taken at z_1 are shared equally between $z_2 = -1$ and $z_2 = 1$ the implication is that there are either 4 or 6 support points. That is, observations are taken at 2 or 3 values z_1 (as in the case of z in the two parameter models). Considerations of the plots of G in Figure (3.1), Figure (3.2) and Figure (3.3) for binary, beta, normal, gamma weight functions suggests that the smallest central ellipsoid will only touch G_{ab} at 4 points, whereas in the case of the Double Reciprocal and Double Exponential weight functions, there are potentially 6 points of contact, including two at $z_1 = 0$. Excepting these two cases we assert that for any a, b there are only four support points and hence observations are taken at only two values of z_1. i.e. the design is of the form

$$\xi = \begin{pmatrix} c & c & d & d \\ -1 & 1 & -1 & 1 \\ p_c & p_c & p_d & p_d \end{pmatrix} \tag{3.1}$$

where $2(p_c + p_d) = 1$.

Let $Supp(\xi^*)$ denote these two z_1-values and let a^*, b^* be their values on $\mathcal{Z}_w$. We further assert that

$$\begin{aligned} Supp(\xi^*) &= \{a^*, b^*\} \qquad a < a^*, b > b^* \\ Supp(\xi^*) &= \{max\{a,\ a^*(b)\}, b\} \qquad a \leq a^*, b \leq b^* \\ Supp(\xi^*) &= \{a, min\{b,\ b^*(a)\}\} \qquad a \geq a^*, b \geq b^* \\ Supp(\xi^*) &= \{a, b\} \qquad a \geq a^*, b \leq b^* \end{aligned}$$

where $b^*(a)$ (along with p_d, p_a) maximises $det(M(\xi))$ with respect to d (over $d \geq a$) where ξ is the design

$$\xi = \begin{pmatrix} a & a & d & d \\ -1 & 1 & -1 & 1 \\ p_a & p_a & p_d & p_d \end{pmatrix},$$

and $a^*(b)$ (along with p_c, p_b) maximises $det(M(\xi))$ with respect to c (over $c \leq b$) where ξ is the design

$$\xi = \begin{pmatrix} c & c & b & b \\ -1 & 1 & -1 & 1 \\ p_c & p_c & p_b & p_b \end{pmatrix}.$$

3.2.2 Justification of the Conjecture

To prove the above conjecture we need to confirm the requirements of the Equivalence Theorem. This requires that the following necessary and sufficient conditions (Kiefer and Wolfwitz, 1960) must be satisfied by an arbitrary design $\xi(z_1,\, z_2)$ if it is to D-optimal.

$$w(z_1)(1,\ z_1,\ z_2)M^{-1}(\xi)\begin{pmatrix}1\\ z_1\\ z_2\end{pmatrix} \le 3 \quad \forall \quad z_1,\, z_2 \in \mathcal{Z} \tag{3.2}$$

$$= 3 \quad \text{if} \quad \xi(z_1,\, z_2) > 0 \tag{3.3}$$

We only need to check this for $z_2 = \pm 1$ and all relevant z_1, in which case equations (3.2) and (3.3) imply

$$\begin{aligned} w(z_1)Q^{\times}(z_1) &\le 3 \quad \forall \quad (z_1, \pm 1) \in \mathcal{Z} \\ &= 3 \quad \text{if} \quad \xi(z_1,\, \pm 1) > 0\ , \end{aligned}$$

where $Q^{\times}(z_1) = (1,\ z_1,\ \pm 1)M^{-1}(\xi)(1,\ z_1,\ \pm 1)^T$, a quadratic function. That is

$$\begin{aligned} v^{\times}(z_1) = Q^{\times}(z_1) - \frac{3}{w(z_1)} &\le 0 \quad \forall \quad (z_1,\, \pm 1) \in \mathcal{Z} \\ &= 0 \quad \text{if} \quad \xi(z_1,\, \pm 1) > 0. \end{aligned}$$

So for an optimal design we wish to see $v^{\times}(z_1) \le 0$ in the case $\mathcal{Z} = \{(z_1 z_2) : a \le z_1 \le b, -1 \le z_2 \le 1\}$. To explore the shape of $v^{\times}(z_1)$, we analyze the derivative of $v^{\times}(z_1)$. This can be written as follows :

$$\frac{\mathbf{d}v^{\times}(z_1)}{\mathbf{d}z_1} = L(z_1) - H_3^{\times}(z_1), \tag{3.4}$$

where $H_3^{\times}(z_1) = \dfrac{-\,3\, w'(z_1)}{[w(z_1)]^2}$ and $L(z_1)$ is an increasing linear function of z_1 because the coefficient of z_1 is the value of the second diagonal element of the inverse of the design matrix $M(\xi)$ which is positive definitive.

The consequence is that $\frac{\mathrm{d}v^{\times}(z_1)}{\mathrm{d}z_1} = 0$ iff $L(z_1) = H_3^{\times}(z_1)$. That is, $\frac{\mathrm{d}v^{\times}(z_1)}{\mathrm{d}z_1} = 0$ when the line $L(z_1)$ crosses $H_3^{\times}(z_1)$.

The striking point is that $H_3^{\times}(z_1) \propto H(z_1)$ [Chapter 2, equation (2.3)]. There is no difference in the shapes of these functions. Thus $L(z_1)$ can only cut $H_3^{\times}(z_1)$ at most three times as was the case for most of our weight functions in the two parameter case.

Therefore we have the same conclusion here (for most of weight functions considered): namely, $H_3^{\times}(-\infty) = -\infty$, $H_3^{\times}(+\infty) = +\infty$ and $H_3^{\times}(z_1)$ is concave increasing up to some point and thereafter convex increasing.

It follows that $v^{\times}(z_1)$ has at most 3 turning points at $\mathcal{Z}_w$. Because $L(z_1)$ first crosses $H_3^{\times}(z_1)$ from above, $v^{\times}(z_1)$ has only one minimum turning point for the same reasons as before. Hence for these weight functions there are only two support points along each horizontal edge identified by two distinct values of z_1 with the weight at these shared equally between $z_2 = \pm 1$. These give a total of 4 support points. We now need to determine the two values of z_1 and the optimal weights. In fact there is an explicit solution for these weights.

3.2.3 Definition of weights

We consider the specific design

$$p = \begin{pmatrix} i & 1 & 2 & 3 & 4 \\ z_{1i} & c & c & d & d \\ z_{2i} & -1 & 1 & -1 & 1 \\ p_i & p_c & p_c & p_d & p_d \end{pmatrix},$$

where p_c, $p_d > 0$ and $2(p_c + p_d) = 1$. The design matrix is

$$M(p) = \Sigma_i^4 p_i \underline{g}_i \underline{g}_i^T,$$

where

$$\underline{g}_i = \sqrt{w(z_{1i})}(1, z_{1i}, z_{2i})^T \qquad i = 1, 2, 3, 4.$$

Therefore,

$$M(p) = \begin{pmatrix} 2p_c w(c) + 2p_d w(d) & 2cp_c w(c) + 2dp_d w(d) & 0 \\ 2cp_c w(c) + 2dp_d w(d) & 2c^2 p_c w(c) + 2d^2 p_d w(d) & 0 \\ 0 & 0 & 2p_c w(c) + 2p_d w(b) \end{pmatrix}.$$

The determinant is

$$|M(p)| = 2^3 (d-c)^2 \, p_c \, w(c) \, p_d \, w(b) \, [p_c \, w(c) + p_d \, w(d)].$$

We need to choose c, d, p_c, p_d to maximize the determinant of the design matrix $|M(p)|$. We can find an explicit solution for the weights: First, we get the **log** of the ***determinant*** function which is a concave function of $M(\cdot)$ and substitute $p_d = \frac{1}{2} - p_c$.

$$\begin{aligned} \ln|M(p)| &= 2\ln 3 + 2\ln(d-c) + \ln p_c + \ln(\frac{1}{2} - p_c) + \ln w(c) + \ln w(d) \\ &\quad + \ln[p_c w(c) + (\frac{1}{2} - p_c) w(d)] \end{aligned}$$

$$\begin{aligned} \frac{\partial \ln|M(p)|}{\partial p_c} &= \frac{1}{p_c} - \frac{1}{(\frac{1}{2} - p_c)} + \frac{w(c) - w(d)}{p_c w(c) + (\frac{1}{2} - p_c) w(d)} \\ &= \frac{1 - 4p_c}{p_c(1 - 2p_c)} + \frac{w(c) - w(d)}{p_c w(c) + (\frac{1}{2} - p_c) w(d)} \end{aligned}$$

Further,

$$\frac{\partial \ln|M(p)|}{\partial p_c} = 0 \qquad if \qquad \frac{1 - 4p_c}{p_c(1 - 2p_c)} + \frac{w(c) - w(d)}{p_c w(c) + (\frac{1}{2} - p_c) w(d)} = 0\,.$$

$$3p_c^2[w(c) - w(d)] - p_c[w(c) - 2w(d)] - \frac{1}{4} w(b) = 0 \qquad (3.5)$$

$$p_c = \frac{[w(c) - 2w(d)] \pm \sqrt{[w(c) - 2w(d)]^2 + 3[w(c) - w(d)] w(d)}}{6[w(c) - w(d)]}.$$

This is an explicit solution for the values of p_c that maximize $|M(p)|$. Of the above two roots, our solution is given by the first root[1] :

$$p_c = \frac{[w(c)-2w(d)]+\sqrt{[w(c)-2w(d)]^2+3[w(c)-w(d)]w(d)}}{6[w(c)-w(d)]}. \quad (3.6)$$

Hence $p_d = \frac{1}{2} - p_c$.

Further we can express the solution for p_c in terms of $r = \dfrac{w(c)}{w(d)}$, namely:

$$p_c = q_3(r) = \frac{(r-2)+\sqrt{(r-2)^2+3(r-1)}}{6(r-1)}. \quad (3.7)$$

3.2.4 Determination of support points

Still the design is

$$\begin{pmatrix} i & 1 & 2 & 3 & 4 \\ z_1 & c & c & d & d \\ z_2 & -1 & 1 & -1 & 1 \\ p & p_c & p_c & p_d & p_d \end{pmatrix}$$

and

$$\begin{aligned} \ln[detM(p)] &= 3\ln 2 + 2\ln(d-c) + \ln p_c + \ln(p_d) \\ &\quad + \ln w(c) + \ln w(d) + \ln[p_c w(c) + (p_d)w(d)], \qquad c < d. \end{aligned} \quad (3.8)$$

We now view this as a function of four sub-functions of c, namely $w(c)$, p_c, p_d (since p_d is a function of c through the condition $p_d + p_c\frac{1}{2}$) and $A(c,d) = 3\ln 2 + 2\ln(d-c)$, so that

$$\begin{aligned} \ln[detM(p)] &= A(c,d) + \ln p_c + \ln(p_d) + \ln w(c) + \ln w(d) \\ &\quad + \ln[p_c w(c) + (p_d)w(d)], \qquad c < d & (3.9) \\ &= F(A(c,d), w(c), w(d), p_c, p_d) & (3.10) \\ &= F & (3.11) \end{aligned}$$

[1] In fact, the use of the second root leads to negative weights which obviously violates the constraints on the weights.

Note that here we have not substituted for p_d in terms of p_c. If we do not make this substitution we need to use a Lagrangian approach to determine the optimal values of p_c, p_d. Some useful formulae emerge if we do this. Since $p_c + p_d = \frac{1}{2}$ the Lagrangian is

$$L(p_c, p_d, \lambda) = F - \lambda(p_c + p_d - 1/2).$$

Having formed our total objective function we now determine the partial derivatives of $L(p_c, p_d, \lambda)$ with respect to p_c, p_d and λ respectively.

$$\begin{aligned} \frac{\partial L(p_c, p_d, \lambda)}{\partial p_c} &= \frac{\partial F}{\partial p_c} - \lambda \\ \frac{\partial L(p_c, p_d, \lambda)}{\partial p_d} &= \frac{\partial F}{\partial p_d} - \lambda \\ \frac{\partial L(p_c, p_d, \lambda)}{\partial \lambda} &= -(p_c + p_d - 1/2) \end{aligned}$$

Hence

$$\left\{ \begin{array}{l} \frac{\partial L(p_c, p_d, \lambda)}{\partial p_c} = 0 \\ \frac{\partial L(p_c, p_d, \lambda)}{\partial p_d} = 0 \end{array} \right. \iff \left\{ \begin{array}{l} \frac{\partial F}{\partial p_c} = \lambda \\ \frac{\partial F}{\partial p_d} = \lambda \end{array} \right. \tag{3.12}$$

To determine λ we note

$$\begin{aligned} p_c \frac{\partial F}{\partial p_c} + p_d \frac{\partial F}{\partial p_d} &= \lambda(p_c + p_d) \\ &= \frac{1}{2}\lambda. \end{aligned}$$

Consequently,

$$\lambda = 2\Big[p_c \frac{\partial F}{\partial p_c} + p_d \frac{\partial F}{\partial p_d}\Big] \tag{3.13}$$

Now

$$\frac{\partial F}{\partial p_c} = \frac{1}{p_c} + \frac{w(c)}{[p_c w(c) + p_d w(d)]} \tag{3.14}$$

$$\frac{\partial F}{\partial p_d} = \frac{1}{p_d} + \frac{w(d)}{[p_c w(c) + p_d w(d)]} \tag{3.15}$$

Multiplying equations (3.14) and (3.15) by p_c p_d respectively, and summing the resulting equations we can write

$$p_c \frac{\partial F}{\partial p_c} + p_d \frac{\partial F}{\partial p_d} = 1 + \frac{p_c w(c)}{[p_c w(c) + p_d w(d)]} + 1 + \frac{p_d w(d)}{[p_c w(c) + p_d w(d)]} = 3$$

which is constant. And from equation (3.13), $\lambda = 6$.

Further

$$\begin{aligned} \frac{\partial F}{\partial p_c} &= \frac{1}{p_c} + \frac{w(c)}{p_c w(c) + p_d w(d)} = 6 \\ &\Rightarrow 1 + \frac{p_c w(c)}{p_c w(c) + p_d w(d)} = 6p_c \\ &\Rightarrow \frac{p_c w(c)}{p_c w(c) + p_d w(d)} = 6p_c - 1\,. \end{aligned} \tag{3.16}$$

Similarly,

$$\begin{aligned} \frac{\partial F}{\partial p_d} &= \frac{1}{p_d} + \frac{w(d)}{p_c w(c) + p_d w(d)} = 6 \\ &\Rightarrow 1 + \frac{w(d) p_d}{p_c w(c) + p_d w(d)} = 6p_d \\ &\Rightarrow \frac{w(d) p_d}{p_c w(c) + p_d w(d)} = 6p_d - 1\,. \end{aligned} \tag{3.17}$$

We note that we will be interested in derivatives of this function with respect to c and or d. To find the best four-point design on $\mathcal{Z}_w$ we need to maximise $\ln[detM(p)]$ w.r.t. c and d or if we wish to find the best four point design subject to c (or d) being a support point we need to maximise F w.r.t. d (or c).

$$\frac{\partial F}{\partial c} = \frac{\partial F}{\partial A(c,d)}\frac{\partial A(c,d)}{\partial c} + \frac{\partial F}{\partial w(c)}\frac{\partial w(c)}{\partial c} + \frac{\partial F}{\partial p_c}\frac{\partial p_c}{\partial c} + \frac{\partial F}{\partial p_d}\frac{\partial p_d}{\partial c} \tag{3.18}$$

Now we can substitute the values from equations (3.16) and (3.17) into equation (3.18) to obtain the following :

$$\frac{\partial F}{\partial c} = \frac{\partial F}{\partial A(c,d)}\frac{\partial A(c,d)}{\partial c} + \frac{\partial F}{\partial w(c)}\frac{\partial w(c)}{\partial c} + 6\frac{\partial p_c}{\partial c} + 6\frac{\partial p_d}{\partial c} \tag{3.19}$$

From the definition of p_c, p_d ($p_c + p_d = 1/2$), we can write the following

$$\begin{aligned} \frac{\partial p_c}{\partial c} + \frac{\partial p_d}{\partial c} &= 0 \\ \frac{\partial p_c}{\partial c} &= -\frac{\partial p_d}{\partial c} \end{aligned}$$

Therefore,

$$\begin{aligned}
\frac{\partial F}{\partial c} &= \frac{\partial F}{\partial A(c,d)}\frac{\partial A(c,d)}{\partial c} + \frac{\partial F}{\partial w(c)}\frac{\partial w(c)}{\partial c} + \underbrace{\frac{\partial p_c}{\partial c}[6-6]}_{=0} \\
&= \frac{\partial F}{\partial A(c,d)}\frac{\partial A(c,d)}{\partial c} + \frac{\partial F}{\partial w(c)}\frac{\partial w(c)}{\partial c} \\
&= \frac{-2}{d-c} + \frac{w'(c)}{w(c)} + \frac{p_c w'(c)}{p_c w(c) + p_d w(d)} \\
&= \frac{-2}{d-c} + \frac{w'(c)}{w(c)}\left[1 + \frac{p_c w(c)}{p_c w(c) + p_d w(d)}\right] \\
&= \frac{-2}{d-c} + \frac{6 p_c w'(c)}{w(c)} \\
&= \frac{-2w(c) + (p_c w'(c)\, 6)\,(d-c)}{w(c)(d-c)} \\
&= \frac{w'(c)\, 6\, p_c}{w(c)(d-c)}\left[(d-c) - \frac{w(c)}{3 p_c w'(c)}\right] \quad if \quad w'(c) \neq 0 \\
&= \frac{p_c w'(c)\, 6}{w(c)(d-c)}\Big[d - h_d(c)\Big]
\end{aligned}$$

where $\quad h_d(c) = c + \dfrac{w(c)}{3\, p_c w'(c)}$. So $\dfrac{\partial[\ln det M(\xi)]}{\partial z_1} \propto p_c w'(z)\, 6\big[d - h_d(c)\big]$.

Further,

$$\frac{\partial[\ln det M(p)]}{\partial c} = 0 \quad if \quad \Rightarrow \quad \frac{p_c w'(c)\, 6}{w(c)(d-c)}\left[(d-c) - \frac{2w(c)}{6\, p_c w'(c)}\right] = 0$$

$$i.e. \quad if \quad d = h_d(c) \quad \big[given \quad w'(c) \neq 0\big].$$

(Note, if $p_c w'(c) = 0$ then $\dfrac{\partial[\ln det M(p)]}{\partial c} = \dfrac{-2}{d-c} \neq 0$. Since $w'(z_{max}) = 0$, so z_{max} is not a solution of $\dfrac{\partial[\ln det M(p)]}{\partial c} = 0$, where z_{max} is the value of z_1 which maximises $w(z_1)$.)

Similarly,

$$\frac{\partial F}{\partial d} = \frac{\partial F}{\partial A(c,d)}\frac{\partial A(c,d)}{\partial d} + \frac{\partial F}{\partial w(d)}\frac{\partial w(d)}{\partial d} + \frac{\partial F}{\partial p_c}\frac{\partial p_c}{\partial d} + \frac{\partial F}{\partial p_d}\frac{\partial p_d}{\partial d} \tag{3.20}$$

Now we can substitute the values from equations (3.16) and (3.17) into equation (3.20) to obtain the following :

$$\frac{\partial F}{\partial d} = \frac{\partial F}{\partial A(c,d)}\frac{\partial A(c,d)}{\partial d} + \frac{\partial F}{\partial w(d)}\frac{\partial w(d)}{\partial d} + 6\frac{\partial p_c}{\partial d} + 6\frac{\partial p_d}{\partial d} \tag{3.21}$$

From the definition of p_c, p_d ($p_c + p_d = 1/2$), we can write the following

$$\begin{aligned}\frac{\partial p_d}{\partial d} + \frac{\partial p_d}{\partial d} &= 0 \\ \frac{\partial p_d}{\partial d} &= -\frac{\partial p_d}{\partial d}\end{aligned}$$

Therefore,

$$\begin{aligned}\frac{\partial F}{\partial d} &= \frac{\partial F}{\partial A(c,d)}\frac{\partial A(c,d)}{\partial d} + \frac{\partial F}{\partial w(d)}\frac{\partial w(d)}{\partial d} + \underbrace{\frac{\partial p_d}{\partial d}[6-6]}_{=0} \\ &= \frac{\partial F}{\partial A(c,d)}\frac{\partial A(c,d)}{\partial d} + \frac{\partial F}{\partial w(d)}\frac{\partial w(d)}{\partial d} \\ &= \frac{2}{d-c} + \frac{w'(d)}{w(d)} + \frac{p_d w'(d)}{p_c w(c) + p_d w(d)} \\ &= \frac{2}{d-c} + \frac{w'(d)}{w(d)}\left[1 + \frac{p_c w(c)}{p_c w(c) + p_d w(d)}\right] \\ &= \frac{2}{d-c} + \frac{6 p_d w'(d)}{w(d)} \\ &= \frac{2w(d) + (p_d w'(d)\,6)\,(d-c)}{w(d)(d-c)} \\ &= \frac{p_d w'(d)\,6}{w(d)(d-c)}\left[(d-c) + \frac{w(d)}{3 p_d w'(d)}\right] \quad if \quad w'(d) \neq 0 \\ &= \frac{p_d w'(d)\,6}{w(d)(d-c)}\left[h_c(d) - c\right]\end{aligned}$$

where $\quad h_c(d) = d + \dfrac{w(d)}{3\,p_d w'(d)}$. So $\quad \dfrac{\partial[\ln det M(\xi)]}{\partial z_1} \propto p_d w'(z)\,6\left[h_c(d) - c\right]$

Further,

$$\frac{\partial[\ln det M(p)]}{\partial d} = 0 \quad if \quad \Rightarrow \quad \frac{p_d w'(d)\,6}{w(d)(d-c)}\left[(d-c) + \frac{w(d)}{3\,p_d w'(d)}\right] = 0$$

$$i.e. \quad if \quad d = h_c(d) \quad \left[given \quad w'(d) \neq 0\right].$$

(Note, if $p_d w'(d) = 0$ then $\dfrac{\partial[\ln det M(p)]}{\partial d} = \dfrac{2}{d-c} \neq 0$. So $(z)_{max}$ is not a solution of $\dfrac{\partial[\ln det M(p)]}{\partial d} = 0$, where z_{max} is the value of z_1 which maximises $w(z_1)$.)

As a result of this, we can be interested in solving one or both of the equations

$$c = h_c(d) \tag{3.22}$$

$$h_d(c) = d \tag{3.23}$$

Clearly, the function $h_c(d)$, $h_d(c)$ play the role of $h(z)$ in the two parameter case but that has now been replaced by a class of functions. It is useful to study $h_t(z_1)$. The solution to these equations clearly depends on the nature of $h_t(z_1)$. We consider the same weight functions as in chapter two. For those that are unimodal and stationary at their maximum the class of functions $h_t(z_1)$ are increasing both over $z_1 \leq z_{max}$ and over $z_1 \geq z_{max}$ with a vertical asymptote at z_{max}. Plots in the case of the weight function for binary logistic regression are shown in Figure (3.6). Further plots are revealed in Chapter 5, Figures (5.2), (5.3), (5.4) and (5.5). This again is useful to us.

Now consider the single equation in z_1

$$h_{z_2}(z_1) \;=\; e.$$

As in the previous chapter there is one solution to this equation say $z_1 = z_L^*(e)$ in the range $z_1 \leq z_{max}$ and one, say $z_1 = z_U^*(e)$, in the range $z_1 \geq z_{max}$. Moreover since $w'(z_L^*(e)) > 0$ and $w'(z_U^*(e)) < 0$ we have $z_L^*(e) < e < z_U^*(e)$. In equations (3.22) and (3.23) we have two versions of the above. Their joint solution, with $z_1 < z_2$, must be $z_1^* = a^*$, $z_2^* = b^*$, $a^* < b^*$, a^*, b^* being the support points of the optimal four-point design on $\mathcal{Z}_w$ as defined in the conjectures above. Note this means

$$h(a^*) \;=\; b^*, \qquad h(b^*) \;=\; a^* \qquad \textit{and} \quad z_1^* = z_L^*(z_2^*), \quad z_2^* = z_U^*(z_1^*)$$

3.2.5 Examination of the conjectures against the Equivalence Theorem

We can now consider checking these conjectures against the Equivalence Theorem. Consider an arbitrary four point design:

$$p = \begin{pmatrix} i & 1 & 2 & 3 & 4 \\ z_{1i} & b_1 & b_1 & b_2 & b_2 \\ z_{2i} & -1 & 1 & -1 & 1 \\ p_i & p(b_1) & p(b_1) & p(b_2) & p(b_2) \end{pmatrix}$$

where $p(b_1)$, $p(b_2) > 0$ are the optimal weights for b_1, b_2 respectively. The information matrix is

$$M(p) = \Sigma_i^4 p_i \underline{g}_i \underline{g}_i^T$$

where

$$\underline{g}_i = \sqrt{w(z_{1i})}(1, z_{1i}, z_{2i})^T \qquad i = 1, 2, 3, 4$$

Therefore,

$$M(p) = \left(\begin{array}{cc|c} 2p(b_1)w(b_1) + 2p(b_2)w(b_2) & 2b_1p(b_1)w(b_1) + 2b_2p(b_2)w(b_2) & 0 \\ 2b_1p(b_1)w(b_1) + 2b_2p(b_2)w(b_2) & 2b_1^2p(b_1)w(b_1) + 2b_2^2p(b_2)w(b_2) & 0 \\ \hline 0 & 0 & 2p(b_1)w(b_1) + 2p(b_2)w(b_2) \end{array}\right)$$

$$M(p) = 2\left(\begin{array}{cc|c} p(b_1)w(b_1) + p(b_2)w(b_2) & b_1p(b_1)w(b_1) + b_2p(b_2)w(b_2) & 0 \\ b_1p(b_1)w(b_1) + b_2p(b_2)w(b_2) & b_1^2p(b_1)w(b_1) + b_2^2p(b_2)w(b_2) & 0 \\ \hline 0 & 0 & p(b_1)w(b_1) + p(b_2)w(b_2) \end{array}\right) \quad (3.24)$$

Let the partition of $M(p)$ in Equation (3.24) be represented by

$$M(p) = 2\begin{pmatrix} \mathcal{S}_0 & 0 \\ 0 & \mathcal{S}_2 \end{pmatrix}$$

where $\mathcal{S}_2 = p(b_1)w(b_1) + p(b_2)w(b_2)$ and $\mathcal{S}_0$ is

$$\mathcal{S}_0 = \begin{pmatrix} S_{011} & S_{012} \\ S_{021} & S_{022} \end{pmatrix}$$

where $\mathcal{S}_{011} = p(b_1)w(b_1) + p(b_2)w(b_2)$, $\mathcal{S}_{012} = b_1p(b_1)w(b_1) + b_2p(b_2)w(b_2)$ and $\mathcal{S}_{022} = b_1^2b(b_1)w(b_1) + b_2^2p(b_2)w(b_2)$. Therefore the inverse of the design matrix is

$$M^{-1}(p) = \frac{1}{2}\begin{pmatrix} \mathcal{S}_0^{-1} & 0 \\ 0 & \mathcal{S}_2^{-1} \end{pmatrix},$$

from the definition of $\mathcal{S}_0$,

$$\mathcal{S}_0^{-1} = \begin{pmatrix} \frac{S_{022}}{|\mathcal{S}_0|} & -\frac{S_{012}}{|\mathcal{S}_0|} \\ \frac{-S_{021}}{|\mathcal{S}_0|} & \frac{S_{011}}{|\mathcal{S}_0|} \end{pmatrix}$$

which yields,

$$M^{-1}(p) = \frac{1}{2}\begin{pmatrix} \frac{S_{022}}{|\mathcal{S}_0|} & -\frac{S_{012}}{|\mathcal{S}_0|} & 0 \\ -\frac{S_{021}}{|\mathcal{S}_0|} & \frac{S_{011}}{|\mathcal{S}_0|} & 0 \\ 0 & 0 & \mathcal{S}_2^{-1} \end{pmatrix}.$$

If the above design is to be D-optimal on a set $\mathcal{Z}$ of values of z_1 for $z_2 = \pm 1$, then, as noted in section (3.2.2), we must have

$$v^{\times}(z_1) \leq 0 \qquad \forall\, z_1 \in \mathcal{Z}$$

where

$$v^{\times}(z_1) = Q^{\times}(z_1) - \frac{3}{w(z_1)}$$

and now

$$
\begin{aligned}
Q^{\times}(z_1) &= \left(\frac{1}{2}\right)(1,\, z_1\, \pm 1)\begin{pmatrix} \frac{S_{022}}{|S_0|} & -\frac{S_{012}}{|S_0|} & 0 \\ -\frac{S_{021}}{|S_0|} & \frac{S_{011}}{|S_0|} & 0 \\ 0 & 0 & S_2^{-1} \end{pmatrix}\begin{pmatrix} 1 \\ z_1 \\ \pm 1 \end{pmatrix} \\
&= \left\{\frac{1}{2}\right\}\left\{\frac{S_{022} - 2z_1 S_{012} + z_1^2 S_{011}}{|S_0|} + \frac{1}{S_2}\right\} \\
&= \left\{\frac{1}{2}\right\}\left\{\frac{1}{p(b_1)p_(b_2)w(b_1)w(b_2)(b_2-b_1)^2}\Big\{[b_1^2 p(b_1)w(b_1) + b_2^2 p(b_2)w(b_2)]\right. \\
&\quad -2\underbrace{b_1 b_2}_{z_1}[b_1 p(b_1)w(b_1) + b_2 p(b_2)w(b_2)] + \underbrace{b_1^2 b_2^2}_{z_1^2}[p(b_1)w(b_1) + p(b_2)w(b_2)]\Big\} \\
&\quad \left. + \left[\frac{1}{p(b_1)w(b_1) + p(b_2)w(b_2)}\right]\right\}
\end{aligned}
$$

Equivalently we must have

$$v(z_1) \;\leq\; 0$$

where $v(z_1) = Q(z_1) - \dfrac{6}{w(z_1)}$ with $Q(z_1) = 2Q^{\times}(z_1)$. In fact $v(z_1)$ must be maximised at b_1, b_2 over $\mathcal{Z}$, a maximum of zero since $v(b_1) = v(b_2) = 0$. So it is convenient to consider the derivative of $v(z_1)$ at b_1, b_2. We recall that

$$
\begin{aligned}
v'(z_1) &= Q'(z_1) + \frac{6w'(z_1)}{[w(z_1)]^2} \\
&= L_3(z_1) - H_3(z_1)
\end{aligned}
$$

where or $L_3(z_1) = Q'(z_1)$ and $H_3(z_1) = \dfrac{-6w'(z_1)}{[w(z_1)]^2}$. Now we can explore $L_3(z_1)$ as follows:

$$
\begin{aligned}
L_3(z_1) &= Q'(z_1) \\
&= \frac{-2S_{021} + 2z_1 S_{011}}{|S_0|} \\
&= \frac{1}{p(b_1)p_(b_2)w(b_1)w(b_2)(b_2-b_1)^2}\Big(-2[b_1 p(b_1)w(b_1) + b_2 p(b_2)w(b_2)] + 2z_1[p(b_1)w(b_1) + p(b_2)w(b_2)]\Big).
\end{aligned}
$$

Therefore $L(b_1)$ and $L(b_2)$ can be written as follows:

$$\begin{aligned}
L_3(b_1) &= \frac{1}{p(b_1)p(b_2)w(b_1)w(b_2)(b_2-b_1)^2}(2b_1[p(b_1)w(b_1)+p(b_2)w(b_2)]-2[b_1p(b_1)w(b_1)+b_2p(b_2)w(b_2)]) \\
&= \frac{2}{p(b_1)p(b_2)w(b_1)w(b_2)(b_2-b_1)^2}[(b_1-b_2)p(b_2)w(b_2)] \\
&= \frac{-2}{p(b_1)w(b_1)(b_2-b_1)}. \\
L_3(b_2) &= \frac{1}{p(b_1)p(b_2)w(b_1)w(b_2)(b_2-b_1)^2}(2b_2[p(b_1)w(b_1)+p(b_2)w(b_2)]-2[b_1p(b_1)w(b_1)+b_2p(b_2)w(b_2)]) \\
&= \frac{2}{p(b_1)p(b_2)w(b_1)w(b_2)(b_2-b_1)^2}[(b_2-b_1)p(b_1)w(b_1)] \\
&= \frac{2}{p(b_2)w(b_2)(b_2-b_1)}.
\end{aligned}$$

We reached the same result in chapter 2.

So

$$\begin{aligned}
v'(b_1) &= L_3(b_1) - H_3(b_1) \\
&= L_3(b_1) + \frac{6w'(b_1)}{[w(b_1)]^2} \\
&= \frac{-2}{p(b_1)w(b_1)(b_2-b_1)} + \frac{6w'(b_1)}{[w(b_1)]^2} \\
&= \frac{1}{w(b_1)}\Big[\frac{-2}{p(b_1)(b_2-b_1)} + \frac{6w'(b_1)}{w(b_1)}\Big] \\
&= \frac{6w'(b_1)}{[w(b_1)]^2(b_2-b_1)}\Big[\frac{-2w(b_1)}{6p(b_1)w'(b_1)} + (b_2-b_1)\Big] \\
&= \frac{6w'(b_1)}{[w(b_1)]^2(b_2-b_1)}\Big[b_2 - [b_1 + \frac{2w(b_1)}{6p(b_1)w'(b_1)}]\Big] \\
&= \frac{6w'(b_1)}{[w(b_1)]^2(b_2-b_1)}[b_2 - h_{b_2}(b_1)].
\end{aligned}$$

Similarly,

$$\begin{aligned}
v'(b_2) &= L_3(b_2) - H_3(b_2) \\
&= L_3(b_2) + \frac{6w'(b_2)}{[w(b_2)]^2} \\
&= \frac{2}{p(b_2)w(b_2)(b_2-b_1)} + \frac{6w'(b_2)}{[w(b_2)]^2} \\
&= \frac{1}{w(b_2)}\Big[\frac{2}{p(b_2)(b_2-b_1)} + \frac{6w'(b_2)}{w(b_2)}\Big] \\
&= \frac{6w'(b_2)}{[w(b_2)]^2(b_2-b_1)}\Big[\frac{2w(b_2)}{6p(b_2)w'(b_2)} + (b_2-b_1)\Big] \\
&= \frac{6w'(b_2)}{[w(b_2)]^2(b_2-b_1)}\Big[[b_2 + \frac{2w(b_2)}{6p(b_2)w'(b_2)}] - b_1\Big] \\
&= \frac{6w'(b_2)}{[w(b_2)]^2(b_2-b_1)}[h_{b_1}(b_2) - b_1].
\end{aligned}$$

Therefore

$$\begin{aligned} v'(b_1) &\propto w'(b_1)[b_2 - h_{b_2}(b_1)] \\ v'(b_2) &\propto w'(b_2)[h_{b_1}(b_2) - b_1] \end{aligned}$$

So the signs of $v'(b_i)$ $i = 1, 2$ depend on the signs of $w'(b_i)$ $i = 1, 2$ and $[b_2 - h_{b_2}(b_1)]$, $[h_{b_1}(b_2) - b_1]$.

3.2.6 Two Conditions for the Conjecture

We can prove our conjecture about the support points if the following two conditions are satisfied:

(i) $h_u(z_1)$ is increasing in z_1 over $z_1 \leq z_{max}$ and $z_1 \geq z_{max}$ for each u.

(ii) For all $z_1 \leq z_{max}$, $h_u(z_1)$ decreases in u over $u \geq z_{max}$, and for all $z_1 \geq z_{max}$ $h_u(z_1)$ decreases in u over $u \leq z_{max}$. We only need to consider these cases since we are considering desugn intervals contaning z_{max}.

We can only provide empirical evidence in support of (i). In figure (3.6) we show plots of $h_u(z_1)$ for the Logistic weight function for a range of values of u. It can be seen that these functions are all increasing. It is also evident that (ii) is true. However, we can provide analytical proof for this.

Proof of condition (ii). To make our proof easy, we write $h_u(z_1)$ in the following form :

$$h_u(z_1) \quad = \quad z_1 + \frac{w(z_1)}{3q_3(r)w'(z_1)}, \qquad r = \frac{w(z_1)}{w(u)}$$

where $q_3(r)$ is the expression encountered in equation (3.7).

From the above expression of $h_u(z_1)$, proving that $q_3(r)$ is increasing would be sufficient to establish that $h_u(z_1)$ is decreasing. In appendix A we prove that $q_3(r)$ is increasing in r.

Thus if $w'(z_1) > 0$, i.e. $z_1 < z_{max}$, $h_u(z_1)$ is decreasing in r, while if $z_1 > z_{max}$ it is increasing in r.

Finally, we note that $w(u)$ is increasing in u over $u < z_{max}$ and decreasing over $u > z_{max}$. Hence $r = \dfrac{w(z_1)}{w(u)}$ is decreasing in u over $z_1 < z_{max}$ and increasing in u over $z_1 > z_{max}$. Hence (ii). □

3.2.7 Confirmation of D-Optimality

Now we consider u_1, u_2 to be the two distinct values of z_1 which produce the support points of the conjectured optimal designs of the various cases of $[a, b]$. Our essential aim is to verify that $v(z_1) \leq 0$ on $[a, b]$. The properties of $v(z)$ [see (Chapter 2, Section (2.2.4))] confirm that $v(z_1) \leq 0$ on $[a, b]$ if $v'(u_1) \leq 0$ and $v'(u_2) \geq 0$.

The confirmation of D-optimality is similar that of the two parameter case. However, it is worth to detaling some of the cases:

We now establish results confirming $v'(u_1) \leq 0$, $v'(u_2) \geq 0$ as appropriate under two assumptions. We assume that

(i-) $h_u(z)$ is increasing in z over $z \leq z_{max}$ and $z \geq z_{max}$ for each u.

(ii-) For all $z \leq z_{max}$, $h_u(z)$ decreases in u over $u \geq z_{max}$ and for all $z \geq z_{max}$ $h_u(z)$ decreases in u over $u \leq z_{max}$.

We assume observations are taken at $z_1 = u_1$, u_2.

Case 1 :

- $u_1 = a > z_{max}$, $u_2 = b > a$. We show that $v'(a) < 0$, where

$$v'(a) = \frac{6w'(a)}{[w(a)]^2(b-a)}[b - h_b(a)].$$

Now since $a > z_{max}$, we have $w'(a) < 0$. So $v'(a) < 0$ is true if $[b - h_b(a)] > 0$.

$$b - h_b(a) = (b-a) - \frac{2w(a)}{6p(a)w'(a)}$$

The right side of the equation is always positive, because $a < b$ and $w'(a) < 0$. Therefore $v'(a) < 0$.

- $u_2 = b < z_{max}$, $u_1 = a < b$. We show that $v'(b) > 0$.

$$v'(b) = \frac{6w'(b)}{[w(b)]^2(b-a)}[h_a(b) - a]$$

Now $w'(b) > 0$. So $v'(b) > 0$ is true if $[h(b) - a] > 0$.

$$h_a(b) - a = (b - a) + \frac{2w(b)}{6p(b)w'(b)}$$

$[h(b) - a]$ is always positive, because $a < b$ and $w'(b) > 0$. Therefore $v'(u_2) > 0$.

Case 2 : $a^* < a < z_{max} < b < b^* \quad u_1 = a, \quad u_2 = b$

Because $a < z_{max}$ and $b > z_{max}$, we have $w'(a) > 0$ and $w'(b) < 0$.

- $[b - h_b(a)] < [b^* - h_b(a)]$ since $b < b^*$.

  $[b^* - h_b(a)] < [b^* - h_{b^*}(a)]$ since $h_u(z_1)$ decreases in u by (ii).

  $[b^* - h_{b^*}(a)] < [b^* - h_{b^*}(a^*)]$ by (i)

  $[b^* - h_{b^*}(a)] < [b^* - h_{b^*}(a^*)] = 0$

  Therefore $v'(z_1) < 0$.

- $[h_a(b) - a] < [h_a(b) - a^*]$ since $a^* < a$.

  $[h_a(b) - a^*] < [h_{a^*}(b) - a^*]$ since $h_u(z_1)$ decreases in u by (ii).

  $[h_{a^*}(b) - a^*] < [h_{a^*}(b^*) - a^*]$ by (i).

  Therefore $v'(z_1) > 0$.

Case 5 : $a^* \leq a \leq z_{max} \leq b = b^*(a) \quad (b^*(a) > b^*) \quad u_1 = a, u_2 = b^*(a)$

Clearly $v'(b_2) = 0$. First $w'(a) > 0$. We want $v'(b_1) = v'(a)$ to be negative. So we need to the investigate the derivative of $v'(a)$. Here

$$\begin{aligned} v'(b_1) = v'(a) &= \frac{w'(a)}{[w(a)]^2(b-a)}[b^*(a) - h_{b^*(a)}(a)] \\ v'(a) &\propto w'(a)[b^*(a) - h_{b^*(a)}(a)] \end{aligned}$$

where $w'(a) > 0$ since $a^* \leq a \leq z_{max}$. Therefore we need to argue that

$$D = b^*(a) - h_{b^*(a)}(a) < 0 \; \forall a^* \leq a \leq z_{max}.$$

From Theorem (2.1) $a = a^*$ is the only solution to $D = 0$. Further, taking $a = z_{max}$, we get $h_{b^*(z_{max})}(z_{max}) = \infty$ and $b^*(z_{max})$ is finite. Therefore

$D < 0$ at $a = z_{max}$ and hence over $a^* \leq a \leq z_{max}$. Therefore $v'(b_1) = v'(a) < 0$. This design is also D-optimal for $b > b^*(a)$.

Case 6 : $a^* \leq a \leq z_{max}$ $b^* < b < b^*(a)$ $u_1 = a$, $u_2 = b$

First $w'(a) > 0$. Secondly $[b - h_b(a)] < b^*(a) - h_b(a) < b^*(a) - h_{b^*(a)}(a)$. From above $b^*(a) - h_{b^*}(a) < 0$, so $v'(a) < 0$. Also we need to show that $v'(b) > 0$. Because of $b \geq z_{max}$, $w'(b) < 0$. We assumed $b < b^*(a)$. If $h_a(\cdot)$ is an increasing function, $h_a(b) < h_a(b^*(a)) = a$. Hence $h_a(b) - a < [h_a(b^*(a)) - a] = 0$.

3.2.8 Some Empirical Results for D-optimal designs

The general objective has been to find empirically D-optimal designs when $\mathcal{Z} = \{(z_1, z_2) : a \leq z_1 \leq b, -1 \leq z_2 \leq 1\}$ for all possible choices of $\boldsymbol{a}$, $\boldsymbol{b}$. In section (3.2.7) we showed that two distinct values of z_1 produce the support points of the conjectured optimal designs of the various cases of $\mathcal{Z} = [a, b]$. Now we will show empirically that the Equivalence Theorem is satisfied by our conjectured optimal designs for all possible design intervals $[a, b]$. There are only four support points and hence observations are taken at only two values of z_1.

Case 1 : $\mathcal{Z} = \mathcal{Z}_w = \{(z_1, z_2) : -\infty \leq z_1 \leq \infty \;\; -1 \leq z_2 \leq 1\}$ and $Supp(p^*) = \{-b^*, b^*\}$

In the case of **symmetric** weight functions $w(z_1)$, z_1- support points are $\pm b^*$ with $z_2 = \pm 1$ and with equal weights of $\frac{1}{4}$ where b^* maximizes $\{detM(p) = b^2[w(b)]^3\}$. As did Sitter and Torsney (1995) we found that the b^* value that maximizes $detM(p)$ is $b = \pm 1.22$ for the logistic regression model.

We checked for the optimality of this design, by checking the Equivalence Theorem for $z_1 \in (-\infty, \infty)$, and $z_2 = \pm 1$. Figure (3.7) presents the variance function. The design is globally D-optimal on $\mathcal{Z}_w$.

Case 2 : $\mathcal{Z} = \{(z_1, z_2) : a \leq z_1 \leq b, \;\; -1 \leq z_2 \leq 1\}$

and $Supp(p^*) = \{max\{a, a^*(b)\}, b\} \; a < a^*, \; b < b^*$

Results are very similar with the next case. So we only include empirical results for that.

Case 3 : $\mathcal{Z} = \{(z_1, z_2) : a \leq z_1 \leq b \;\; -1 \leq z_2 \leq 1\}$

and $Supp(p^*) = \{a, min\{b, b^*(a)\} \; a > a^*, \; b > b^*$

Here we have to choose $b^*(a)$, p_a, p_b for fixed a to maximize $det(M(p))$ where

$$det(M(p)) = (b-a)^2 p_a p_b w(a) w(b)[p_a w(a) + p_b w(b)].$$

Recall that there is an explicit solution for the optimal weights p_a, p_b given any a, b for any weight function and assuming a design of the above form, namely,

$$p_a = \frac{[2w(b) - w(a)] \pm \sqrt{[2w(b) - w(a)]^2 - 3[w(b) - w(a)]w(b)}}{2 * 3[w(b) - w(a)]}$$

So we could substitute for p_a and p_b in terms of a, b and maximize the resultant function with respect to b. This was done using a simple search to find b^*. However, a possible alternative to this is the following ***alternating algorithm.***

Alternating Algorithm Steps :

1. *a fixed.*

2. *Choose* $b^{(0)}$ *initial value for b.*

3. *Let* $p_a^{(0)}$, $p_b^{(0)}$ *be the optimal weights for a,* $b^{(0)}$.

4. *Keeping* $p_a^{(0)}$, $p_b^{(0)}$ *fixed, use the Newton-Raphson Method to maximize* $det(M(p))$ *with respect to b. Let the solution be* $b^{(1)}$.

5. *Let* $p_a^{(1)}$, $p_b^{(1)}$ *be the optimal weights for a,* $b^{(1)}$.

6. *Keeping* $p_a^{(1)}$, $p_b^{(1)}$ *fixed use the Newton-Raphson Method to maximize* $det(M(p))$ *with respect to b (using* $b^{(1)}$ *as initial approximation). Let solution be* $b^{(2)}$.

   $\vdots$

The optimal design for $b = 3$ and $a = -1.22, -1.20, -1.10, -1.00, -0.90$ $\cdots$ 1.10,1.20 were calculated using this **alternating algorithm**. Results are summarized in Table (3.1). Relevant variance functions are plotted in Figures (3.8) and (3.9). These show that the necessary and sufficient condition of the equivalence theorem is satisfied.

Case 4 : $\mathcal{Z} = \{(z_1, z_2) : a \leq z_1 \leq b \ -1 \leq z_2 \leq 1\}$
and $Supp(p^*) = \{a, b\}$ $a > a^*$, $b < b^*$
For this z_1-interval, end points are support points, and the equivalence theorem is satisfied in the examples considered: $a = -1$, $b = 1$; $a = -0.75$, $b = 0.75$; $a = -0.50$, $b = 0.50$. See Figure (3.10).

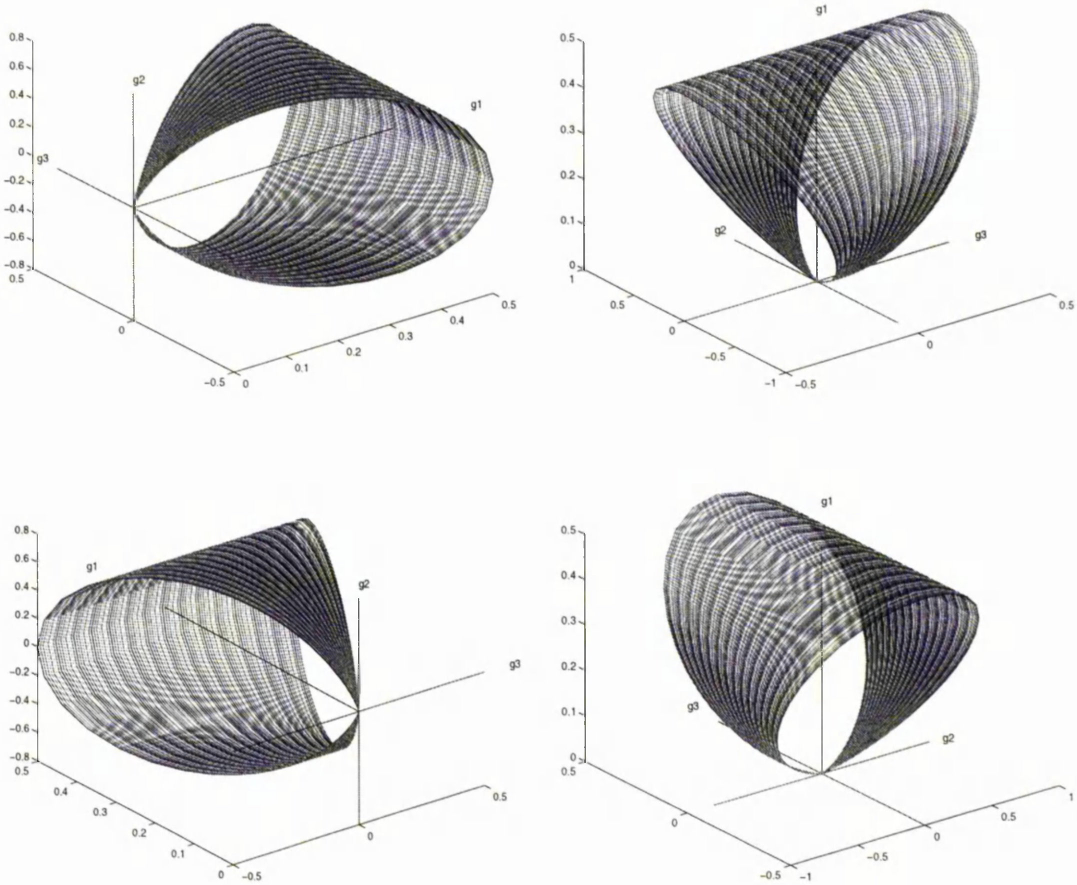
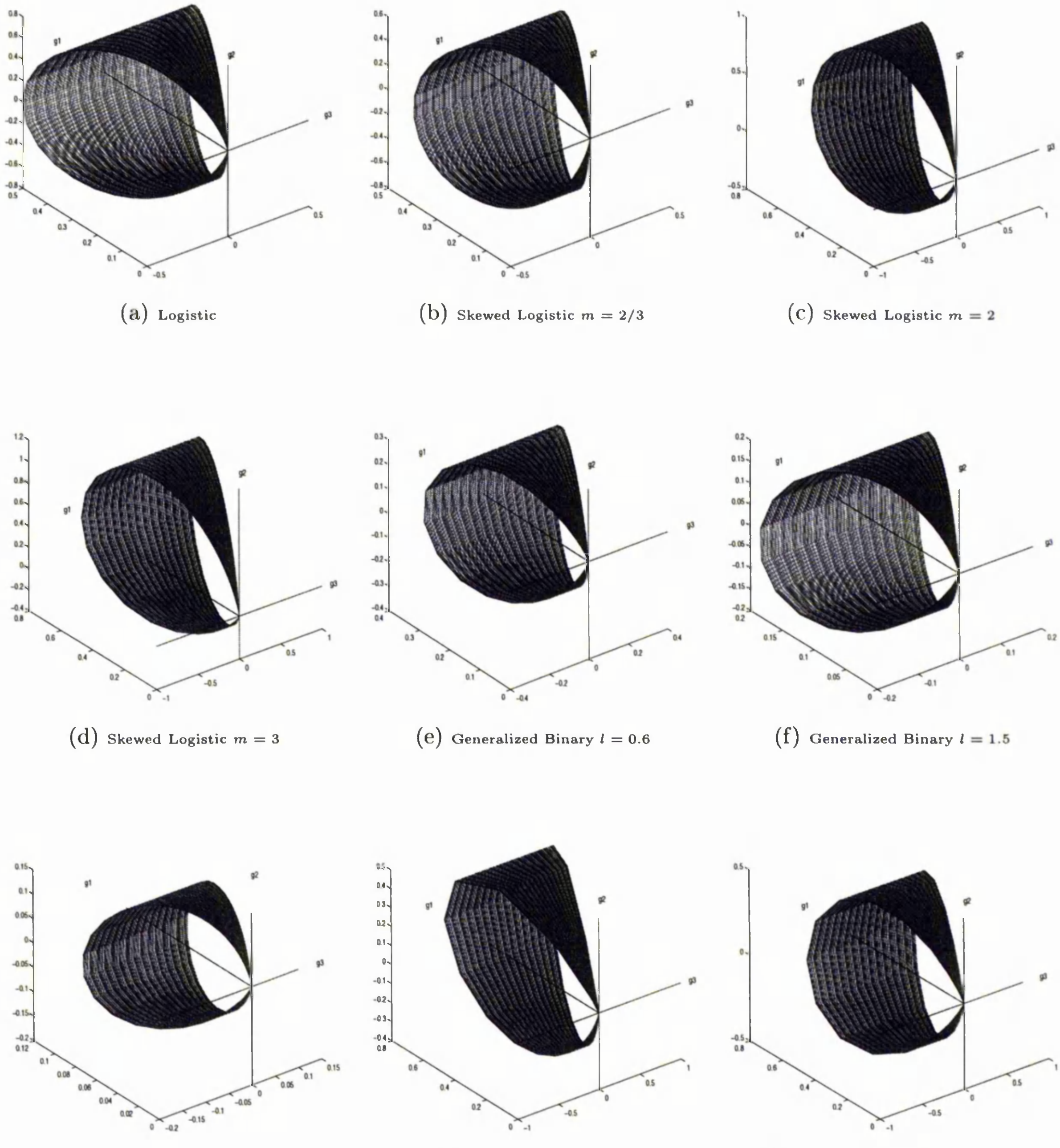


Figure 3.1: Plots of 3-D $G(\mathcal{Z})$ for the Logistic Weight Function with different orientations obtained by considering different permutations of axes, $g_1 = \sqrt{w(z_1)}$, $g_2 = z_1\sqrt{w(z_1)}$, $g_3 = z_2\sqrt{w(z_1)}$ and $-20 \leq z_1 \leq 20$, $-1 \leq z_2 \leq 1$.



(a) Logistic

(b) Skewed Logistic $m = 2/3$

(c) Skewed Logistic $m = 2$

(d) Skewed Logistic $m = 3$

(e) Generalized Binary $l = 0.6$

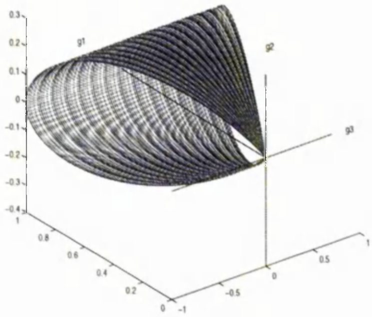
(f) Generalized Binary $l = 1.5$

(g) Generalized Binary $l = 3$

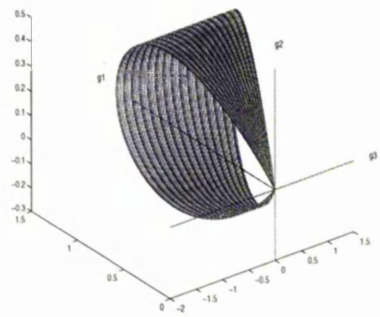
(h) Complementary log-log

(i) Probit

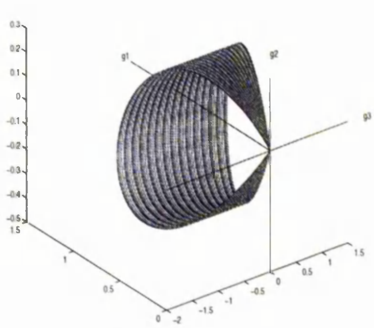
Figure 3.2: Plots of 3-D $G(\mathcal{Z})$ for Group I-binary weight functions. $g_1 = \sqrt{w(z_1)}$, $g_2 = z_1\sqrt{w(z_1)}$, $g_3 = z_2\sqrt{w(z_1)}$ and $-20 \leq z_1 \leq 20$, $-1 \leq z_2 \leq 1$ (Note : $\boldsymbol{l}$ represents the parameter $\boldsymbol{\lambda}$ of the Generalized Binary case).



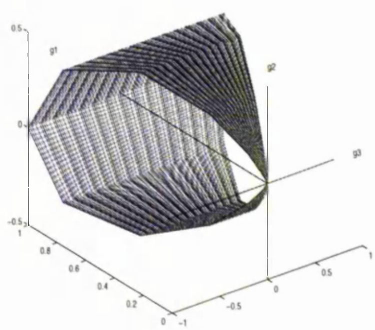
(a) Beta, $a = 3$, $b = 3$



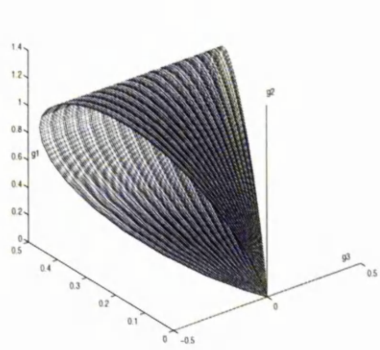
(b) Beta, $a = 0.5$, $b = 1.3$



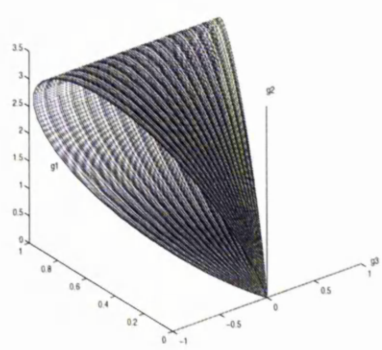
(c) Beta, $a = 0.9$, $b = 0.3$



(d) Normal

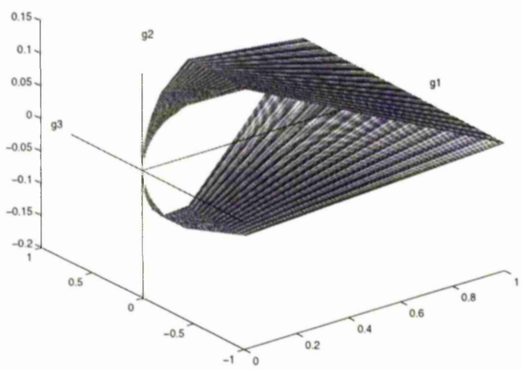


(e) Gamma $g = 0.8$

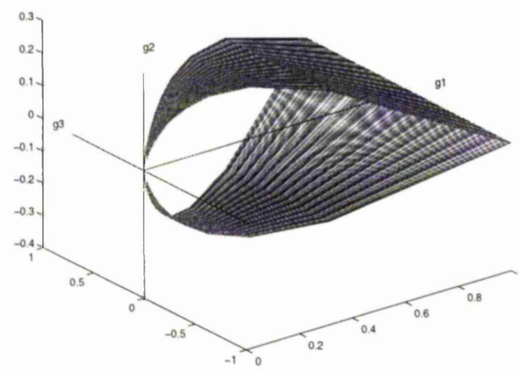


(f) Gamma $g = 1.7$

Figure 3.3: 3-D plots of $G(\mathcal{Z})$ for Group III: Density weight functions. $g_1 = \sqrt{w(z_1)}$, $g_2 = z_1\sqrt{w(z_1)}$, $g_3 = z_2\sqrt{w(z_1)}$.
(a represents α, b represents β, g represents γ and interval $-1 \leq z_1, z_2 \leq 1$; interval $-20 \leq z_1 \leq 20$, $-1 \leq z_2 \leq 1$; and interval $0 \leq z_1 \leq 20$, $-1 \leq z_2 \leq 1$ are used to draw the plots for Beta, Normal and Gamma weight functions respectively).

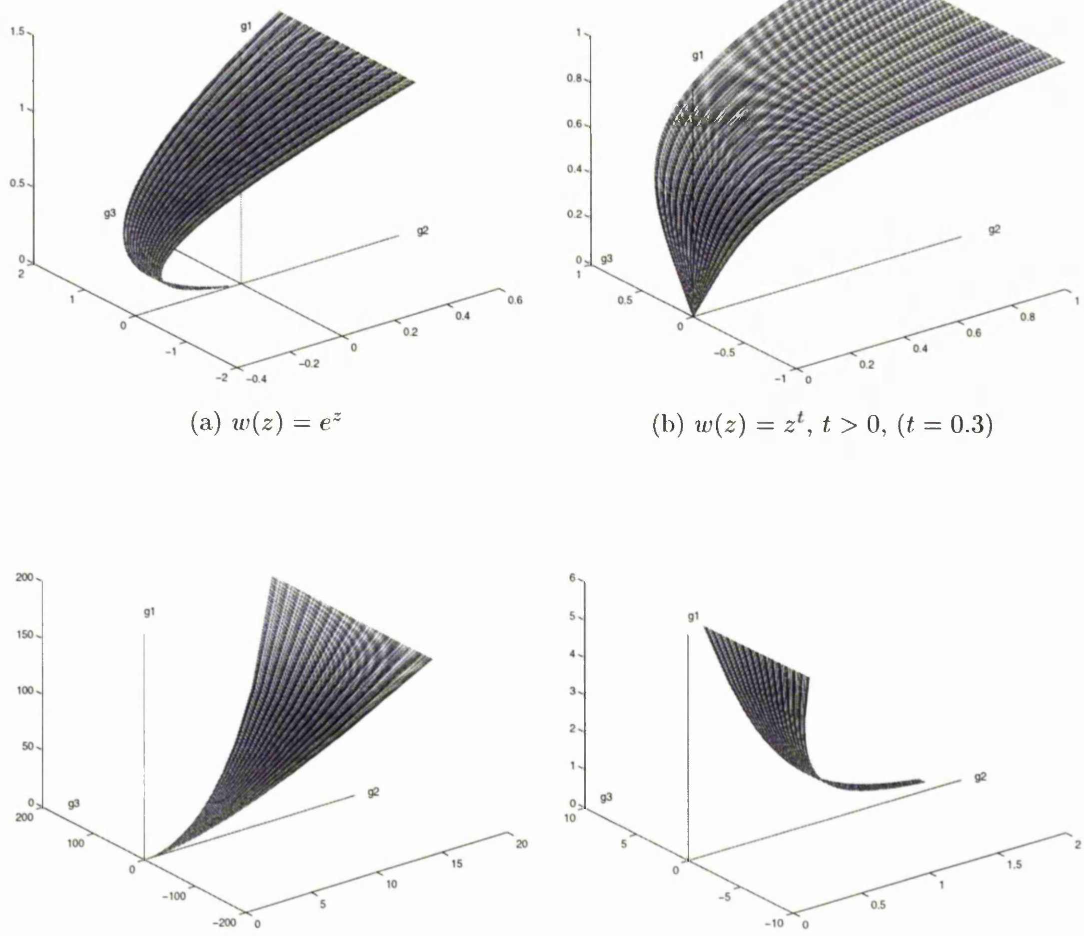


(a) Double Reciprocal



(b) Double Exponential

Figure 3.4: 3-D plots of $G(\mathcal{Z})$ for Group II; Double Reciprocal & Double Exponential Binary weight functions and $g_1 = \sqrt{w(z_1)}$, $g_2 = z_1\sqrt{w(z_1)}$, $g_3 = z_2\sqrt{w(z_1)}$ and $-20 \leq z_1 \leq 20$, $-1 \leq z_2 \leq 1$.



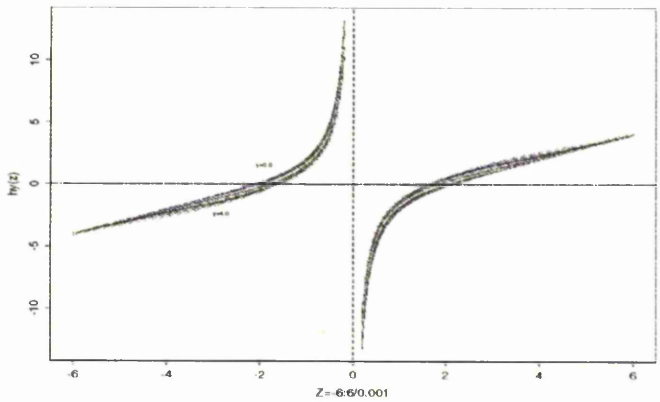
(a) $w(z) = e^z$

(b) $w(z) = z^t$, $t > 0$, $(t = 0.3)$

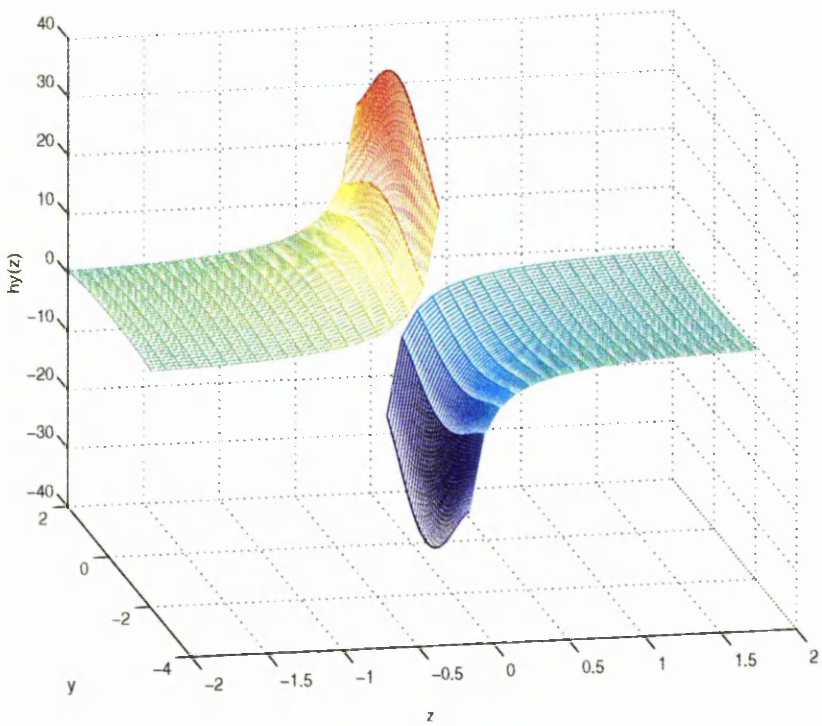
(c) $w(z) = z^t$, $t < -2$, $(t = -2.2)$

(d) $w(z) = z^t$, $-2 < t < 0$, $(t = -0.7)$

Figure 3.5: 3-D plots of $G(\mathcal{Z})$ for Group IV weight functions and $g_1 = \sqrt{w(z_1)}$, $g_2 = z_1\sqrt{w(z_1)}$, $g_3 = z_2\sqrt{w(z_1)}$: details are: $w(z) = e^z$, $-4.5 \leq z_1 \leq 0.3$, $-1 \leq z_2 \leq 1$; $w(z) = z^t$, $t > 0$, $(t = 0.3)$, $0 \leq z_1 \leq 1$, $-1 \leq z_2 \leq 1$; $w(z) = z^t$, $t < -2$, $(t = -2.2)$, $0.1 \leq z_1 \leq 3$, $-1 \leq z_2 \leq 1$; and $w(z) = z^t$, $-2 < t < 0$, $(t = -0.7)$, $0.1 \leq z_1 \leq 6$, $-1 \leq z_2 \leq 1$.



(a) Plot of $h_y(z)$ function for the Logistic Weight Function, with y=0.0,0.05,0.1,0.15,0.20,0.25,0.30,0.35,0.40 and number of parameter k is 3.



(b) 3-D Plots of $h_y(z)$ for the Logistic Weight Function, $h_y(z) = z + \frac{w(z)}{3p_y(z)w'(z)}$ $-2.05 \leq y \leq 1.95$ and $-2 \leq z \leq 2$.

Figure 3.6: Two Different plots of $h_y(z)$ for the Logistic Weight Function.

Three parameter case: Logistic weight Function, $z_1 \in [a, \infty)$ for fixed $a > -b^*$ optimal b $p_b(a)$ and $p_a(b)$ value.			
fixed a value	$b^*(a)$	$p_b(a)$	$p_a(b)$
-1.22291	1.222905	0.250000	0.250000
-1.20000	1.236604	0.251243	0.248757
-1.10000	1.298286	0.256628	0.243372
-1.00000	1.362882	0.261888	0.238112
-0.90000	1.430150	0.266959	0.233041
-0.80000	1.499816	0.271783	0.228217
-0.70000	1.571598	0.276319	0.223681
-0.60000	1.645228	0.280539	0.219461
-0.50000	1.720466	0.284429	0.215571
-0.40000	1.797109	0.287988	0.212012
-0.30000	1.874993	0.291222	0.208778
-0.20000	1.953995	0.294149	0.205851
-0.10000	2.034021	0.296786	0.203214
0	2.115009	0.299156	0.200844
0.10000	2.196914	0.301281	0.198719
0.20000	2.279705	0.303185	0.196815
0.30000	2.363363	0.304889	0.195111
0.40000	2.447872	0.306412	0.193588
0.50000	2.533219	0.307775	0.192225
0.60000	2.619389	0.308995	0.191005
0.70000	2.706368	0.310085	0.189915
0.80000	2.794137	0.311062	0.188938
0.90000	2.882678	0.311936	0.188064
1.10000	-1.298286	0.256628	0.243372
1.20000	-1.236604	0.251243	0.248757

Table 3.1: For the Logistic weight function: D-optimal support points and weights.

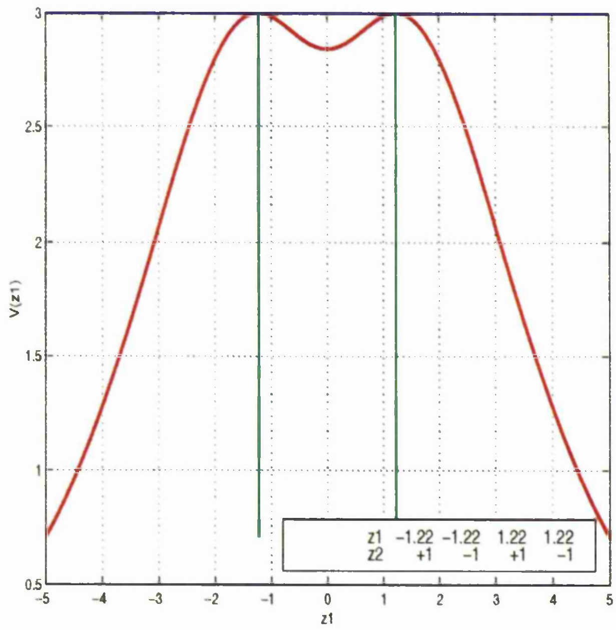


Figure 3.7: Plot of the variance function for the global symmetric D-optimal four-point design on $\mathcal{Z} = \mathcal{Z}_w = \{(z_1, z_2) : -\infty \leq z_1 \leq \infty \; -1 \leq z_2 \leq 1\}$ for the logistic weight function. Note this plot is only for $z_2 = \pm 1$.

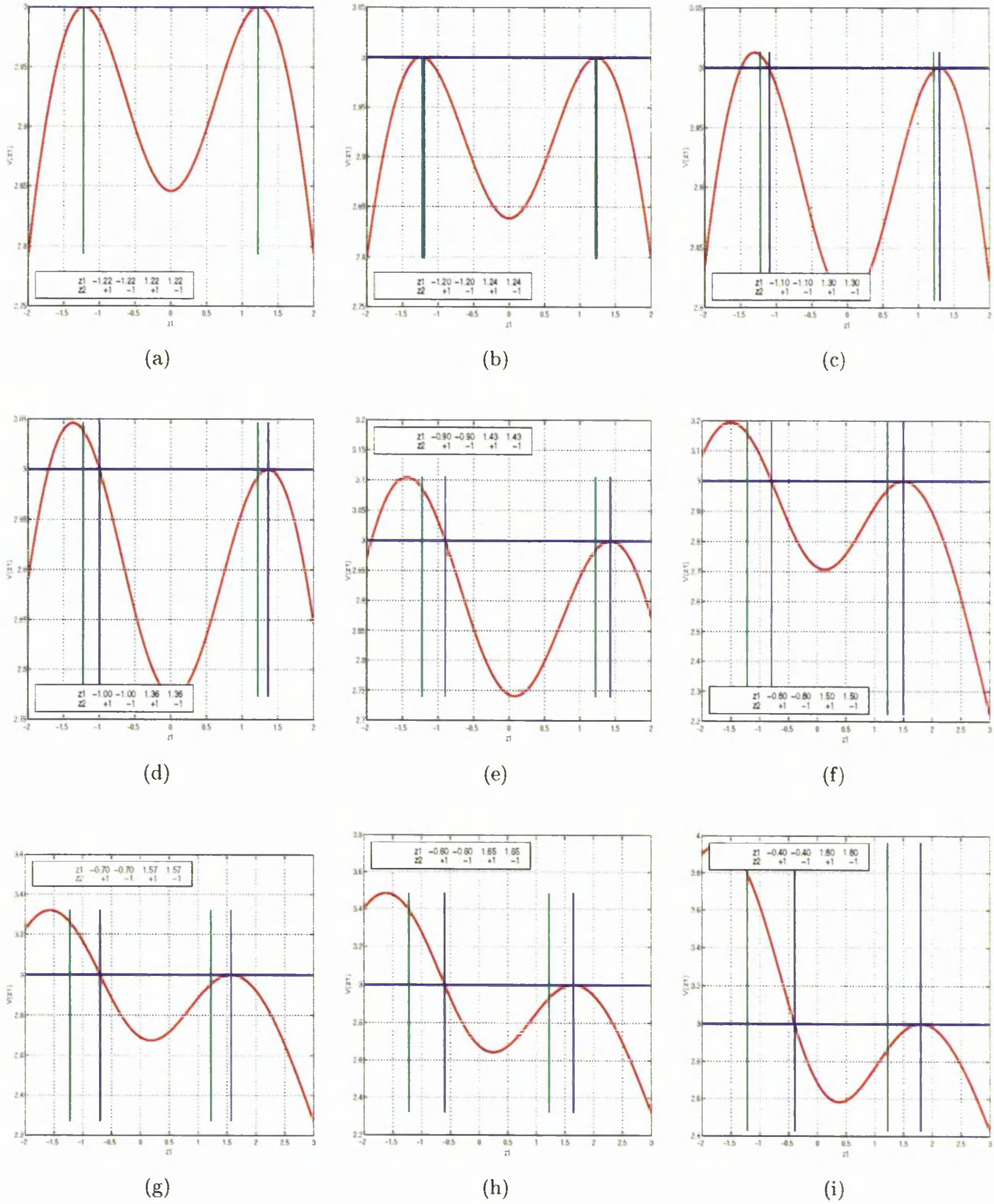


Figure 3.8:

Some plots of the variance function, $\mathcal{V}(z_1)$, under an optimal design on $\mathcal{Z} = \{(z_1, z_2) : a \leq z_1 \leq b \; z_2 = \pm 1\}$, $a > a^*$, $b > b^*(a)$ for the logistic weight function, $(k = 3)$.

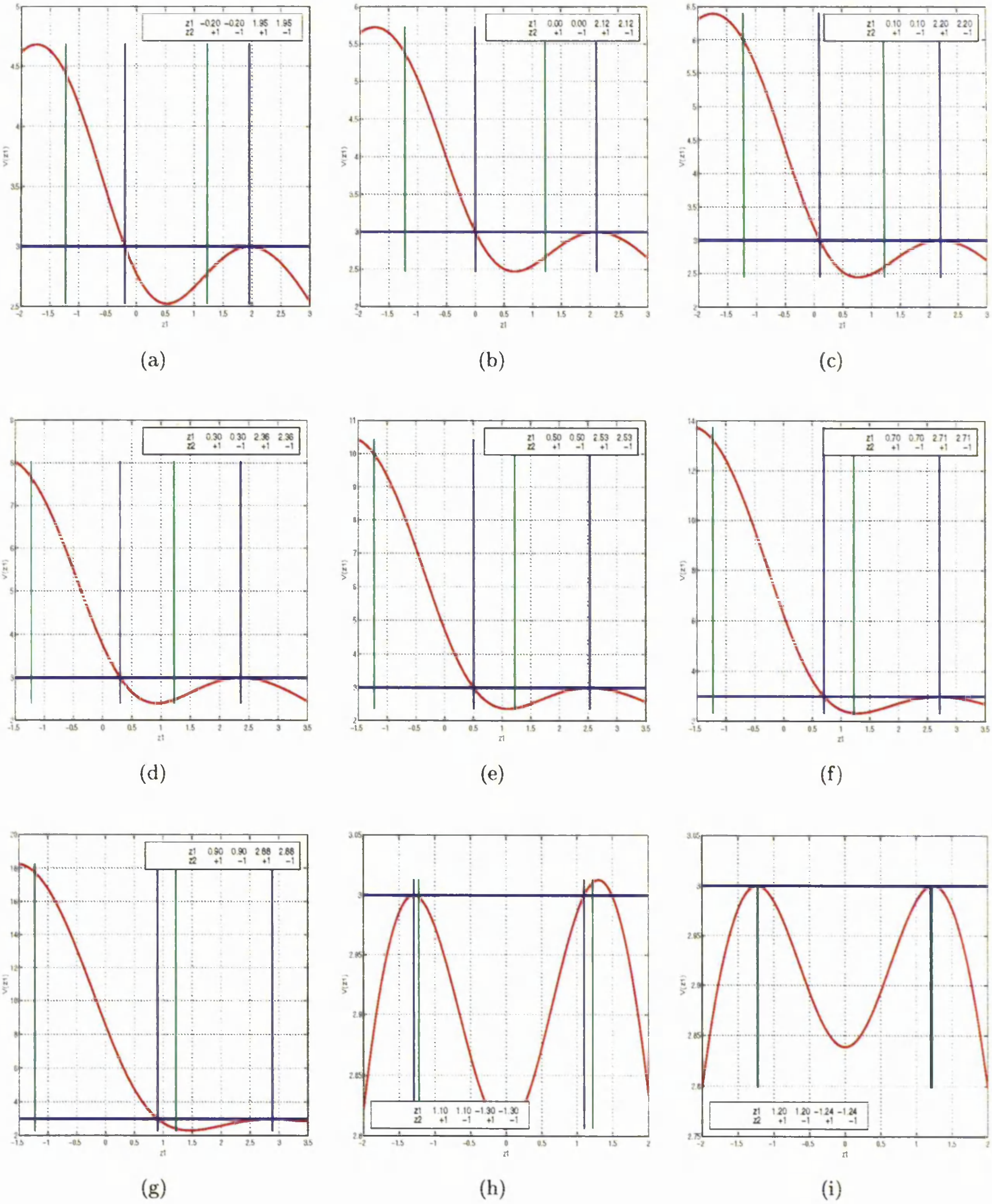
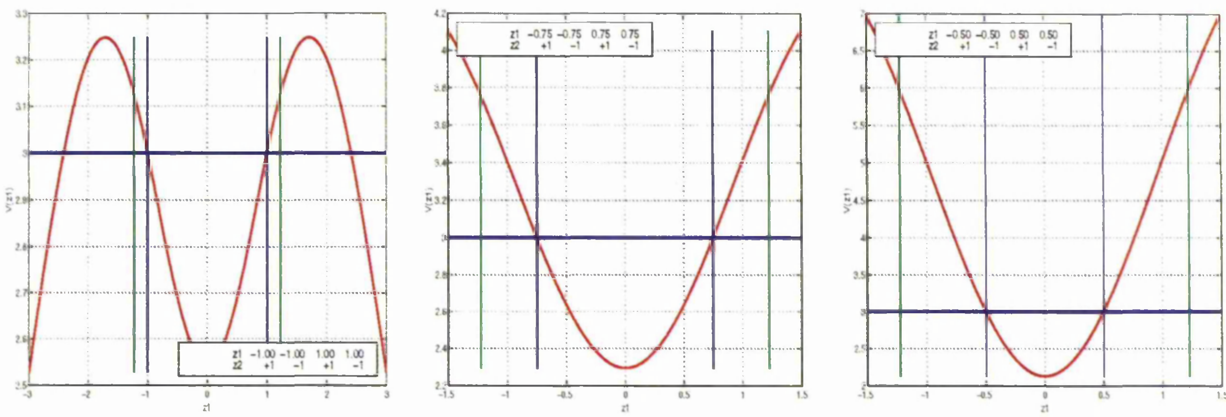


Figure 3.9: Some plots of the variance function, $\mathcal{V}(z_1)$, under an optimal design on $\mathcal{Z} = \{(z_1, z_2) : a \leq z_1 \leq b \; z_2 = \pm 1\}$ for the logistic weight function ($k = 3$) : $a > a^*$, $b > b^*(a)$ in plots (a) to (g); $b < b^*$, $a < a^*(b)$ in plots (h), (i).



(a) (b) (c)

Figure 3.10: Some plots of the variance function, $\mathcal{V}(z_1)$, of the D-optimal design on the $\mathcal{Z} = \{(z_1, z_2) : a \leq z_1 \leq b \; z_2 = \pm 1 \; a > a^*, \, b < b^*$ for the logistic weight function where a, b are support points.

Chapter 4

A design region for z_1, z_2 in the form of a polygon : The Case of the Three Parameter Model

4.1 Bounded design spaces

We first transformed $\mathcal{X}$ to the new design space, $\mathcal{Z}$, with two new design variables $z_1 = \alpha + \beta_1 x_1 + \beta_2 x_2$, $z_2 = a + bx_1 + cx_2$ where a, b, c are arbitrary constants to be chosen by the experimenter. We also considered the further transformation $\mathcal{Z} \longrightarrow G = G(\mathcal{Z})$. This set G needs to be bounded and then we have the characterisation that D-optimal designs have as support points the points of contact between G and the smallest ellipsoid centered on the origin containing G (Silvey,1980). A minimum requirement for G to be bounded is that z_2 be bounded. Without loss of generality, we assumed that $-1 \leq z_2 \leq 1$. Bounds are actually not necessary on z_1.

So initially, we assumed $\mathcal{X}$ such that, $\mathcal{Z} = \mathcal{Z}_w$, $G_w = G(\mathcal{Z}_w)$, where $\mathcal{Z}_w$ is the widest possible design space. Then we considered the case $\mathcal{Z} = \{(z_1, z_2) : a \leq z_1 \leq b, -1 \leq z_2 \leq 1\}$ so that z_1 is potentially bounded. This is the case of a subset of G_w which is a 'vertical' (in the g_3-direction) portion of G_w.

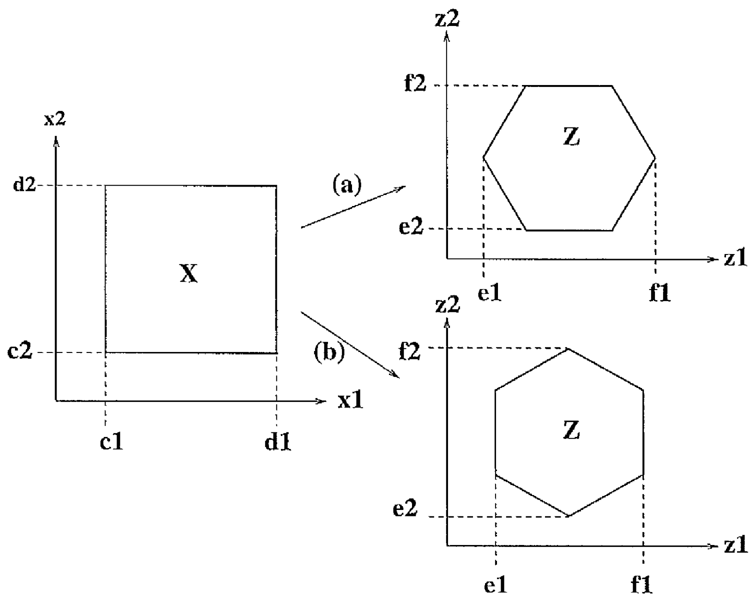


Figure 4.1: Diagram illustrating the transformation from $\mathcal{X}$ design space to $\mathcal{Z}$ design space, which creates a polygon

Now we consider other possibilities. The most likely scenario is that $\mathcal{X}$ is the set of pairs (x_1, x_2), satisfying $c_i \leq x_i \leq d_i$, $i = 1, 2$, so that it is rectangular. Thus $\mathcal{X} = \{(x_1, x_2) : c_i \leq x_i \leq d_i, \quad i = 1, 2\}$. This transforms $\mathcal{X}$, into a polygon $\mathcal{Z}$ in the variables $z_1 = \alpha + \beta x_1 + \gamma x_2$, $z_2 = a + bx_1 + cx_2$. It is a polygon with at most 6 sides [See Figure 4.1]. The number of sides will depend on the choice of z_1 and z_2. For example if $z_1 = x_1$ (if $\alpha = \gamma = 0$, $\beta = 1$) and $z_2 = x_2$ then $\mathcal{Z}$ is a rectangle. Of course the definition of z_1 is fixed, but z_2 is a free choice for us and the number of sides of the polygon may depend on this choice.

Other possibilities are that a bounded set $\mathcal{X}$ may be defined by other linear constraints. For example there may be limits on x_1+x_2 if x_1, x_2 are the component values of a mixture of two drugs. These could lead to quadritalerals, pentagons or hexagons as the form of $\mathcal{Z}$.

We consider simple cases first. A general point is that if $\mathcal{Z}$ is bounded, finite limits will be imposed on z_1 and z_2, say $e_i \leq z_i \leq f_i$, $i = 1, 2$. Thus $\mathcal{Z}$ is contained in the rectangle $\{(z_1, z_2) : e_i \leq z_i \leq f_i \quad i = 1, 2\}$. Again without loss of generality we assume $e_2 = -1$, $f_2 = 1$. Hence $G = G(\mathcal{Z}) \subseteq G(\mathcal{Z}_w)$. In the following examples we take $w(z)$ to be the binary logistic regression weight function.

4.1.1 Examples

Example 4.1. *First we recall the global D-optimal design on G_w for Logistic regression in the case $k = 3$. This is :*

$$\begin{pmatrix} i & 1 & 2 & 3 & 4 \\ z_{1i} & a^* & a^* & b^* & b^* \\ z_{2i} & -1 & 1 & -1 & 1 \\ p_i & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$$

where $a^* = -1.22$ and $b^* = 1.22$ are the support points. The design region for z_1, z_2 has the form of a infinite rectangle. We note that this remains the optimal design for any finite rectangle $\mathcal{Z} = \{(z_1, z_2),\ A \leq z_1 \leq B,\ -1 \leq z_2 \leq 1\}$ if $A \leq -1.22$, $B \geq 1.22$ e.g. $A = -B = 3$. See Figure 4.2

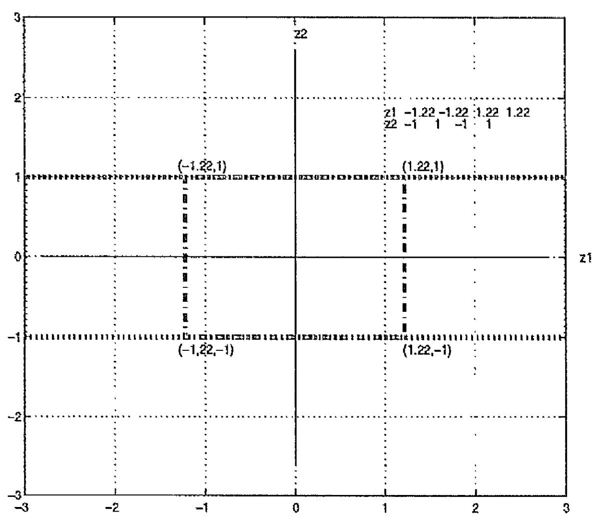
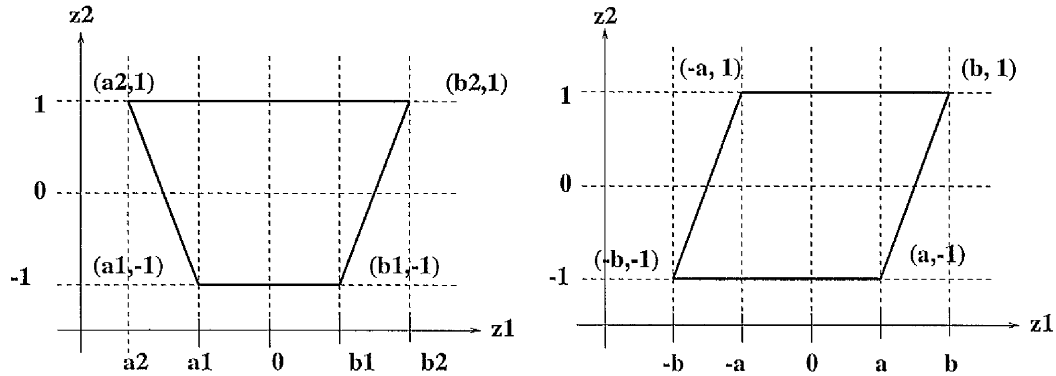


Figure 4.2: Design Region for the case of two design variables using the Logistic weight function.



(a) $\mathcal{Z}$ is quadriteral (b) $\mathcal{Z}$ is parallelogram

Figure 4.3:

Example 4.2. *Now suppose $\mathcal{Z}$ is the quadrilateral with vertices $(a_1, -1), (a_2, 1), (b_1, -1), (b_2, 1)$ [See Figure 4.3 a]. The above design is still optimal if $a_1, a_2 < -1.22$ and $b_1, b_2 > 1.22$. So we want to consider a_1 and/or $a_2 > -1.22$ and/or b_1 and/or $b_2 > 1.22$. Consider the case $b_1 = -b_2 = -b$, $a_2 = -a_1 = -a$.*

If these 4 corners are the support points the symmetry implies that the structure of the design should be as follows:

$$\begin{pmatrix} i & 1 & 2 & 3 & 4 \\ z_{1i} & -b & -a & a & b \\ z_{2i} & -1 & 1 & -1 & 1 \\ p_i & p_b & p_a & p_a & p_b \end{pmatrix}$$

where p_b, $p_a > 0$, $2p_b + 2p_a = 1$, and $p_a = \frac{1}{2} - p_b$ [See, Figure 4.3 b]. The information matrix is

$$M(\xi) = \Sigma_i^4 p_i \underline{g}_i \underline{g}_i^T$$

where

$$\underline{g}_i = \sqrt{w(z_{1i})}\underline{v}_i \qquad \underline{v}_i = (1, z_{1i}, z_{2i})^T \qquad i = 1, 2, 3, 4$$

Therefore,

$$M(\xi) = \begin{pmatrix} 2p_b w(b) + 2p_a w(a) & 0 & 0 \\ 0 & 2p_b b^2 w(b) + 2p_a a^2 w(a) & 2p_b b w(b) + 2p_a a w(a) \\ 0 & 2p_b b w(b) + 2p_a a w(a) & 2p_b w(b) + 2p_a w(a) \end{pmatrix}$$

The determinant is

$$|M(\xi)| \;=\; 2^3\,(b+a)^2\,p_b\,p_a\,w(b)\,w(a)\,[p_b\,w(b)\;+p_a\,w(a)]$$

We need to choose p_b *to maximise* $\mathbf{det\{M(\boldsymbol{\xi})\}}$. *This is exactly the same as* $|M(\xi)|$ *in chapter 3, section 3.2.3 . Hence we have the same explicit solution for the weights:*

$$p_b \;=\; \frac{[w(b)-2w(a)]+\sqrt{[w(b)-2w(a)]^2+3[w(b)-w(a)]w(a)}}{6[w(b)-w(a)]} \qquad (4.1)$$

Hence $p_a \;=\; \frac{1}{2}-p_b$. *Now our question is :* ***Is this design D-optimal on*** $\mathcal{Z}$ ***or* G*?*** *We need to check the Equivalence Theorem. Given the minimal ellipsoid characterisation of D-optimalty we only need to check along the edges of* $\mathcal{Z}$. *Each edge either corresponds to* $z_2 = \pm 1$ *(as before) or can be viewed as defining* z_2 *as a linear function of* z_1*, say* $z_2 = mz_1 + c$*, for some range of values of* z_1*, say* $A \leq z_1 \leq B$. *Hence we are interested in checking for* $A \leq z_1 \leq B$ *the value of*

$$\mathcal{V}(z_1) \;=\; \underline{g}^T M^{-1}(\xi^*)\underline{g}$$

where $\underline{g}^T \;=\; \sqrt{w(z_1)}(1,\, z_1,\, S(z_1))$, $S(z_1) \;=\; mz_1 + c$ *and* ξ^* *is the conjectured optimal design. We require* $\mathcal{V}(z_1) \leq 3$ *for* $A \leq z_1 \leq B$.

For **Example 4.2** we checked the equivalence theorem for the following value(s) of b and a :

$b = 2$, $a = 1.22,\ 1,\ 0.75,\ 0.50,\ 0.25,\ 0,\ -0.25,\ -0.50$

Complete plots consist of possibly 4 distinct $\mathcal{V}(z_1)$-curves. In Figure (4.10) we show plots depicting the four relevant curves simultaneously. It is clear that in some cases these curves are partly above 3, and hence the Equivalence theorem is not satisfied. Only for $b = 2$, $a = 0$ is the **'four-corner'**design **optimal** [See Figure 4.14].

Further Example 4.1. *In Figure (4.11) we show 6 other quadrilateral choices for* $\mathcal{Z}$. *Optimal weights under the designs with the four corners as support points are given in Table (4.1). These are not optimal for* $\mathcal{Z}$*; see Figure (4.12). For these designs since* $z_1 = 0$ *there is an explicit solution for the optimal weights as above.*

Now a question of interest is ' **What is the optimal design when the 'four-corner' design is not optimal?'**.

4.2 Determination of design using an algorithm

Initially, we use a class of algorithms ((Torsney, 1983) and (Torsney and Alahmadi,1992)) to find these optimal designs. The algorithm is indexed by a function which depends on derivatives and a free parameter (say δ) for a constrained maximisation problem which requires the calculation of an optimizing probability distribution. Such algorithms are needed since in general there is no explicit solution for optimal designs or weights.

First we must establish conditions of optimality. It helps to consider the following general problems, of which the design problem is an example.

Problem 4.1 (P1). *Maximise a criterion* $\Phi(p)$ *subject to the constraints* $\sum_{j=1}^{J} p_j = 1$ *and* $p_j \geq 0$.

Problem 4.2 (P2). *Maximise* $\Psi(X)$ *over the polygon whose vertices are the points* $G(v_1), G(v_2), \cdots, G(v_j)$*, where* $G(\cdot)$ *is a given one to one function and* $\mathcal{V} = \{v_1, v_2, \cdots, v_j\}$ *is a known set of vectors (or matrices) vertices of fixed dimension. This is solve (P1) for:*

$$\Phi \;=\; \Psi\{E_p[G(v)]\}, \qquad X = E_p[G(v)] = \sum_{j=1}^{J} G(v_j).$$

4.2.1 Optimality Conditions

We concentrate on Problem (P2) and define optimality conditions in terms of point to point directional derivatives.

4.2.2 Directional Derivatives

Let

$$\begin{aligned} f(X,Y,\zeta) &= \Psi\{(1-\zeta)X+\zeta Y\} \\ F_{\Psi}(X,Y) &= \left.\frac{\partial f(X,Y,\zeta)}{\partial \zeta}\right|_{\zeta=0^+} = \lim_{\zeta\to 0}\frac{1}{\zeta}[\Psi\{(1-\zeta)X+\zeta Y\}-\Psi(X)] \end{aligned}$$

$F_{\Psi}(X,Y)$ is known as the directional derivative of $\Psi(\cdot)$ at X in the direction of Y, [(Whittle 1973)].

Let$F_j = F_{\Psi}\{X, G(v_j)\}$. We call F_j a vertex directional derivative of $\Psi(\cdot)$ at X. If $\Psi(\cdot)$ is differentiable, then so is the function $\Phi(p) = \Psi\{E_p[G(v)]\}$ and

$$F_j = d_j - \sum_{i=1}^{J} p_i d_i,$$

where $d_j = \frac{\partial\Phi}{\partial p_j}$.

4.2.3 Condition for Local Optimality

If $\Psi(\cdot)$ is differentiable at $\mathcal{X} = E_p\{\mathcal{G}(v)\}$, then $\Psi(\mathcal{X})$ is a local maximum of $\Psi(\cdot)$ in the feasible region of problem (P2) if,

$$F_j = F_{\Psi}\{X, G(v_j)\} \begin{cases} = 0 & if \quad p_i^* > 0 \\ \leq 0 & if \quad p_i^* = 0 \end{cases} \tag{4.2}$$

If $\Psi(\cdot)$ is concave on its feasible region then the first order stationary conditions (4.2) is both necessary and sufficient for a solution to problem (P2). Indeed this is the General Equivalence Theorem in optimal design.

4.2.4 A Class of Algorithms

Problem (P1) has a distinctive set of constraints, which are that the variables $p_1, p_2, \cdots p_j$ must be positive and sum to 1. An iteration which preserves these and has respectable properties is

$$p_j^{r+1} = p_j^r f(d_j^{(r)}, \delta) \;/\; \sum_{i=1}^{J} p_i^{(r)} f(d_i^{(r)}, \delta) \tag{4.3}$$

where $d_j^{(r)} = \frac{\partial \phi}{\partial p_j} |_{p=p^{(r)}}$, while $f(d, \delta)$ satisfies the following conditions:

(i) $f(d, \delta) > 0$,

(ii) $f(d, \delta) > 0$ is strictly increasing in d for some set of δ-values, say $\delta > 0$,

(iii) the variable δ is a free parameter.

Properties of the iteration:

a - $p^{(r)}$ is always feasible.

b - $F_\phi\{p^{(r)}, p^{(r+1)}\} \geq 0$ with equality when the d_j corresponding to nonzero p_j is equal (in which case $p^{(r+1)} = p^{(r)}$).

c - Let $v_1, v_2, \dots v_J$ be the vertices of the feasible region of (P1) and $\mathcal{V}$ be the induced design space. Let $supp(p) = \{v_j \in \mathcal{V} : \ p_j > 0 \}$ denote the support of the distribution p. Under the above iteration $supp(p^{(r+1)}) \subseteq supp(p^{(r)})$.

d - An iterate $p^{(r)}$ is a fixed point of the iteration if the derivatives $\partial\phi/\partial p_j^{(r)}$ corresponding to nonzero $p_j^{(r)}$ are all equal. This is a necessary but not a sufficient condition for $p^{(r)}$ to solve (P1).

4.2.5 Results of the algorithm and explicit solution for the weights for this new design

The optimal support points found using this algorithm in all cases of Example 4.2 are summarized in Table 4.2 and 4.3.

We note the following:

i - Support points consist of two corners and one point on each of two opposite sides. [See Figure 4.4]

ii - The designs are symmetric as is to be expected given the symmetry of $w(z)$ and of $\mathcal{Z}$ about $z_1 = 0$. In fact the designs are of the following form:

$$\xi = \begin{pmatrix} z_{1i} & -b & -a & a & b \\ z_{2i} & -d & -c & c & d \\ p_i & q & p & p & q \end{pmatrix}, \qquad (4.4)$$

where $i = 1, \cdots, 4$.

iii- It turns out that we can find an explicit solution for the weights, q and p given a, b, c, d: For the above design the determinant $|M(\xi)|$ has the following form (Fedorov 1972 page 83-84):

$$\begin{aligned} |M(\xi)| &= qw(b)p^2w^2(a)[detV_{123}]^2 \\ &+ q^2w^2(b)pw(a)[detV_{124}]^2 \\ &+ q^2w^2(b)pw(a)[detV_{134}]^2 \\ &+ qw(b)p^2w^2(a)[detV_{234}]^2 \end{aligned}$$

where V_{123}, V_{124}, V_{134} and V_{234} are all possible 3×3 minor matrices of the above design. For the above design $[detV_{ijk}]^2 = 4(bc - ad)^2 \ \forall \ i, j, k$. Therefore, the determinant is

$$|M(\xi)| = 8(bc - ad)^2 qw(b)pw(a)[qw(b) + pw(a)]$$

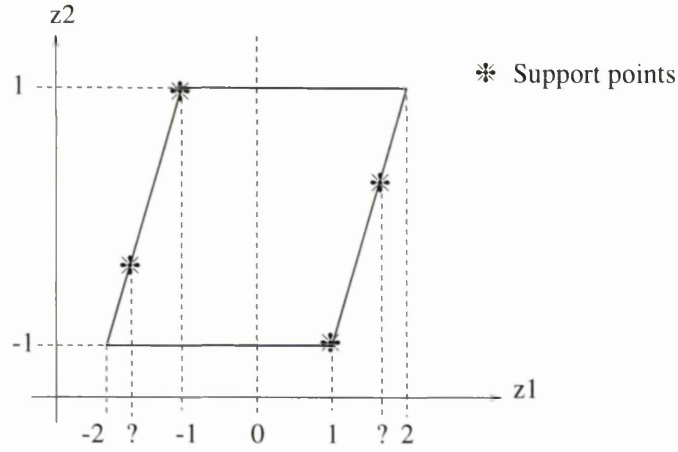


Figure 4.4: Support point for optimal design

Note that, as a function of q, p, this is independent of c,d. In fact we have exactly the same explicit solution as earlier for the weights for given a,b : namely

$$q = \frac{[w(b) - 2w(a)] + \sqrt{[w(b) - 2w(a)]^2 + 3[w(b) - w(a)]w(a)}}{6[w(b) - w(a)]}.$$

iv - We note again that to confirm D-optimality of the above designs, we only need to check the equivalence theorem along the four edges of the parallelograms. We have again produced **'four curve'** plots of $\mathcal{V}(z_1)$ in Figure (4.13), Figure (4.14) and Figure (4.15).

4.2.6 Conjectured D-optimal designs for polygonal $\mathcal{Z}$

The first observation above supports the following assertion for any polygonal $\mathcal{Z} = [A, B]$.

Theorem 4.1. *Suppose design space $\mathcal{Z}$ is a polygon and that the function $H(z_1) = \frac{-w(z_1)}{[w(z_1)]^2}$, $H(-\infty) = -\infty$, $H(\infty) = \infty$ and $H(z)$ is concave increasing up to some point and thereafter is convex increasing. Then, there can be at most 2 support points along any edge of design space $\mathcal{Z}$.*

4.2.7 Proof

First, we note that if an edge is vertical, support points can only be at its end-points since the minimal ellipsoid centered on the origin containing $\mathcal{Z}$ could not touch the edge internally.

Any other edge must be defined by a linear equation, say $z_2 = S(z_1) = mz_1 + c$ over some range of values $A \leq z_1 \leq B$. Then if a design ξ^* is to be D-optimal on $\mathcal{Z}$, the equivalence theorem must be satisfied along this edge.

According to the theorem the design ξ^* is D-optimal iff

$$w(z_1)(1,\ z_1,\ S(z_1))M^{-1}(\xi^*)\begin{pmatrix}1\\ z_1\\ S(z_1)\end{pmatrix} \leq 3 \quad \forall \quad A \leq z_1 \leq B \qquad (4.5)$$

$$= 3 \quad \text{if} \quad \xi(z_1,\ S(z_1)) > 0 \qquad (4.6)$$

where $S(z_1) = mz_1 + c$.

Let

$$\begin{aligned}\mathcal{V}(z_1) &= w(z_1)(1,\ z_1,\ S(z_1))M^{-1}(\xi^*)(1,\ z_1,\ S(z_1))^T\\ \mathcal{V}(z_1) &= w(z_1)Q(z_1)\end{aligned}$$

where $Q(z_1) = (1,\ z_1,\ S(z_1))M^{-1}(\xi^*)(1,\ z_1,\ S(z_1))^T$. So the Equivalence theorem is satisfied iff

$$\begin{aligned}\mathcal{V}(z_1) &\leq 3 \qquad \forall \quad A \leq z_1 \leq B\\ &= 3 \qquad if \quad \xi(z_1, S(z_1)) > 0.\end{aligned}$$

This is true iff

$$\begin{aligned}v(z_1) &\leq 0 \qquad \forall \quad A \leq z_1 \leq B\\ &= 0 \qquad if \quad \xi(z_1, S(z_1)) > 0.\end{aligned}$$

where $v(z_1) = \frac{1}{3}Q(z_1) - \frac{1}{w(z_1)}$. **So for an optimal design we wish to see** $\boldsymbol{v(z_1) \leq 0\ \forall\ z_1 \in \mathcal{Z} = [A, B]}$. To explore the shape of $v(z_1)$ we analyze the derivative of $v(z_1)$. This can be written as follows:

$$\begin{aligned} v'(z_1) &= \frac{1}{3}Q'(z_1) - \{\frac{-w'(z_1)}{[w(z_1)]^2}\} \\ &= L(z_1) - H_3(z_1) \end{aligned}$$

where $H_3(z_1) = \frac{-w'(z_1)}{[w(z_1)]^2}$ and $L(z_1) = \frac{1}{3}Q'(z_1)$. Now

$$\begin{aligned} Q(z_1) &= (1,\, z_1,\, S(z_1))M^{-1}(\xi)(1,\, z_1,\, S(z_1))^T \\ &= [(0,\, z_1,\, mz_1) + (1,\, 0,\, c)]M^{-1}(\xi)[(0,\, z_1,\, mz_1) + (1,\, 0,\, c)]^T \\ &= [z_1(0,\, 1,\, m) + (1,\, 0,\, c)]M^{-1}(\xi)[z_1(0,\, 1,\, m) + (1,\, 0,\, c)]^T \\ &= (\underline{a}z_1 + \underline{b})M^{-1}(\xi)(\underline{a}z_1 + \underline{b})^T \end{aligned}$$

where $\underline{a} = (0,\, 1,\, m)^T$ and $\underline{b} = (1,\, 0,\, c)^T$. Hence $Q(z_1)$ is the quadratic function:

$$Q(z_1) \;=\; \underline{a}^T M^{-1}(\xi)\underline{a}z_1^2 + 2(\underline{a}^T M^{-1}(\xi)\underline{b})z_1 + \underline{b}^T M^{-1}(\xi)\underline{b}$$

Since the coefficient of z_1^2, $(\underline{a}^T M^{-1}(\xi)\underline{a})$, is positive, $L(z_1)$ is an increasing line. As a result, $v'(z_1) = 0$ iff $L(z_1) = H_3(z_1)$. That is, $v'(z_1) = 0$ when the line $L(z_1)$ crosses $H_3(z_1)$.

A question of interest is **"How many times can an increasing line $L(z_1)$ cross the function $H_3(z_1)$?"**

Recall that, $H_3(z_1) \propto H(z_1)$ [chapter 2 equation 2.3 and chapter 3, equation 3.3]. The similarity between these H functions leads to similar conclusions, namely that $H_3(-\infty) = -\infty$, $H_3(\infty) = \infty$ and H_3 is concave increasing up to z_{max} (or on$[-\infty, z_{max}]$) then convex increasing if $H(z)$ possesses these properties. So, as for the two-parameter case, $v(z_1)$ can have, **over the real line**, at most 3 TP's. Moreover since $L(z_1)$ first crosses H_3 from above the first TP, and hence the third, are maximal TP's, leaving the middle one as a minimal TP. Clearly, any number of these or none of them may occur in a particular line segment [A,B]. The various possibilities are depicted in Figure 4.16. These confirm that $v(z_1)$ can be zero at at most two points in $[A, B]$.

Since $v(z_1) = 0$ at support points the theorem is proved.

We do not have recommendations to make about where support points might lie along an edge except that they can only be at endpoints of vertical edges. Also the minimal ellipsoid characterization may shed light on where support points lie.

Clearly it would be possible to make progress in special cases like the symmetrical $\mathcal{Z}$ of Example 4.2. There must be a solution in these cases along the lines of Theorem (2.2).

However we settle for reporting results in some asymmetrical examples. These further support the theorem.

4.3 Some More Examples

Example 4.3. *Now suppose $\mathcal{Z}$ is the trapezium with vertices $(-2,-1)$,$(-1,1)$, $(1,1)$, $(2,-1)$. [See Figure 4.5]. Again first consider the design with the four corners as support points, namely*

$$\xi = \begin{pmatrix} z_{1i} & -b & -a & a & b \\ z_{2i} & -d & -c & c & d \\ p_i & q & p & p & q \end{pmatrix}$$

where $b = 2$, $a = 1$, $d = -1$, $c = 1$ and $i = 1, \cdots, 4$. We note that the symmetry about $z_1 = 0$ justifies assuming two pairs of equal weights. Correspondingly we again have an explicit solution for these weights ***given a,b,c,d***. *The determinant $|M(\xi)|$ is*

$$|M(\xi)| = 8(d-c)^2 qw(b)pw(a)[a^2pw(a) + b^2qw(b)].$$

In this case $|M(\xi)|$, as a function of q and p, is independent of c, d. So the optimal weights depend only on a and b. The solution is

$$q = \frac{[b^2w(b) - 2a^2w(a)] + \sqrt{[b^2w(b) - 2a^2w(a)]^2 + 3[b^2w(b) - a^2w(a)]a^2w(a)}}{6[b^2w(b) - a^2w(a)]}.$$

In fact the solution is identical to that in equation 6.6 (chapter 3, $q(r)$) but with $r = \frac{[b^2w(b)]}{[a^2w(a)]}$.

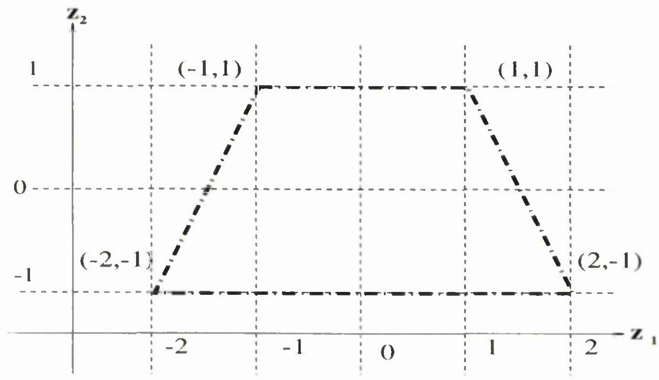


Figure 4.5: Trapezium design space I

This design proves to be non-optimal for $\mathcal{Z}$. Therefore we found the D-optimal design using the algorithm in section (4.2). This is summarized in Table (4.4). Additionally, we can see relevant variance functions in Figure 4.17. There are two support points along three edges, including two vertices.

Example 4.4. *Further, we will consider another trapezium with vertices* $(-2,-0.5)$ $(-2,0.5)$ $(2,-1)$ $(2,1)$. *[See Figure 4.6]. A symmetric design with the four corners as support points*

$$\xi = \begin{pmatrix} z_{1i} & a & a & b & b \\ z_{2i} & -c & c & -d & d \\ p_i & p & p & q & q \end{pmatrix}$$

where $a = -2$, $b = 2$, *differently from the above example,* $c = 0.5$, $d = 1$ *and* $i = 1, \cdots, 4$.

Note that the structure of the design is similar to the above. Hence there is an explicit solution for the optimal weights given the support points. First the determinant function for the design is

$$|M(\xi)| = 8(a-b)^2 pqw(a)w(b)[c^2w(a)p + d^2w(b)q].$$

The optimal solution is

$$q = \frac{[c^2w(a) - 2d^2w(b)] + \sqrt{[c^2w(a) - 2d^2w(b)]^2 + 3[c^2w(a) - d^2w(b)]d^2w(b)}}{6[c^2w(a) - d^2w(b)]}.$$

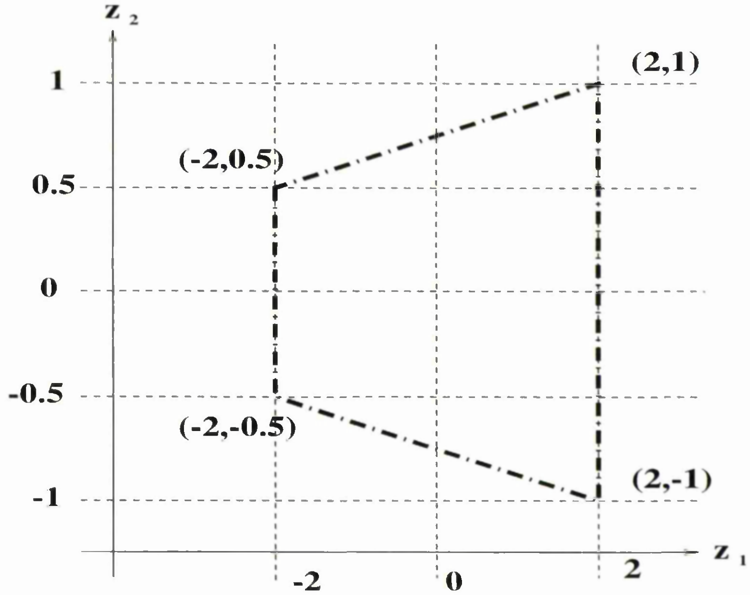


Figure 4.6: Trapezium design space II

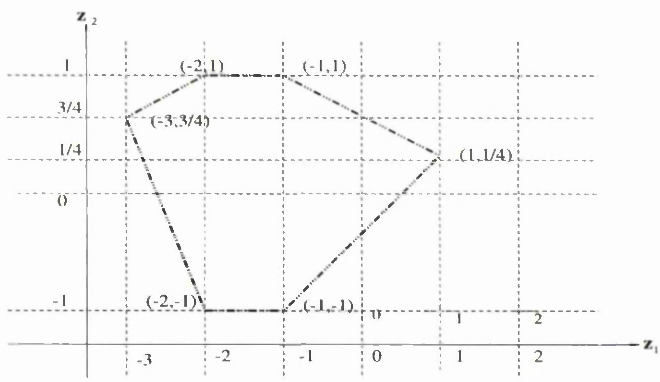


Figure 4.7: An arbitrary Polygon, 1

This now depends on c and d. In fact the solution is again identical to that in equation (3.6) [Chap 3, $q(r)$] but with $r = \dfrac{d^2 w(b)}{c^2 w(a)}$. Again the resultant design is not optimal for $\mathcal{Z}$. Using the algorithm we find the D-optimal design. This is summarized in Table (4.4). Also, the variance function can be seen from Figure (4.18). Results are similar to **Example (4.1)**. There are two support points on the bounds of z_2 (or; there are two support points on the top and bottom edge).

Example 4.5. *Now suppose $\mathcal{Z}$ is the polygon with vertices $(-2,-1), (-1,-1), (1,1/4), (-1,1), (-2,1), (-3,3/4)$ [See Figure 4.7]. For this example we found the D-optimal design using the algorithm. We summarize the*

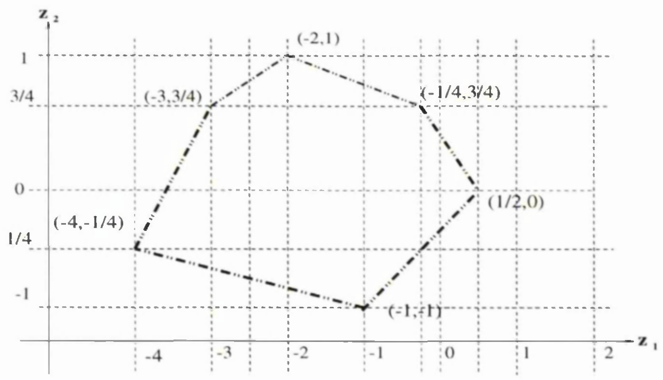


Figure 4.8: An arbitrary Polygon, 2

results in Table (4.4). There are three support points with equal weights. Two of them are on the vertices and one of them lies on the top horizontal edge. Along the six edges it can be seen from the Figure (4.19) that the equivalence theorem is satisfied only at these points. There are two edges with no support points, there are three edges with one and one edge with two. Of course, vertices are at the end of two lines. So there is at most two support points along any edge.

Example 4.6. *In this example, we consider the design space $\mathcal{Z}$ to be the polygon with vertices $(-1,-1), (1/2,0), (-1/4,3/4), (-2,1), (-3,3/4), (-4,1/4)$ [See Figure 4.8]. In the same way as in the preceding example we applied the algorithm to find the optimal design. The D-optimal design has four support points at vertices. See Table 4.4 and Figure 4.20. There is at most two support points along any edge, including vertices.*

Example 4.7. *Lastly, we consider any polygon containing the supports on $\mathcal{Z}_w$ then the global optimal design must still be optimal and this only has support points along the horizantal edges of $\mathcal{Z}$, two on each of these; [See Figure 4.9].*

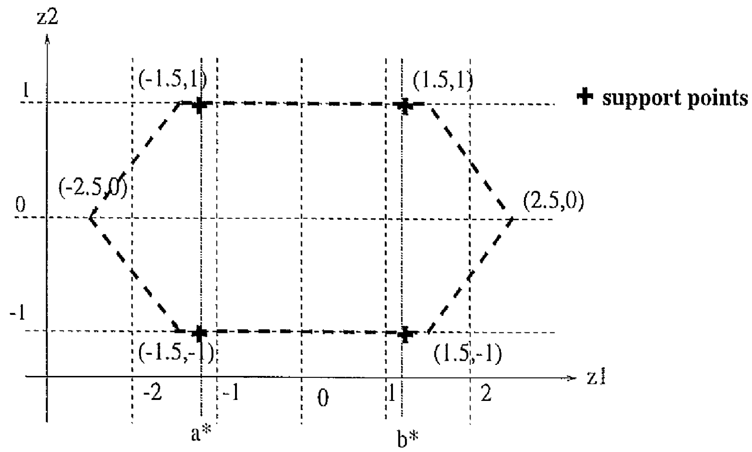


Figure 4.9: Polygon containing supports on $\mathcal{Z}_w$

Three parameter case: Logistic weight function, polygon design space optimal four-corner design with optimal weights							
VERTICES OF DESIGN REGION					SUPPORT POINTS		D-OPTIMAL WEIGHTS
	1	1	1	1			
z_1	−2	−1	1	2	z_1	z_2	p_j
z_2	−1	−1	1	1			
					-2.0000	-1.0000	0.214346
					-1.0000	-1.0000	0.285654
					1.0000	1.0000	0.285654
					2.0000	1.0000	0.214346
	1	1	1	1			
z_1	−2	−1	1	2	z_1	z_2	p_j
z_2	1	1	−1	−1			
					-2.0000	1.0000	0.214346
					-1.0000	1.0000	0.285654
					1.0000	-1.0000	0.285654
					2.0000	-1.0000	0.214346
	1	1	1	1			
z_1	−2	−1	1	2	z_1	z_2	p_j
z_2	−1	1	−1	1			
					-2.0000	-1.0000	0.214346
					-1.0000	1.0000	0.285654
					1.0000	-1.0000	0.285654
					2.0000	1.0000	0.214346
	1	1	1	1			
z_1	−2	−1	1	2	z_1	z_2	p_j
z_2	1	−1	1	−1			
					-2.0000	1.0000	0.214346
					-1.0000	-1.0000	0.285654
					1.0000	1.0000	0.285654
					2.0000	-1.0000	0.214346
	1	1	1	1			
z_1	−2	−1	1	2	z_1	z_2	p_j
z_2	−1	1	1	−1			
					-2.0000	-1.0000	0.291533
					-1.0000	1.0000	0.208467
					1.0000	1.0000	0.208467
					2.0000	-1.0000	0.291533
	1	1	1	1			
z_1	−2	−1	1	2	z_1	z_2	p_j
z_2	1	−1	−1	1			
					-2.0000	1.0000	0.291533
					-1.0000	-1.0000	0.208467
					1.0000	-1.0000	0.208467
					2.0000	1.0000	0.291533

Table 4.1: For Logistic Regression model D-optimal support points and weights.

Three parameter case: For Logistic weight Function, a design region for z_1,z_2 in the form of a parellogram: D-optimal weights and support points.				
VERTICES OF DESIGN REGION	SUPPORT POINTS		D-OPTIMAL WEIGHTS	VARIANCE FUNCTION
$\begin{matrix} & 1 & 1 & 1 & 1 \\ z_1 & -2 & -1.2 & 1.2 & 2 \\ z_2 & -1 & -1 & 1 & 1 \end{matrix}$	z_1	z_2	p_j	$\mathcal{V}(z_1)$
	-1.2000	-1.0000	0.2509	3.0000
	1.2275	0.5172	0.2492	3.0000
	-1.2275	-0.5172	0.2492	3.0000
	1.2000	1.0000	0.2509	3.0000
$\begin{matrix} & 1 & 1 & 1 & 1 \\ z_1 & -2 & -1 & 1 & 2 \\ z_2 & -1 & -1 & 1 & 1 \end{matrix}$	z_1	z_2	p_j	$\mathcal{V}(z_1)$
	-1.0000	-1.0000	0.2622	3.0000
	1.3773	0.5821	0.2378	3.0000
	-1.3773	-0.5821	0.2378	3.0000
	1.0000	1.0000	0.2622	3.0000
$\begin{matrix} & 1 & 1 & 1 & 1 \\ z_1 & -2 & -0.75 & 0.75 & 2 \\ z_2 & -1 & -1 & 1 & 1 \end{matrix}$	z_1	z_2	p_j	$\mathcal{V}(z_1)$
	-0.7500	-1.0000	0.2730	3.0000
	1.5040	0.6393	0.2270	3.0000
	-1.5040	-0.6393	0.2270	3.0000
	0.7500	1.0000	0.2730	3.0000
$\begin{matrix} & 1 & 1 & 1 & 1 \\ z_1 & -2 & -0.5 & 0.5 & 2 \\ z_2 & -1 & -1 & 1 & 1 \end{matrix}$	z_1	z_2	p_j	$\mathcal{V}(z_1)$
	-0.5000	-1.0000	0.2838	3.0000
	1.7004	0.7603	0.2162	3.0000
	-1.7004	-0.7603	0.2162	3.0000
	0.5000	1.0000	0.2838	3.0000

Table 4.2: For Logistic weight function D-optimal support points and weights.

Three parameter case: For Logistic weight Function, a design region for z_1,z_2 in the form of a parellogram: D-optimal weights and support points.

Vertices of Design Design					Support points		D-optimal weights	Variance function
	1	1	1	1				
z_1	-2	-0.25	0.25	2	z_1	z_2	p_j	$\mathcal{V}(z_1)$
z_2	-1	-1	1	1				
					-0.2500	-1.0000	0.2916	3.0000
					1.8767	0.8904	0.2085	3.0000
					-1.8767	-0.8904	0.2085	3.0000
					0.2500	1.0000	0.2916	3.0000
	1	1	1	1				
z_1	-2	0	0	2	z_1	z_2	p_j	$\mathcal{V}(z_1)$
z_2	-1	-1	1	1				
					-2.0000	-1.0000	0.2040	3.0000
					0	-1.0000	0.2959	3.0000
					0	1.0000	0.2959	3.0000
					2.0000	1.0000	0.2040	3.0000
	1	1	1	1				
z_1	-2	0.25	-0.25	2	z_1	z_2	p_j	$\mathcal{V}(z_1)$
z_2	-1	-1	1	1				
					-1.8767	-1.0000	0.2085	3.0000
					0.2500	-1.0000	0.2916	3.0000
					-0.2500	1.0000	0.2916	3.0000
					1.8767	1.0000	0.2085	3.0000
	1	1	1	1				
z_1	-2	0.5	-0.5	2	z_1	z_2	p_j	$\mathcal{V}(z_1)$
z_2	-1	-1	1	1				
					-1.7004	-1.0000	0.2162	3.0000
					0.5000	-1.0000	0.2838	3.0000
					-0.5000	1.0000	0.2838	3.0000
					1.7004	1.0000	0.2162	3.0000

Table 4.3: For Logistic weight function D-optimal support points and weights.

Three parameter case: For Logistic weight Function, a design region for z_1,z_2 in the form of a polygon: D-optimal weights and support points.

Vertices of design region							Support points		D-optimal weights	Variance function
	1	1	1	1						
z_1	−2	−1	1	2			z_1	z_2	p_j	$\mathcal{V}(z_1)$
z_2	−1	1	1	−1						
							-1.2000	1.0000	0.2258	3.0000
							1.2000	1.0000	0.2258	3.0000
							-1.1363	-1.0000	0.2741	3.0000
							1.1363	-1.0000	0.2741	3.0000
	1	1	1	1						
z_1	−2	−2	2	2			z_1	z_2	p_j	$\mathcal{V}(z_1)$
z_2	−0.5	0.5	−1	1						
							-1.2829	0.5896	0.2040	3.0000
							1.2829	0.9104	0.2960	3.0000
							-1.2829	-0.5896	0.2040	3.0000
							1.2829	-0.9104	0.2960	3.0000
	1	1	1	1	1	1				
z_1	−2	−1	1	−1	−2	−3	z_1	z_2	p_j	$\mathcal{V}(z_1)$
z_2	−1	−1	1/4	1	1	3/4				
							-1.0000	-1.0000	0.3333	3.0000
							1.0000	0.2500	0.3333	3.0000
							-1.3000	1.0000	0.3333	3.0000
	1	1	1	1	1	1				
z_1	−1	1/2	−1/4	−2	−3	−4	z_1	z_2	p_j	$\mathcal{V}(z_1)$
z_2	−1	0	3/4	1	3/4	1/4				
							-1.9500	1.0000	0.2894	3.0000
							-0.2500	0.7500	0.1127	3.0000
							0.5000	0.0000	0.2718	3.0000
							-1.0000	-1.0000	0.3260	3.0000

Table 4.4: The trapezium and arbitrary polygon 1 and polygon 2. D-optimal support points and weights for logistic weight function.

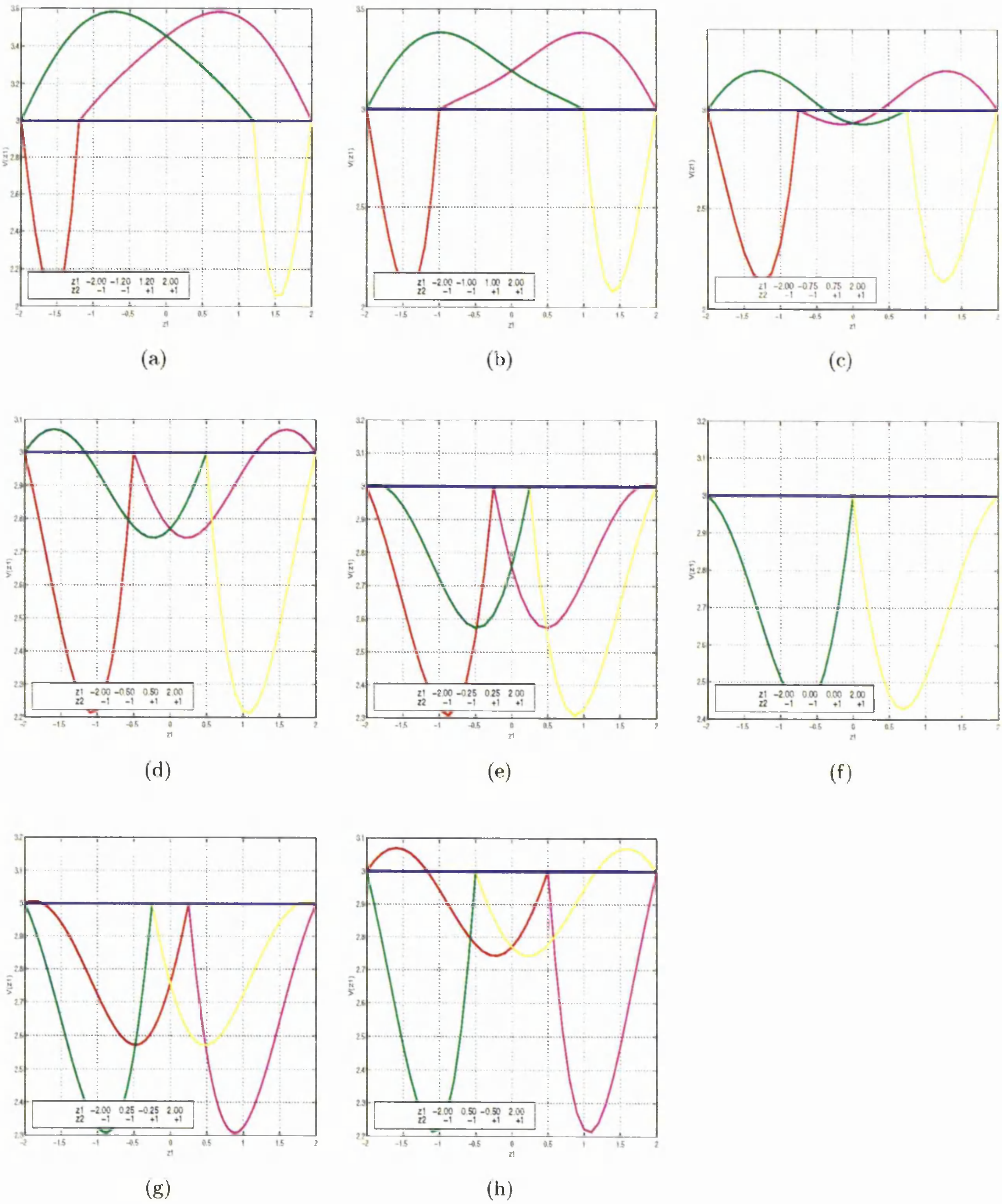
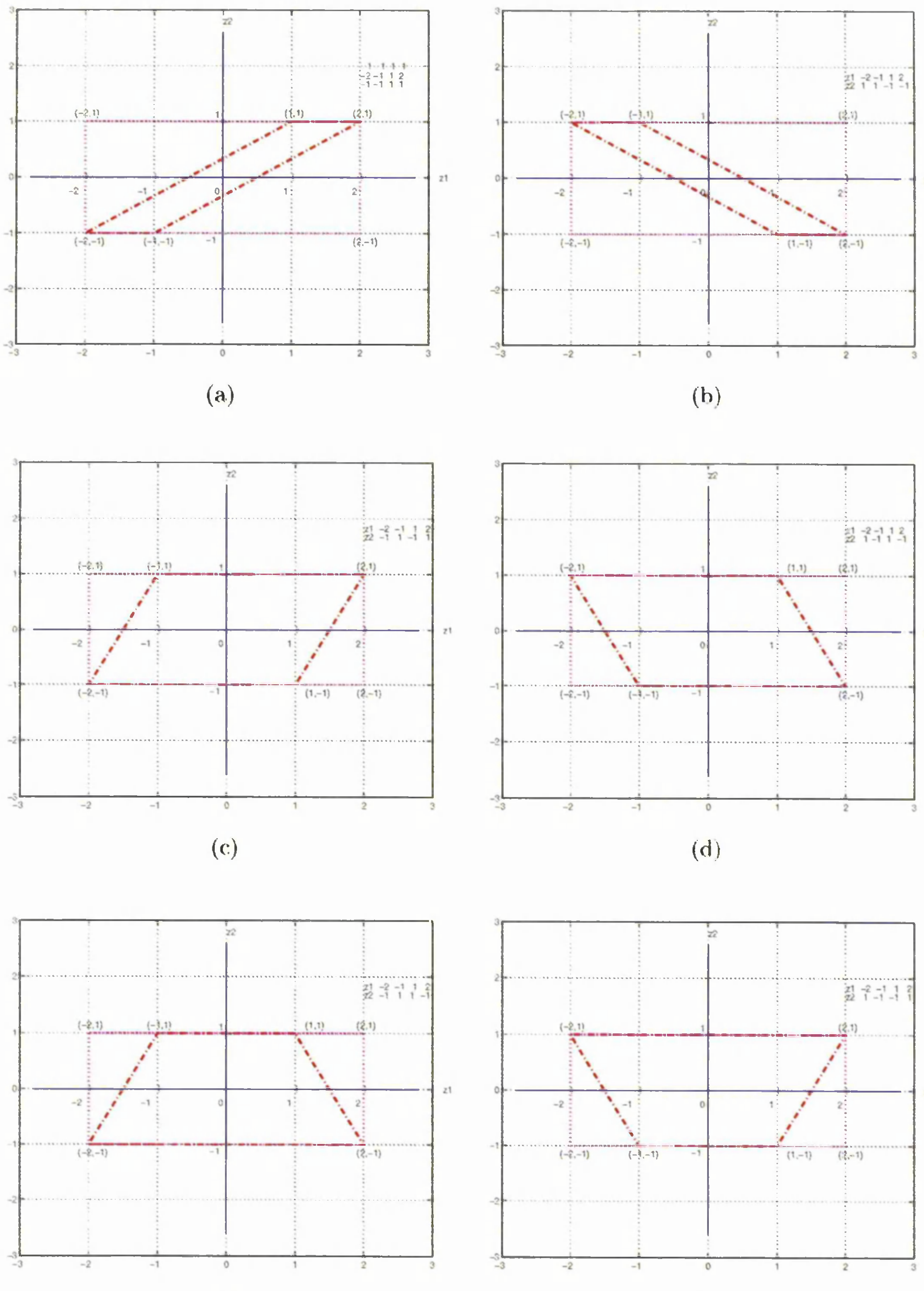


Figure 4.10: To check **'four corner'** design Variance Function for different a value $b = 2$.



(a) (b) (c) (d) (e) (f)

Figure 4.11: Design Region for the case of two design variables (z_1, z_2) using the Logistic weight function in the form of parallelogram.

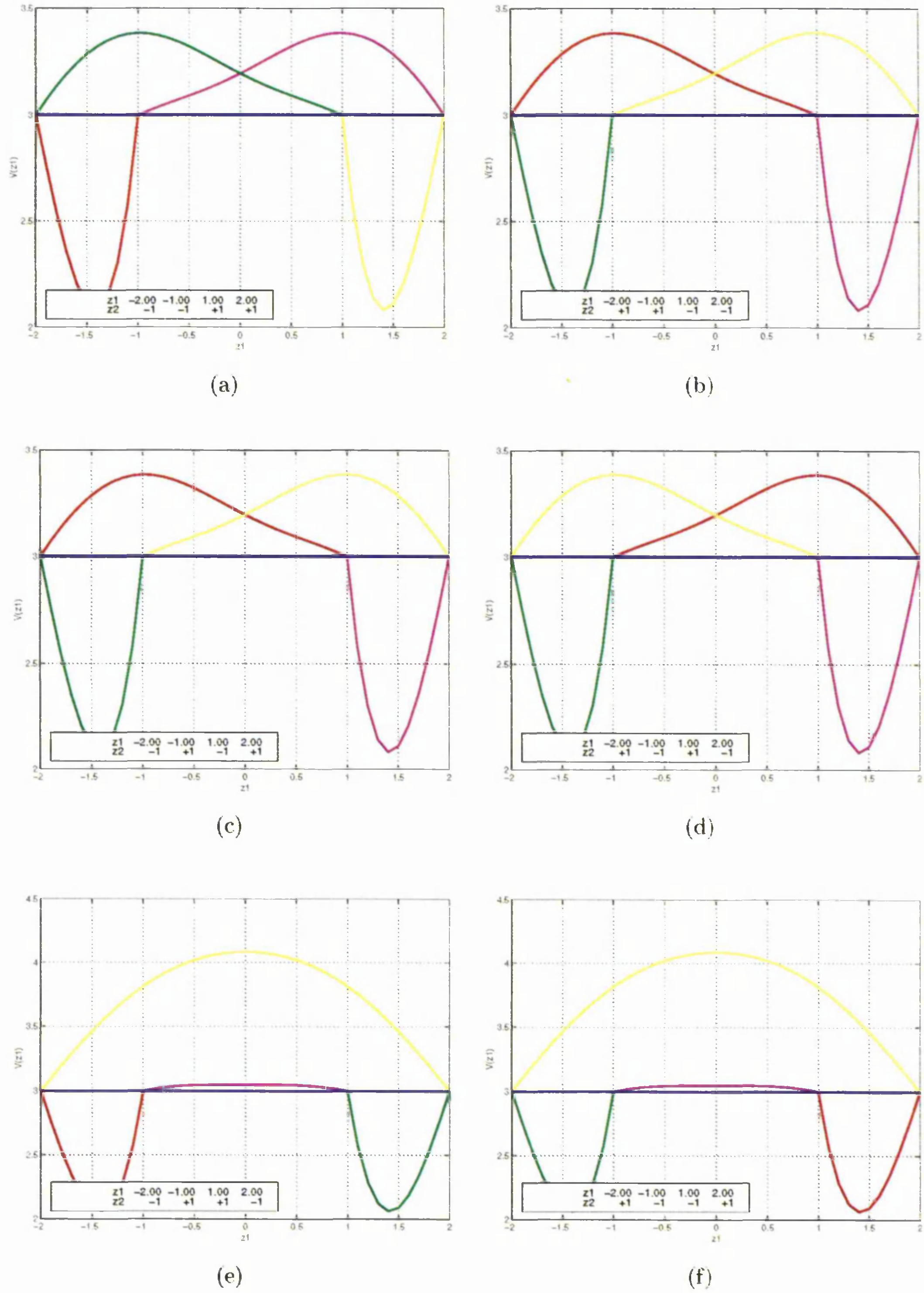


Figure 4.12: Plots of the Variance function for 6 different parallelograms.

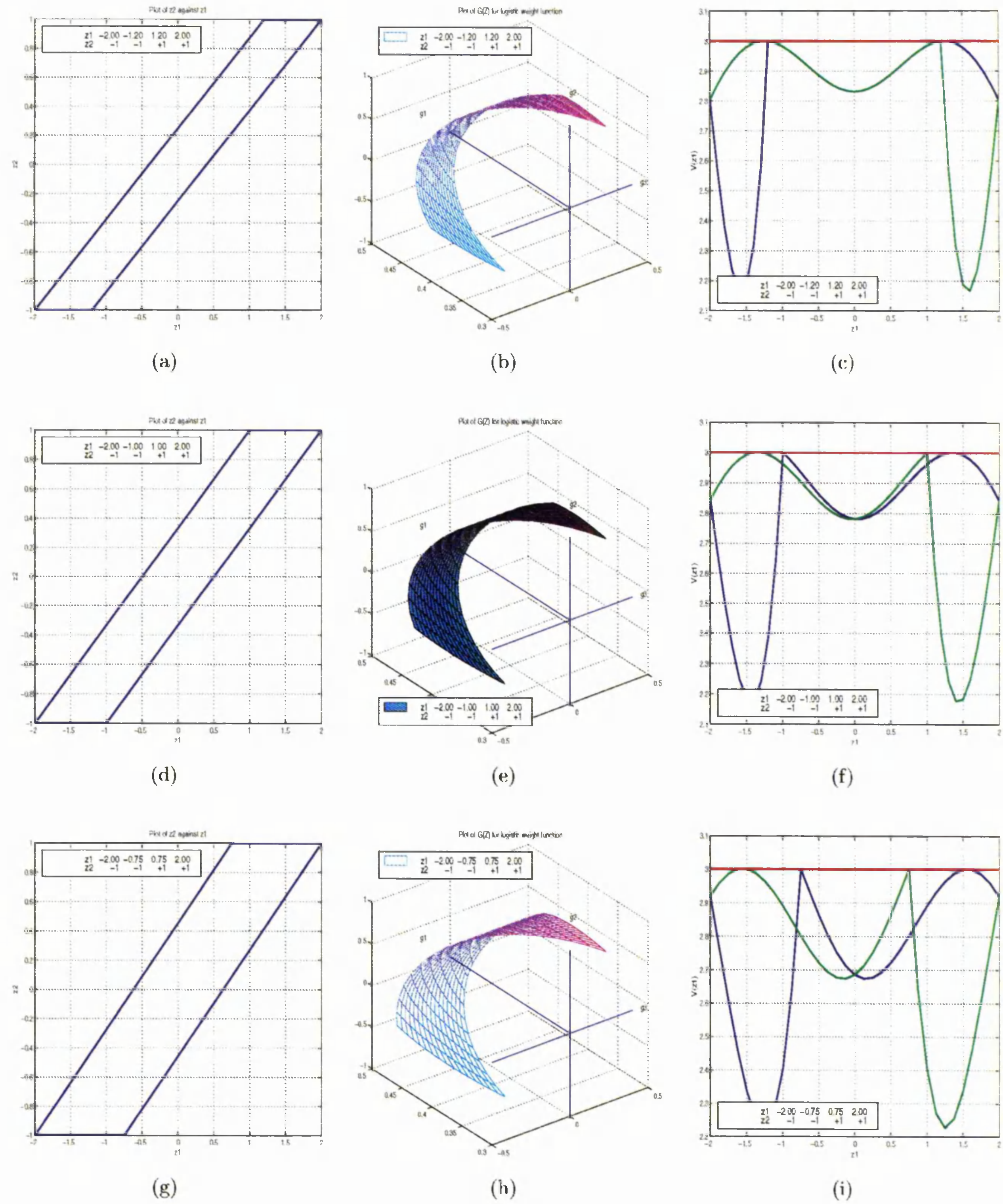


Figure 4.13: Plots of the variance function for a D-optimal design on a parallelogram and Design space for the Logistic weight function.

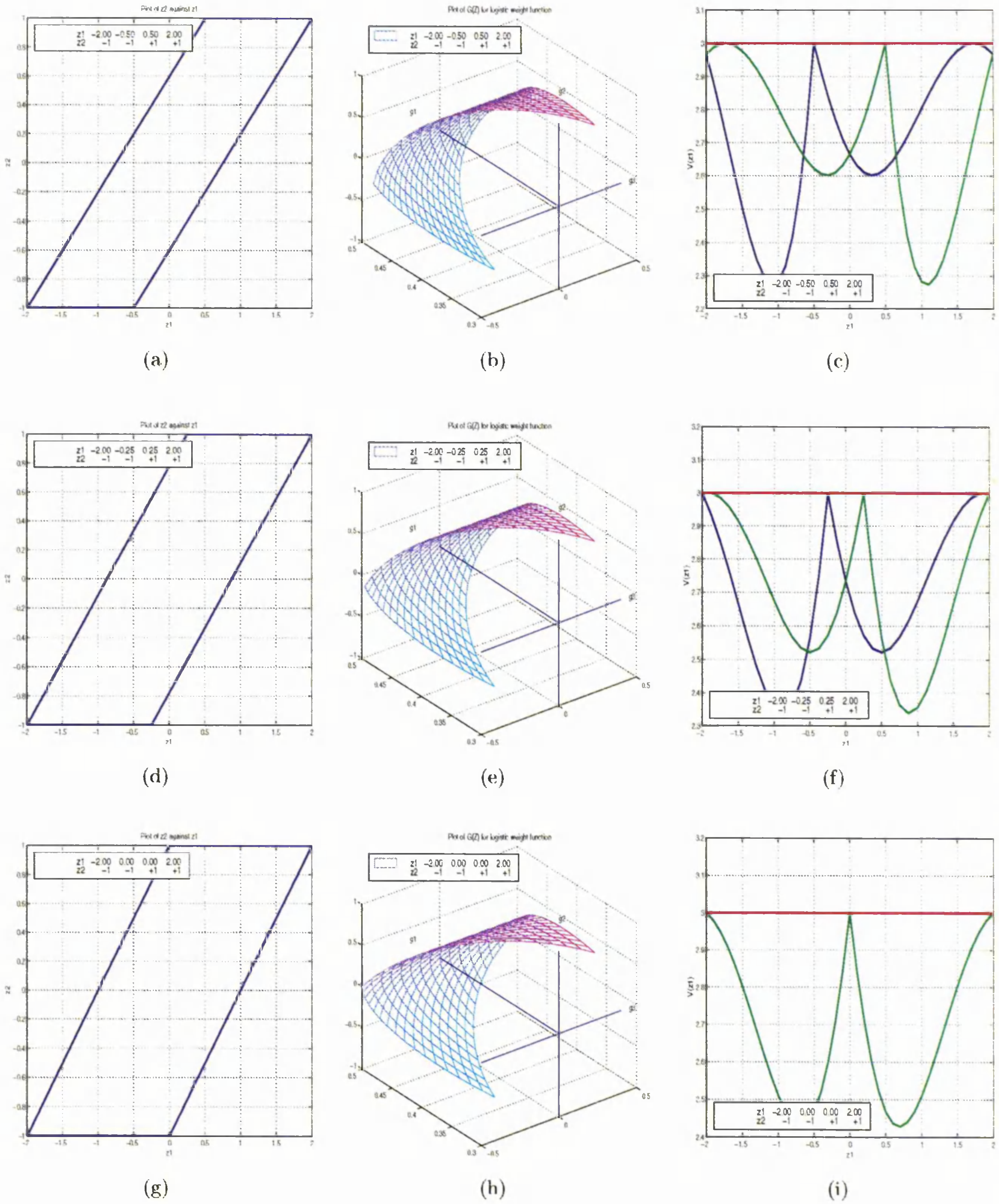


Figure 4.14: Plots of the variance function for a D-optimal design on a parallelogram and Design space for the Logistic weight function.

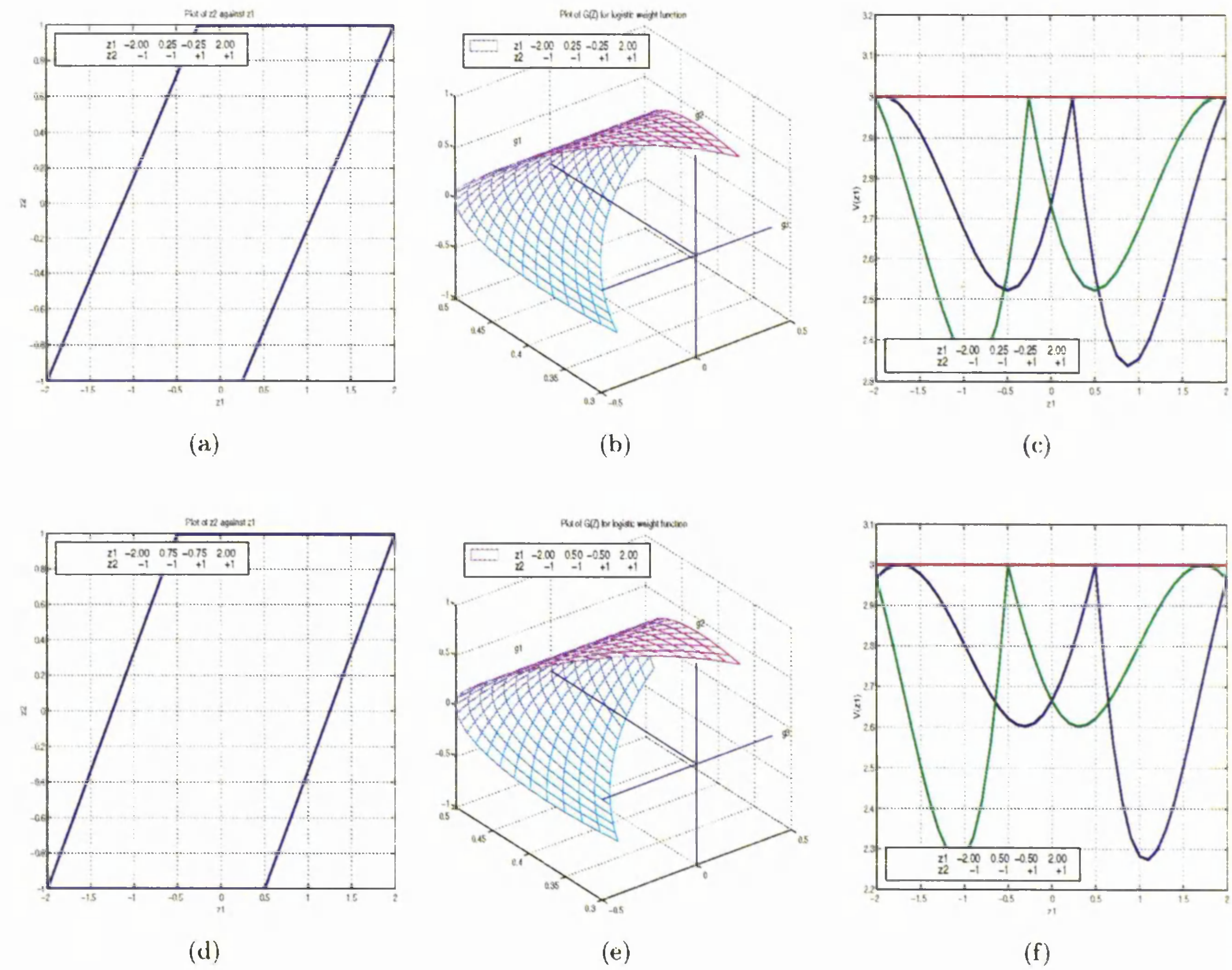


Figure 4.15: Plots of the variance function for a D-optimal design on a parallelogram and Design space for the Logistic weight function.

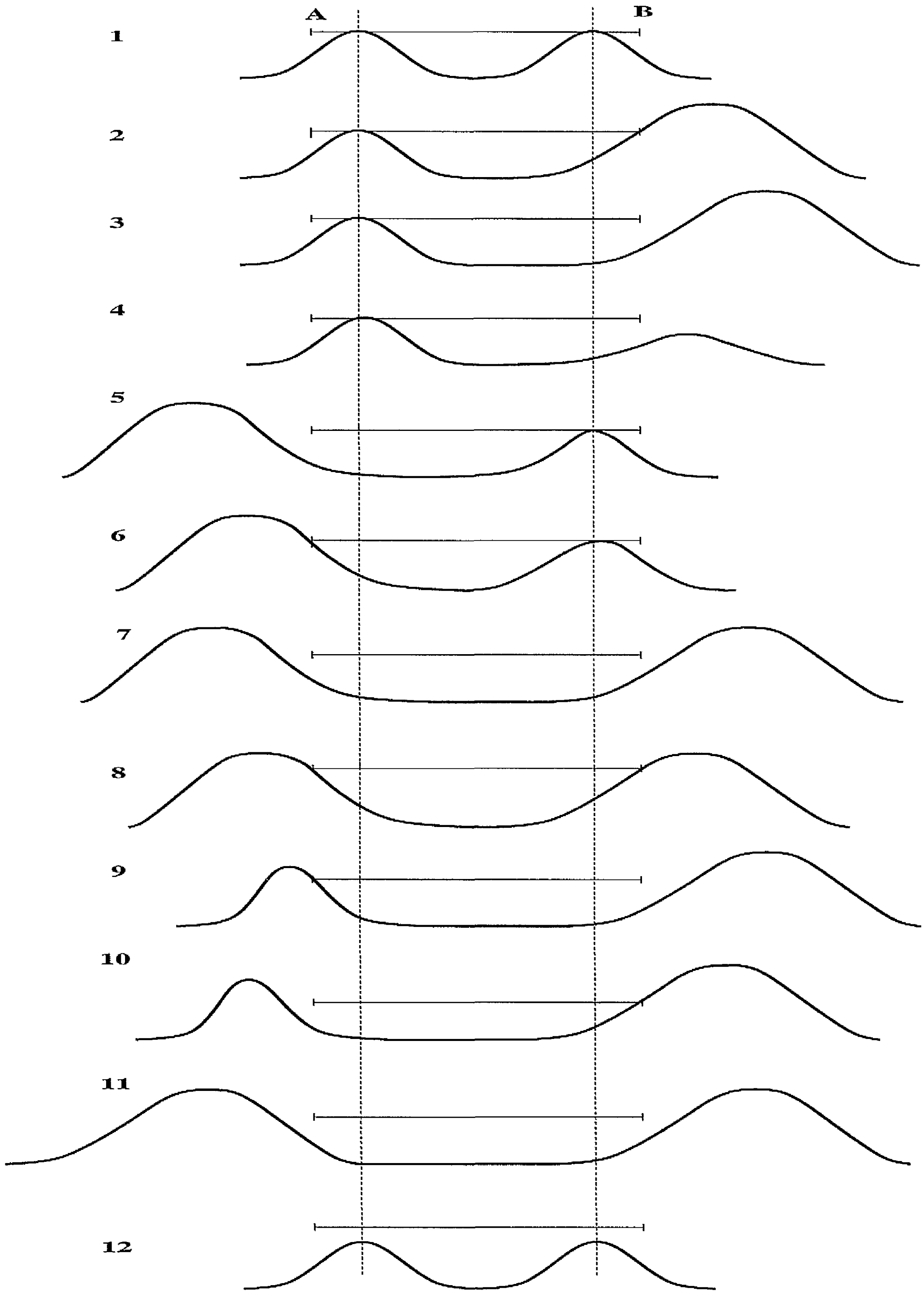


Figure 4.16: Possible shapes for the variance function $\mathcal{V}(z_1)$ under an optimal design.

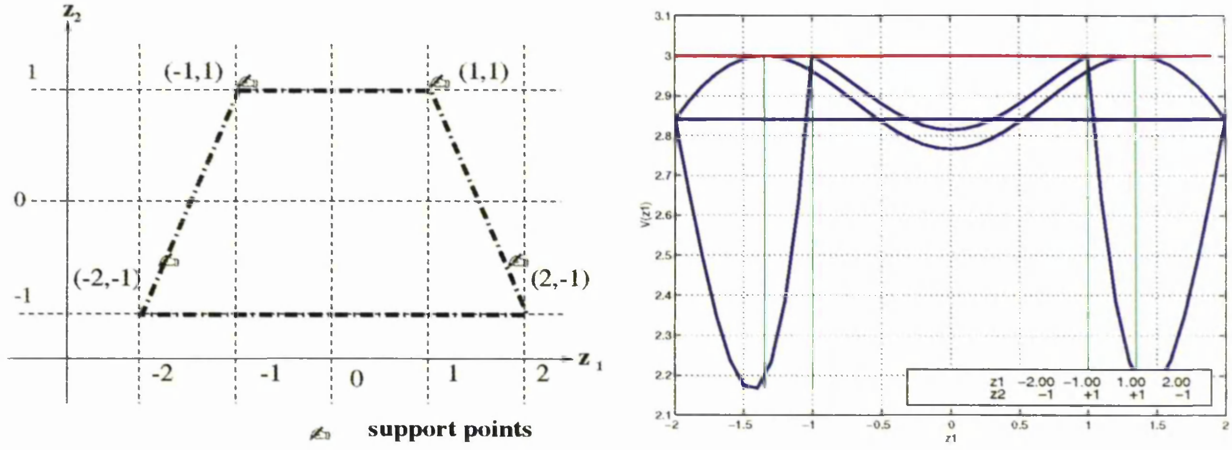


Figure 4.17: Plot of the variance function for a D-optimal design on arbitrary trapezium I for the logistic weight function.

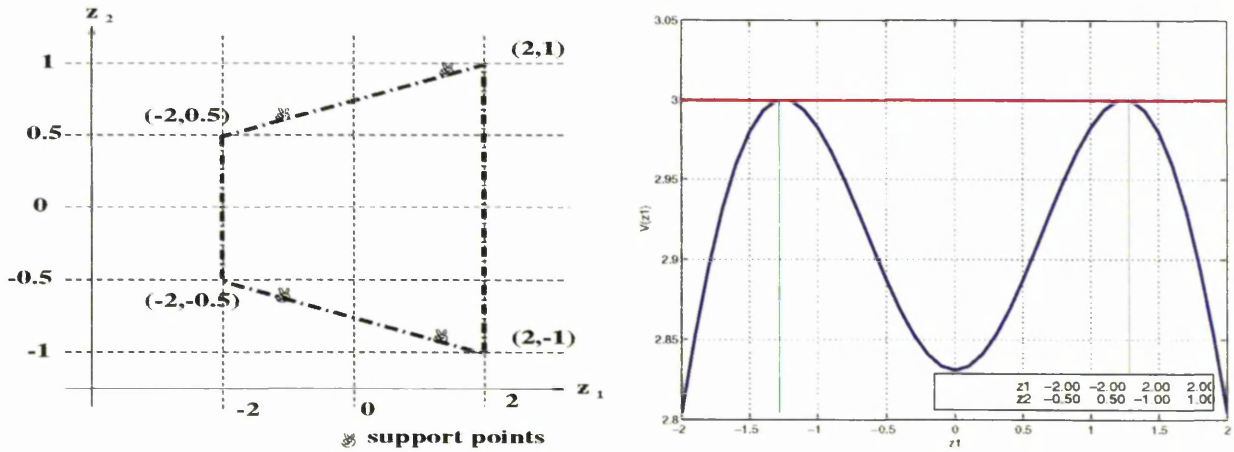


Figure 4.18: Plot of the variance function for a D-optimal design on arbitrary trapezium II for the logistic weight function.

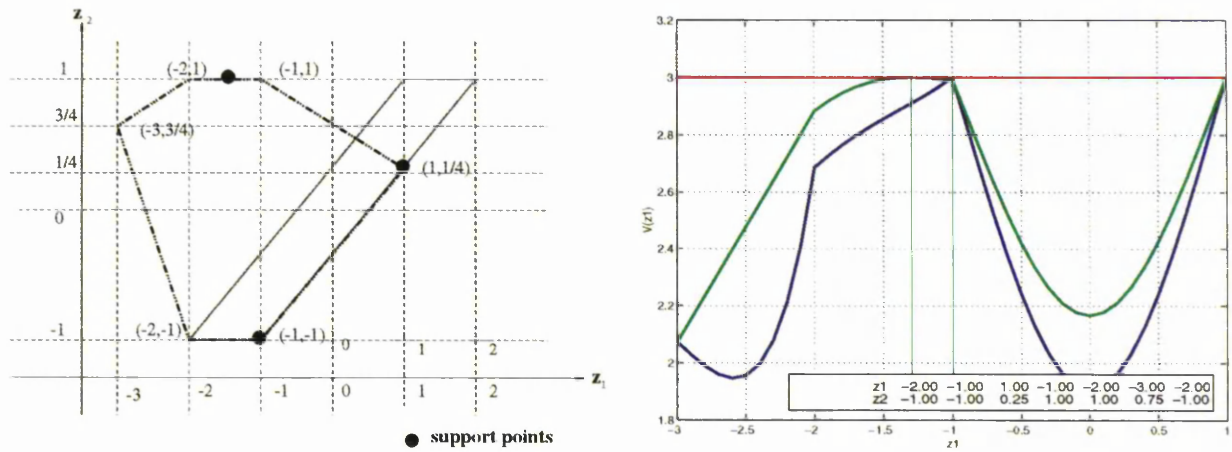


Figure 4.19: Plot of the variance function for a D-optimal design on arbitrary polygon 1 for the logistic weight function.

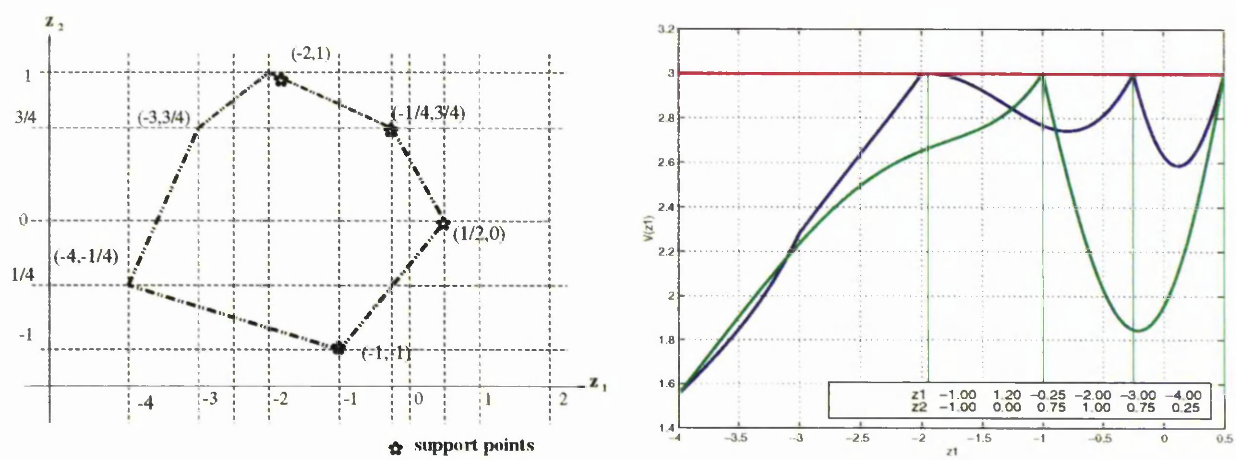


Figure 4.20: Plot of the variance function for a D-optimal design on arbitrary polygon 2 for the logistic weight function.

Chapter 5

Construction of D-optimal Design For Multiple Parameter Weighted Regression Model

5.1 Multiparameter Binary Regression

In multiple binary regression, we generally consider a model in which an observed value Y depends on a vector of $\underline{x}$ of l explanatory variables $\underline{x} = (x_1, \cdots, x_l)$ which are selected from a design space $\mathcal{X} \in R^l$. The outcome is binary, with probabilities

$$\mathbf{Pr}(Y = 0|\underline{x}) = 1 - \pi(\underline{x}) \qquad \mathbf{Pr}(Y = 1|\underline{x}) = \pi(\underline{x}).$$

Namely, $Y \sim Bi(1, \pi(\underline{x}))$. We deal with the relationship between the response probability $\pi(\underline{x})$ and the explanatory or design variables $\underline{x} = (x_1, \cdots, x_l)$. We assume $\pi(\underline{x}) = F(\alpha + \beta_1 x_1 + \cdots + \beta_l x_l)$, where $F(\cdot)$ is a cumulative distribution. So this is a GLM under which the dependence of π on $\underline{x} = (x_1, \cdots, x_l)$ is through the linear function

$$z_1 = \alpha + \beta_1 x_1 + \cdots + \beta_l x_l$$

for unknown parameters $\alpha, \beta_1, \cdots, \beta_l$. So

$$\begin{aligned} E(Y|\underline{x}) &= \pi(\underline{x}) = F(\alpha + \beta_1 x_1 + \cdots + \beta_l x_l) \\ V(Y|\underline{x}) &= \pi(\underline{x})[1 - \pi(\underline{x})] \end{aligned}$$

5.2 Design for k parameter Binary Regression

We now apply the theory of section 2.2-2.4 of Chapter 2 and section 3.2 of Chapter 3 to the multiparameter case. The informaton matrix for the above model is

$$I(x, \underline{\theta}) = \frac{f^2(z_1)}{F(z_1)[1 - F(z_1)]} \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_l \end{pmatrix} (1\ , x_1, \cdots, x_l)$$

where $f(z) = F'(z)$ and

$$\begin{aligned} \eta &= \pi(\underline{x}) \\ &= F(\alpha + \beta_1 x_1 + \cdots + \beta_l x_l), \qquad z_1 = \alpha + \beta_1 x_1 + \cdots + \beta_l x_l \\ &= F(z_1) \end{aligned}$$

and

$$\begin{aligned} a(\underline{x}, \underline{\theta}) &= V(Y|\underline{x}) \\ &= \pi(\underline{x})[1 - \pi(\underline{x})] \\ &= F(\alpha + \beta_1 x_1 + \cdots + \beta_l x_l)[1 - F(\alpha + \beta_1 x_1 + \cdots + \beta_l x_l)] \\ &= F(z_1)[1 - F(z_1)]. \end{aligned}$$

Also

$$\begin{aligned} \eta_\theta &= \left[\frac{\partial F(z_1)}{\partial z_1}\frac{\partial z_1}{\partial \alpha}, \frac{\partial F(z_1)}{\partial z_1}\frac{\partial z_1}{\partial \beta_1} \cdots \frac{\partial F(z_1)}{\partial z_1}\frac{\partial z_1}{\partial \beta_l}\right]^T \\ &= \left[f(z_1), f(z_1)x_1, \cdots, f(z_1)x_l\right]^T \\ &= f(z_1)\begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_l \end{pmatrix}. \end{aligned}$$

Now let the vector

$$\underline{v} = \frac{1}{\sqrt{V(Y|x)}}\left[\frac{\partial F(z_1)}{\partial z_1}\frac{\partial z_1}{\partial \alpha}, \frac{\partial F(z_1)}{\partial z_1}\frac{\partial z_1}{\partial \beta_1}\frac{\partial F(z_1)}{\partial z_1}\cdots\frac{\partial F(z_1)}{\partial z_1}\frac{\partial z_1}{\partial \beta_l}\right]^T$$

$$= \frac{f(z_1)}{\sqrt{F(z_1)[1-F(z_1)]}}\begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_l \end{pmatrix}.$$

Clearly, z_1 plays a similar role as $z = \alpha + \beta x$ and $z_1 = \alpha + \beta_1 x_1 + \beta_2 x_2$ in the two parameter case and three parameter case respectively.

We now consider the transformation

$$\begin{pmatrix} 1 \\ z_1 \\ z_2 \\ \vdots \\ z_l \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ \alpha & \beta_1 & \beta_2 & \cdots & \beta_l \\ b_{31} & b_{32} & b_{33} & \cdots & b_{3k} \\ \vdots & \vdots & \vdots & \vdots & \\ b_{l1} & b_{l2} & b_{l3} & \cdots & b_{lk} \end{pmatrix}\begin{pmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \\ x_l \end{pmatrix}$$

$$= \mathcal{B}\begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_l \end{pmatrix}$$

where $k = l + 1$ and b_{ij}, $i = 3, \cdots, l$ $j = 1, \cdots, k$ are arbitrary constants to be chosen by the experimenter. So we have defined further variables z_j, $j = 2, \cdots, l$. We have transformed to l new design variables $z_1, \cdots, z_l$. Their design space will be the image of $\mathcal{X}$ under the transformation. Denote this by $\mathcal{Z}$. Hence

$$\underline{g}(\underline{z}) = \mathcal{B}\underline{v}$$

$$= \frac{f(z_1)}{\sqrt{F(z_1)[1-F(z_1)]}}\begin{pmatrix} 1 \\ z_1 \\ \vdots \\ z_l \end{pmatrix}$$

and

$$\underline{v} \;=\; \mathcal{B}^{-1}\underline{g}(z)\ .$$

We consider the D-optimal linear design problem with design vectors

$$\underline{g} \;=\; \sqrt{w(z_1)}(1,\ z_1, \cdots, z_l)^T \qquad (z_1, \cdots, z_l)^T \in \mathcal{Z}$$

where $w(z_1) = \frac{f^2(z_1)}{F(z_1)[1-F(z_1)]}$, which corresponds to a weighted linear regression design problem with weight function $w(z_1)$.

With this transformation the design problem is equivalent to a weighted linear design problem with weight function $w(z_1) = \dfrac{f^2(z_1)}{F(z_1)[1-F(z_1]}$, where $f(z_1) = F'(z_1)$, is the density of $F(\cdot)$. Again it is useful to define the induced design space

$$G = G(\mathcal{Z}) \;=\; \{\underline{g_z} = (g_1, \cdots, g_k)^T : g_1 = \sqrt{w(z_1)},\ g_j = z_{j-1}\sqrt{w(z_1)},$$
$$j = 2, \cdots, k,\ \ z \in \mathcal{Z}\}.$$

5.3 Characterization of the Optimal Design

Let ξ^*, be a design measure on $\mathcal{Z}$. ξ^* is D-optimal iff

$$\underline{g}^T(\underline{z})M^{-1}(\xi^*)\underline{g}(\underline{z}) \;\leq\; k \qquad \xi^*(\underline{z}) = 0 \tag{5.1}$$
$$= \; k \qquad \xi^*(\underline{z}) > 0$$

where $\underline{g}^T(\underline{z}) = \sqrt{w(z_1)}(1, z_1, z_2, \cdots, z_l)$, and $k = l + 1$.

We once again resort to Silvey's minimal ellipsoid concept encountered in the previous sections. As we said earlier, G must be bounded. For most of our weight functions $g_1 = \sqrt{w(z_1)}$, $g_2 = z_1\sqrt{w(z_1)}$ are bounded for all z_1 but $g_j = z_{j-1}\sqrt{w(z_1)}$ $j = 3, \cdots, k$ will be unbounded if z_j is not restricted to a finite set. Without loss of generality we assume $-1 \leq z_j \leq 1,\ j = 2, \cdots, l$. So the largest possible G is

$$G = G(\mathcal{Z}) \;=\; \Big\{\underline{g_z} = (g_1, \cdots, g_k)^T : g_1 = \sqrt{w(z_1)}, \quad g_j = z_{j-1}\sqrt{w(z_1)},$$
$$-\infty < z_1 < \infty \quad -1 \leq z_j \leq 1 \quad j = 2, \cdots, k\Big\}.$$

We again wish to establish $D-$optimal designs for all possible design intervals for z_1. So also of interest is

$$G = G_{ab} \quad = \quad \Big\{ g_z = (g_1, \cdots, g_k)^T : g_1 = \sqrt{w(z_1)}, \quad g_j = z_{j-1}\sqrt{w(z_1)}, \\ z_1 \in [a,b], \quad -1 \le z_j \le 1 \quad j = 2, \cdots, k \Big\}.$$

From this geometrical consideration it is clear that such an ellipsoid will only pass through boundary points of G. This intersection can only occur at points where $z_j = \pm 1 \; j = 2, \cdots, l$. It was possible to see this in the three parameter case. Since G was a vertical surface, an ellipsoid centered at the origin which contains G could only make contact with it on the upper or lower ridges.

Case 1 : $\mathcal{Z} = \mathcal{Z}_w \; = \; \{(z_1 \cdots, z_l) : -\infty < z_1 < \infty, \, -1 \le z_j \le 1, \, j = 2, \cdots, l\}$

We consider $Z = Z_w$ initially for the Binary and Beta, Gamma, Normal weight functions.

The design space G induced by this rectangle is then an l dimensional hyper-planar object perpendicular to the (g_1, g_2) plane which tracks the trajectory defined by (g_1, g_2) over the range of z_1, and G is a closed region. The smallest central ellipsoid can only intersect G on its boundaries. Thus the D-optimal design must have support points on the boundary of Z.

Case 2 : $Z \; = \; Z_{ab} \; = \; \{(z_1 \cdots, z_l) \; : \; a \le z_1 \le b, \; -1 \le z_j \le 1 \quad j = 2, \cdots, l\}$.

We now consider the case $z_1 \in$ [a,b] so that

$$G = G_{ab} \quad = \quad \Big\{ (g_1, \cdots, g_k)^T : g_1 = \sqrt{w(z_1)} \quad g_j = z_{j-1} g_1 \\ z_1 \in [a,b], \quad -1 \le z_j \le 1, \quad j = 2, \cdots, k \Big\}.$$

This is the case of a subset of $G(Z_w)$, a 'vertical' portion of $G(Z)$. We have the same argument as in **Case 1**. Since G_{ab} is also a vertical hyper-planar object, the smallest ellipsoid, centred on the origin, containing it can only touch it on its boundaries and thus the D-optimal design must have support point on the boundary of Z_{ab}.

The next point is **How many support points are there?** We claim for **Case 2** that for many of our weight functions we will take observations at only **two** values of z_1 and that **one** optimal design consist of dividing the total weight at each of these values equally across the 2^{k-2} combinations $z_j = \pm 1$ $j = 2, \cdots, l$. Thus for any interval $[a, b]$ of z_1-values we are arguing that the design is of the form

$$\xi = \begin{pmatrix} i & 1 & 2 & 3 & \cdots & M & N & N+1 & N+2 & \cdots & L \\ z_{1i} & c & c & c & \cdots & c & d & d & d & \cdots & d \\ z_{2i} & -1 & 1 & -1 & \cdots & 1 & -1 & 1 & -1 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ z_{li} & -1 & 1 & -1 & \cdots & 1 & -1 & 1 & -1 & \cdots & 1 \\ p_i & p_c & p_c & p_c & \cdots & p_c & p_d & p_d & p_d & \cdots & p_d \end{pmatrix}$$

where $2^{k-2}(p_c + p_d) = 1$ and $M = 2^{k-2}$, $N = 2^{k-2} + 1$ and $L = 2^{k-1}$.

Let ξ^* denote the D-optimal design and let $Supp(\xi^*)$ denote the two values of z_1 at which observations are taken under ξ^*. Further, let a^*, b^* be their values on Z_w. We further assert that

$$\begin{aligned} Supp(\xi^*) &= \{a^*, b^*\} \qquad a < a^*, b > b^* \\ Supp(\xi^*) &= \{max\{a,\ a^*(b)\}, b\} \qquad a < a^*, b < b^* \\ Supp(\xi^*) &= \{a, min\{b,\ b^*(a)\}\} \qquad a > a^*, b > b^* \\ Supp(\xi^*) &= \{a, b\} \qquad a > a^*, b < b^* \end{aligned}$$

where $b^*(a)$ (along with p_d, p_a) maximises $det(M(\xi))$ with respect to d (over

$d \geq a$) where ξ is the design

$$\xi = \begin{pmatrix} i & 1 & 2 & 3 & \cdots & M & N & N+1 & N+2 & \cdots & L \\ z_{1i} & a & a & a & \cdots & a & d & d & d & \cdots & b \\ z_{2i} & -1 & 1 & -1 & \cdots & 1 & -1 & 1 & -1 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ z_{li} & -1 & 1 & -1 & \cdots & 1 & -1 & 1 & -1 & \cdots & 1 \\ p_i & p_a & p_a & p_a & \cdots & p_a & p_d & p_d & p_d & \cdots & p_d \end{pmatrix}$$

and $a^*(b)$ (along with p_c, p_b) maximises $det(M(\xi))$ with respect to c (over $c \leq b$) where ξ is the design

$$\xi = \begin{pmatrix} i & 1 & 2 & 3 & \cdots & M & N & N+1 & N+2 & \cdots & L \\ z_{1i} & c & c & c & \cdots & c & b & b & b & \cdots & b \\ z_{2i} & -1 & 1 & -1 & \cdots & 1 & -1 & 1 & -1 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ z_{li} & -1 & 1 & -1 & \cdots & 1 & -1 & 1 & -1 & \cdots & 1 \\ p_i & p_c & p_c & p_c & \cdots & p_c & p_b & p_b & p_b & \cdots & p_b \end{pmatrix}.$$

5.4 Justification of the Conjecture

To prove the above conjecture we need to check the Equivalence theorem. According to this a necessary and sufficent condition for a design $\xi(\underline{z})$ to be D-optimal is

$$w(z_1)(1,\ z_1, \cdots, z_l)M^{-1}(\xi)\begin{pmatrix} 1 \\ z_1 \\ \vdots \\ z_l \end{pmatrix} \leq k \quad \forall \quad z_1, \cdots, z_l \in Z \tag{5.2}$$

$$= k \quad \text{if} \quad \xi(z_1, \cdots, z_l) > 0. \tag{5.3}$$

We only need to check this for $z_j = \pm 1$ for $j = 2, \cdots, l$, in which case equation (5.2) and (5.3) imply

$$\begin{aligned} w(z_1)Q^\times(z_1) &\leq k \quad \forall \quad \underline{z} \in \mathcal{Z}^\times \\ &= k \quad \text{if} \quad \xi(\underline{z}) > 0 \end{aligned}$$

where $\mathcal{Z}^\times = \{\underline{z} \in \mathcal{Z} : z_j = \pm 1,\ j = 2, \cdots, l\}$. i.e.

$$\begin{aligned} v^\times(z_1) = Q^\times(z_1) - \frac{k}{w(z_1)} &\leq 0 \quad \forall \quad \underline{z} \in \mathcal{Z}^\times \\ &= 0 \quad \text{if} \quad \xi(\underline{z}) > 0 \end{aligned}$$

where $Q^\times(z_1) = (1,\ z_1,\ \pm 1, \cdots, \pm 1)M^{-1}(\xi)(1,\ z_1,\ \pm 1 \cdots, \pm 1)^T$, a quadratic function. **So for an optimal design we wish to see** $v^\times(z_1) \leq 0$ in the case $Z = \{(z_1, \cdots, z_l) : a \leq z_1 \leq b, -1 \leq z_j \leq 1,\ j = 2, \cdots, l\}$.

Now

$$\frac{\mathbf{d}v^\times(z_1)}{\mathbf{d}z_1} = L(z_1) - H_m^\times(z_1), \tag{5.4}$$

where $H_m^\times(z_1) = \frac{-k\, w'(z_1)}{[w(z_1)]^2}$ and $L(z)$ is an increasing linear function of z_1 because the coefficient of z_1 is the value of the second diagonal element of the inverse of the design matrix $M(\xi)$ which is positive definitive. Consequently, $\frac{\mathbf{d}v^\times(z_1)}{\mathbf{d}z_1} = 0$ iff $L(z_1) = H_m^\times(z_1)$. That is, $\frac{\mathbf{d}v^\times(z_1)}{\mathbf{d}z_1} = 0$ when the line $L(z_1)$ crosses $H_m^\times(z_1)$. The important point is that $H_m^\times(z_1) \propto H_3^\times(z_1) \propto H(z_1)$ (chapter 2 equation 2.3 and chapter 3). There is no difference in the shapes of these functions. Thus $L(z_1)$ can only cut $H_m^\times(z_1)$ at most three times in the case for most of our weight functions in the two and three parameter case.

So we have the same conclusion here : $H_m^\times(-\infty) = -\infty$, $H_m^\times(+\infty) = +\infty$ and $H_m^\times(z_1)$ is concave increasing up to some point and thereafter convex increasing. It follows that $v^\times(z_1)$ has at most 3 turning points on Z_w. Further, because $L(z_1)$ first crosses $H_m^\times(z_1)$ from above, $v^\times(z_1)$ has only one minimum turning point for the same reasons as before. Hence for these weight functions there are only two support points on each of the boundary hyperplanes $z_j = \pm 1$, $j = 2, \cdots, l$, being identified by two distinct values of z_1 with the weight at these split equally between $z_j = \pm 1$ $j = 2, \cdots, l$. We now need to determine these two values of z_1 and the optimal weights. In fact there is an explicit solution for these weights.

5.5 Definition of Weights

We consider a design of the form :

$$\begin{pmatrix} i & 1 & 2 & 3 & \cdots & M & N & N+1 & N+2 & \cdots & L \\ z_{1i} & a & a & a & \cdots & a & b & b & b & \cdots & b \\ z_{2i} & -1 & 1 & -1 & \cdots & 1 & -1 & 1 & -1 & \cdots & 1 \\ \cdots & & & & & & & & & & \\ z_{li} & -1 & 1 & -1 & \cdots & 1 & -1 & 1 & -1 & \cdots & 1 \\ p_i & p_a & p_a & p_a & \cdots & p_a & p_b & p_b & p_b & \cdots & p_b \end{pmatrix}$$

under which weight p_a is assigned to each of the $2^{(k-2)}$ design points with $z_1 = u$, $z_j = \pm 1$ $j = 2, \cdots, l$; $u = a, b$ and $M = 2^{k-2}$, $N = 2^{k-2} + 1$ and $L = 2^{k-1}$. So p_a, $p_b > 0$ and $2^{(k-2)}(p_a + p_b) = 1$.

The design matrix is

$$M(p) = \sum_{i}^{2^{k-1}} p_i \underline{g}_i \underline{g}_i^T$$

where

$$\underline{g}_i = \sqrt{w(z_{1i})}(1, z_{1i}, z_{2i}, \cdots, z_{li})^T \quad i = 1, 2, \cdots, 2^{k-1}.$$

Therefore,

$$M(p) = 2^{(k-2)} \begin{pmatrix} p_a w(a) + p_b w(b) & a p_a w(a) + b p_b w(b) & 0 & \cdots & 0 \\ a p_a w(a) + b p_b w(b) & b^2 p_a w(a) + b^2 p_b w(b) & 0 & \cdots & 0 \\ 0 & 0 & p_a w(a) + p_b w(b) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & p_a w(a) + p_b w(b) \end{pmatrix}.$$

The determinant is

$$|M(p)| = 2^{k(k-2)} p_a p_b w(a) w(b) (b-a)^2 [p_a w(a) + p_b w(b)]^{(k-2)}.$$

We need to choose a, b, p_a, p_b to maximize the determinant of $M(p)$. We can find an explicit solution for the weights: First, we get the log of the *determinant* function which is a concave function of $M(\cdot)$:

$$\begin{aligned}\ln|M(p)| &= k(k-2)\ln 2 + \ln p_a + \ln p_b + \ln w(a) + \ln w(b) + 2\ln(b-a) \\ &\quad +(k-2)\ln[p_a w(a) + p_b w(b)]\end{aligned}$$

Substituting for p_b from

$$2^{(k-2)}(p_a + p_b) = 1 \qquad \Leftrightarrow \qquad p_b = \frac{1}{2^{(k-2)}} - p_a,$$

the determinant function becomes

$$\begin{aligned}\ln|M(p)| &= k(k-2)\ln 2 + \ln p_a + \ln(\frac{1}{2^{(k-2)}} - p_a) + \ln w(a) + \ln w(b) + 2\ln(b-a) \\ &\quad +(k-2)\ln[p_a w(a) + (\frac{1}{2^{(k-2)}} - p_a)w(b)].\end{aligned}$$

$$\frac{\partial \ln|M(p)|}{\partial p_a} = \frac{1}{p_a} - \frac{1}{\frac{1}{2^{(k-2)}} - p_a} + \frac{(k-2)[w(a) - w(b)]}{[p_a w(a) + (\frac{1}{2^{(k-2)}} - p_a)w(b)]}$$

Further,

$$\frac{\partial \ln|M(p)|}{\partial p_a} = 0 \qquad if \qquad \frac{1}{p_a} - \frac{1}{\frac{1}{2^{(k-2)}} - p_a} + \frac{(k-2)[w(a) - w(b)]}{[p_a w(a) + (\frac{1}{2^{(k-2)}} - p_a)w(b)]} = 0$$

$$-k2^{2(k-2)}[w(a) - w(b)]p^2(a) + 2^{(k-2)}[(k-1)w(a) - (k+1)w(b)]p_a + w(b) = 0 \tag{5.5}$$

$$k2^{(k-2)}[w(a) - w(b)]p^2(a) - [(k-1)w(a) - (k+1)w(b)]p_a - \frac{1}{2^{(k-2)}}w(b) = 0 \tag{5.6}$$

$$p_a = \frac{A \pm \sqrt{\{A\}^2 + 4k2^{(k-2)}[w(a) - w(b)]w(b)}}{2k2^{(k-2)}[w(a) - w(b)]}$$

where

$$A = [(k-1)w(a) - (k+1)w(b)].$$

This is an explicit solution for the values of p_a that maximize $|M(p)|$. Of the above two roots our solution is given by the first root because the second root leads to negative weights, namely:

$$p_a = \frac{A + \sqrt{\{A\}^2 + 4k2^{(k-2)}[w(a) - w(b)]w(b)}}{2k2^{(k-2)}[w(a) - w(b)]} \tag{5.7}$$

where

$$A = [(k-1)w(a) - (k+1)w(b)].$$

Then $p_b = \frac{1}{2^{(k-2)}} - p_a$.

Further we can express the solution for p_a in terms of $r = r(z) = \frac{w(a)}{w(b)}$, namely:

$$p_a = q(r(z)) = \frac{(k-1)(r(z)-1) - 2 + \sqrt{(k-1)^2(r(z)-1)^2 + 4r(z)}}{2k2^{(k-2)}(r(z)-1)}. \tag{5.8}$$

5.6 Determination of support points

Still the design is

$$\begin{pmatrix} i & 1 & 2 & 3 & \cdots & M & N & N+1 & N+2 & \cdots & L \\ z_{1i} & a & a & a & \cdots & a & b & b & b & \cdots & b \\ z_{2i} & -1 & 1 & -1 & \cdots & 1 & -1 & 1 & -1 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ z_{li} & -1 & 1 & -1 & \cdots & 1 & -1 & 1 & -1 & \cdots & 1 \\ p_i & p_a & p_a & p_a & \cdots & p_a & p_b & p_b & p_b & \cdots & p_b \end{pmatrix}$$

where $M = 2^{k-2}$, $N = 2^{k-2} + 1$ and $L = 2^{k-1}$ and

$$\ln|M(p)| = k(k-2)\ln 2 + 2\ln(b-a) + \ln p_a + \ln p_b + \ln w(a) + \ln w(b)$$
$$+(k-2)\ln[p_a w(a) + p_b w(b)] \qquad a < b.$$

We now view this as a function of four sub-functions of a, namely $w(a)$, p_a, p_b and $A(a,b) = k(k-2)\ln 2 + 2\ln(b-a)$, so that

$$\begin{aligned}
\ln|M(p)| &= A(a,b) + \ln p_a + \ln p_b + \ln w(a) + \ln w(b) \\
&\quad +(k-2)\ln[p_a w(a) + p_b w(b)] \qquad a < b \\
&= F(A(a,b), w(a), w(b), p_a, p_b) \\
&= F
\end{aligned}$$

Note that here we have not substituted for p_b in terms of p_a. If we do not make this substitution we need to use a Lagrangian approach to determine the optimal values of p_a p_b. Some useful formulae emerge if we do this. Since $p_a + p_b = \frac{1}{2^{(k-2)}}$ the Lagrangian is

$$L(p_a, p_b, \lambda) = F - \lambda(p_a + p_b - \frac{1}{2^{(k-2)}}).$$

Having formed our total objective function we now determine the partial derivatives of $L(p_a, p_b, \lambda)$ with respect to p_a p_b and λ respectively.

$$\frac{\partial L(p_a,p_b,\lambda)}{\partial p_a} = \frac{\partial F}{\partial p_a} - \lambda \tag{5.9}$$

$$\frac{\partial L(p_a,p_b,\lambda)}{\partial p_b} = \frac{\partial F}{\partial p_b} - \lambda \tag{5.10}$$

$$\frac{\partial L(p_a,p_b,\lambda)}{\partial \lambda} = -(p_a + p_b - \frac{1}{2^{(k-2)}}). \tag{5.11}$$

Hence

$$\left\{ \begin{array}{l} \frac{\partial L(p_a,p_b,\lambda)}{\partial p_a} = 0 \\ \frac{\partial L(p_a,p_b,\lambda)}{\partial p_b} = 0 \end{array} \right. \iff \left\{ \begin{array}{l} \frac{\partial F}{\partial p_a} = \lambda \\ \frac{\partial F}{\partial p_b} = \lambda \end{array} \right. \tag{5.12}$$

To determine λ we note

$$p_a\frac{\partial F}{\partial p_a} + p_b\frac{\partial F}{\partial p_b} = \lambda(p_a + p_b) \tag{5.13}$$

$$= \frac{1}{2^{(k-2)}}\lambda \tag{5.14}$$

Consequently,

$$\lambda = 2^{(k-2)}\left[p_a\frac{\partial F}{\partial p_a} + p_b\frac{\partial F}{\partial p_b}\right] \tag{5.15}$$

Now

$$\frac{\partial F}{\partial p_a} = \frac{1}{p_a} + \frac{(k-2)w(a)}{[p_a w(a) + p_b w(b)]} \tag{5.16}$$
$$\frac{\partial F}{\partial p_b} = \frac{1}{p_b} + \frac{(k-2)w(b)}{[p_a w(a) + p_b w(b)]} \tag{5.17}$$

Multiplying equations (5.16) and (5.17) by p_a p_b respectively, and summing the resulting equations we can write

$$p_a \frac{\partial F}{\partial p_a} + p_b \frac{\partial F}{\partial p_b} = 1 + \frac{(k-2)p_a w(a)}{[p_a w(a) + p_b w(b)]} + 1 + \frac{(k-2)p_b w(b)}{[p_a w(a) + p_b w(b)]} = k$$

which is constant. And from equation (5.15), $\lambda = 2^{(k-2)}k$.

Further

$$\begin{aligned} \frac{\partial F}{\partial p_a} &= \frac{1}{p_a} + \frac{(k-2)w(a)}{p_a w(a) + p_b w(b)} = 2^{(k-2)}k \\ &\Rightarrow 1 + \frac{(k-2)p_a w(a)}{p_a w(a) + p_b w(b)} = 2^{(k-2)}k p_a \\ &\Rightarrow \frac{(k-2)p_a w(a)}{p_a w(a) + p_b w(b)} = 2^{(k-2)}k p_a - 1 \end{aligned} \tag{5.18}$$

Similarly,

$$\begin{aligned} \frac{\partial F}{\partial p_b} &= \frac{1}{p_b} + \frac{(k-2)w(b)}{p_b w(b) + p_b w(b)} = 2^{k-2}k \\ &\Rightarrow 1 + \frac{(k-2)w(b)p_b}{p_b w(b) + p_b w(b)} = 2^{(k-2)}k p_b \\ &\Rightarrow \frac{(k-2)w(b)p_b}{p_b w(b) + p_b w(b)} = 2^{(k-2)}k p_b - 1 \end{aligned} \tag{5.19}$$

We note that we will be interested in derivatives of this function with respect to a and/or b. To find the best four-point design on Z_w, we need to maximise $\ln[detM(p)]$ w.r.t. a and b or if we wish to find the best four point design subject to a (or b) being a support point we need to maximise F w.r.t. b (or a).

$$\frac{\partial F}{\partial a} = \frac{\partial F}{\partial A(a,b)}\frac{\partial A(a,b)}{\partial a} + \frac{\partial F}{\partial w(a)}\frac{\partial w(a)}{\partial a} + \frac{\partial F}{\partial p_a}\frac{\partial p_a}{\partial a} + \frac{\partial F}{\partial p_b}\frac{\partial p_b}{\partial a} \tag{5.20}$$

Now we can substitute the values from equations (5.18) and (5.19) into equation (5.20) to obtain the following :

$$\frac{\partial F}{\partial a} = \frac{\partial F}{\partial A(a,b)}\frac{\partial A(a,b)}{\partial a} + \frac{\partial F}{\partial w(a)}\frac{\partial w(a)}{\partial a} + 2^{(k-2)}k\frac{\partial p_a}{\partial a} + 2^{k-2}k\frac{\partial p_b}{\partial a} \tag{5.21}$$

From the definition of p_a, p_b ($p_a + p_b = \left(\frac{1}{2}\right)^{(k-2)}$), we can write the following

$$\frac{\partial p_a}{\partial a} + \frac{\partial p_b}{\partial a} = 0$$
$$\frac{\partial p_a}{\partial a} = -\frac{\partial p_b}{\partial a}$$

Therefore,

$$\begin{aligned}
\frac{\partial F}{\partial a} &= \frac{\partial F}{\partial A(a,b)}\frac{\partial A(a,b)}{\partial a} + \frac{\partial F}{\partial w(a)}\frac{\partial w(a)}{\partial a} + \underbrace{\frac{\partial p_a}{\partial a}[2^{(k-2)}k - 2^{(k-2)}k]}_{=0} \\
&= \frac{\partial A(a,b)}{\partial a} + \frac{\partial F}{\partial w(a)}\frac{\partial w(a)}{\partial a} \\
&= \frac{-2}{b-a} + \frac{w'(a)}{w(a)} + \frac{(k-2)p_a w'(a)}{p_a w(a) + p_b w(b)} \\
&= \frac{-2}{b-a} + \frac{w'(a)}{w(a)}\left[1 + \frac{(k-2)p_a w(a)}{p_a w(a) + p_b w(b)}\right] \\
&= \frac{-2}{b-a} + \frac{2^{(k-2)}k p_a w'(a)}{w(a)} \\
&= \frac{-2w(a) + (p_a w'(a)\, 2^{(k-2)}k)\,(b-a)}{w(a)(b-a)} \\
&= \frac{w'(a)\, 2^{(k-2)}k\, p_a}{w(a)(b-a)}\left[(b-a) - \frac{2w(a)}{2^{(k-2)}k p_a w'(a)}\right] \quad if \quad w'(a) \neq 0 \\
&= \frac{w'(a)\, 2^{(k-2)}k\, p_a}{w(a)(b-a)}\left[(b-a) - \frac{w(a)}{2^{k-3}k p_a w'(a)}\right] \quad if \quad w'(a) \neq 0 \\
&= \frac{p_a w'(a)\, 2^{(k-2)}k}{w(a)(b-a)}\left[b - h_b(a)\right]
\end{aligned}$$

where $\quad h_b(a) = a + \frac{w(a)}{2^{k-3}k\, p_a w'(a)}$. So $\frac{\partial[\ln det M(\xi)]}{\partial z_1} \propto p_a w'(z)\, 2^{(k-2)}k\big[b - h_b(a)\big]$

Further,

$$\frac{\partial[\ln det M(p)]}{\partial a} = 0 \quad if \quad \Rightarrow \quad \frac{p_a w'(a)\, 2^{(k-2)}k}{w(a)(b-a)}\left[(b-a) - \frac{w(a)}{2^{(k-3)}k\, p_a w'(a)}\right] = 0$$
$$i.e. \quad if \quad b = h_b(a) \quad \big[given \quad w'(a) \neq 0\big].$$

(Note, if $p_a w'(a) = 0$, $\frac{\partial[\ln det M(p)]}{\partial a} = \frac{-2}{b-a} \neq 0$. So, z_{max} is not a solution of $\frac{\partial[\ln det M(p)]}{\partial a} = 0$, where z_{max} is the value of z_1 that maximises $\boldsymbol{w(z_1)}$.

Similarly,

$$\frac{\partial F}{\partial b} = \frac{\partial F}{\partial A(a,b)}\frac{\partial A(a,b)}{\partial b} + \frac{\partial F}{\partial w(b)}\frac{\partial w(b)}{\partial b} + \frac{\partial F}{\partial p_a}\frac{\partial p_a}{\partial b} + \frac{\partial F}{\partial p_b}\frac{\partial p_b}{\partial b} \quad (5.22)$$

Now we can substitute the values from equations (5.18) and (5.19) into equation (5.20) to obtain the following :

$$\frac{\partial F}{\partial b} = \frac{\partial A(a,b)}{\partial b} + \frac{\partial F}{\partial w(b)}\frac{\partial w(b)}{\partial b} + 2^{(k-2)}k\frac{\partial p_a}{\partial b} + 2^{(k-2)}k\frac{\partial p_b}{\partial b} \quad (5.23)$$

From the definition of p_a, p_b ($p_a + p_b = \frac{1}{2^{(k-2)}}$), we can write the following

$$\frac{\partial p_a}{\partial b} + \frac{\partial p_b}{\partial b} = 0$$
$$\frac{\partial p_a}{\partial b} = -\frac{\partial p_b}{\partial b}$$

Therefore,

$$\begin{aligned}
\frac{\partial F}{\partial b} &= \frac{\partial A(a,b)}{\partial b} + \frac{\partial F}{\partial w(b)}\frac{\partial w(b)}{\partial b} + \underbrace{\frac{\partial p_b}{\partial b}[2^{(k-2)}k - 2^{(k-2)}k]}_{=0} \\
&= \frac{\partial A(a,b)}{\partial b} + \frac{\partial F}{\partial w(b)}\frac{\partial w(b)}{\partial b} \\
&= \frac{2}{b-a} + \frac{w'(b)}{w(b)} + \frac{(k-2)p_b w'(b)}{p_a w(a) + p_b w(b)} \\
&= \frac{2}{b-a} + \frac{w'(b)}{w(b)}\left[1 + \frac{(k-2)p_a w(a)}{p_a w(a) + p_b w(b)}\right] \\
&= \frac{2}{b-a} + \frac{2^{(k-2)}kp_b w'(b)}{w(b)} \\
&= \frac{2w(b) + (p_b w'(b)\,2^{(k-2)}k)\,(b-a)}{w(b)(b-a)} \\
&= \frac{w'(b)(p_b w'(b)\,2^{(k-2)}k)}{w(b)(b-a)}\left[(b-a) + \frac{2w(b)}{2^{(k-2)}kp_b w'(b)}\right] \quad if \quad w'(b) \neq 0 \\
&= \frac{w'(b)(p_b w'(b)\,2^{(k-2)}k)}{w(b)(b-a)}\left[(b-a) + \frac{w(b)}{2^{k-3}kp_b w'(b)}\right] \quad if \quad w'(b) \neq 0 \\
&= \frac{p_b w'(b)\,2^{(k-2)}k}{w(b)(b-a)}\left[h_a(b) - a\right]
\end{aligned}$$

where $h_a(b) = b + \frac{w(b)}{2^{k-3}k\,p_b w'(b)}$. So $\frac{\partial[\ln det M(\xi)]}{\partial z_1} \propto p_b w'(z)\,2^{(k-2)}k\left[h_a(b) - a\right]$

Further,

$$\frac{\partial[\ln detM(p)]}{\partial b} = 0 \quad if \quad \Rightarrow \quad \frac{p_b w'(b)\, 2^{(k-2)} k}{w(b)(b-a)} \left[(b-a) + \frac{w(b)}{2^{(k-3)} k\, p_b w'(b)}\right] = 0$$

$$i.e. \quad if \quad b = h_a(b) \quad \left[given \quad w'(b) \neq 0\right].$$

(Note, if $p_b w'(b) = 0$, $\frac{\partial[\ln detM(p)]}{\partial b} = \frac{2}{b-a} \neq 0$. So, z_{max} is not a solution of $\frac{\partial[\ln detM(p)]}{\partial b} = 0$.)

As a result of this, we can be interested in solving one or both of the equations

$$a = h_a(b) \tag{5.24}$$

$$h_b(a) = b \tag{5.25}$$

Clearly, the function $h_a(b)$, $h_b(a)$ play the role of $h(z)$ in the two parameter case but that has now been replaced by a class of functions as in the three parameter case. It is useful to study $h_t(z_1)$. The solutions to these equations clearly depend on the nature of $h_t(z_1)$. We consider the same weight functions as in chapter two. Plots in the case of the weight function for binary logistic regression are shown in Figures (5.2), (5.3), (5.4) and (5.5). This again is useful to us.

Now consider the single equation in z_1

$$h_{z_2}(z_1) = e.$$

As in the previous chapter there is one solution to this equation say $z_1 = z_L^*(e)$ in the range $z_1 \leq z_{max}$ and one, say $z_1 = z_U^*(e)$, in the range $z_1 \geq z_{max}$. Moreover since $w'(z_L^*(e)) > 0$ and $w'(z_U^*(e)) < 0$ we have $z_L^*(e) < e < z_U^*(e)$. In equations 5.24 and 5.25 we have two versions of the above. Their joint solution with $z_1 < z_2$, must be $z_1^* = a^*$, $z_2^* = b^*$, $a^* < b^*$, a^*, b^* being the support points of the optimal 2^{k-1} design on Z_w as defined in the conjectures above. Note that this means

$$h(a^*) = b^*, \qquad h(b^*) = a^* \qquad and \quad z_1^* = z_L^*(z_2^*), \quad z_2^* = z_U^*(z_1^*).$$

5.7 Examination of the conjecture against the equivalence theorem

We now begin to check the conjectures in section (5.3) against the equivalence theorem. We consider an arbitrary k parameter design as follows:

$$\begin{pmatrix} i & 1 & 2 & 3 & \cdots & M & N & N+1 & N+2 & \cdots & L \\ z_{1i} & a & a & a & \cdots & a & b & b & b & \cdots & b \\ z_{2i} & -1 & 1 & -1 & \cdots & 1 & -1 & 1 & -1 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ z_{li} & -1 & 1 & -1 & \cdots & 1 & -1 & 1 & -1 & \cdots & 1 \\ p_i & p_a & p_a & p_a & \cdots & p_a & p_b & p_b & p_b & \cdots & p_b \end{pmatrix}$$

where $p_a, p_b > 0$ are the optimal weights for a and b and $M = 2^{k-2}$, $N = 2^{k-2}+1$ and $L = 2^{k-1}$. The design matrix is

$$M(p) = \sum_{i}^{2^{k-1}} p_i \underline{g}_i \underline{g}_i^T$$

where

$$\underline{g}_i = \sqrt{w(z_{1i})}(1, z_{1i}, z_{2i}, \cdots, z_{li})^T \quad i = 1, 2, \cdots, 2^{k-1}.$$

$$M(p) = 2^{(k-2)} \begin{pmatrix} p_a w(a) + p_b w(b) & a p_a w(a) + b p_b w(b) & 0 & \cdots & 0 \\ a p_a w(a) + b p_b w(b) & b_1^2 p_a w(a) + b^2 p_b w(b) & 0 & \cdots & 0 \\ 0 & 0 & p_a w(a) + p_b w(b) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & p_a w(a) + p_b w(b) \end{pmatrix} \quad (5.26)$$

The design matrix can partioned as follows:

$$M(p) = 2^{(k-2)} \begin{pmatrix} \mathcal{S}_0 & 0 \\ 0 & \mathcal{S}_2 \end{pmatrix}.$$

Here $\mathcal{S}_0$ is the 2×2 matrix

$$\mathcal{S}_0 = \begin{pmatrix} S_{011} & S_{012} \\ S_{021} & S_{022} \end{pmatrix} = \begin{pmatrix} p_a w(a) + p_b w(b) & a p_a w(a) + b p_b w(b) \\ a p_a w(a) + b p_b w(b) & b_1^2 p_a w(a) + b^2 p_b w(b) \end{pmatrix}.$$

From the definition of $\mathcal{S}_0$,

$$\mathcal{S}_0^{-1} = \begin{pmatrix} \frac{S_{022}}{|\mathcal{S}_0|} & -\frac{S_{012}}{|\mathcal{S}_0|} \\ \frac{-S_{021}}{|\mathcal{S}_0|} & \frac{S_{011}}{|\mathcal{S}_0|} \end{pmatrix}$$

where the determinant of $\mathcal{S}_0$ is :

$$\begin{aligned} |\mathcal{S}_0| &= [p_a w(a) + p_b w(b)][p_a w(a)a^2 + p_b w(b)b^2] - [p_a w(a)a + p_b w(b)]^2 \\ &= p_a p_b w(a) w(b)(b-a)^2. \end{aligned}$$

Further, $\mathcal{S}_2 = cI$ where I is the $(k-2) \times (k-2)$ identity matrix and $c = p_a w(a) + p_b w(b)$.

Therefore

$$M^{-1}(p) = \frac{1}{2^{(k-2)}} \begin{pmatrix} \frac{S_{022}}{|\mathcal{S}_0|} & -\frac{S_{012}}{|\mathcal{S}_0|} & 0 & 0 & \cdots & 0 \\ -\frac{S_{021}}{|\mathcal{S}_0|} & \frac{S_{011}}{|\mathcal{S}_0|} & 0 & 0 & \cdots & 0 \\ 0 & 0 & \frac{1}{c} & 0 & \cdots & 0 \\ 0 & 0 & 0 & \frac{1}{c} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \frac{1}{c} \end{pmatrix}.$$

If the above design is to be D-optimal on a set of values of z_1 for z_j $j = 2, \cdots, l$, say the set Z, then, as noted in section 5.4, we must have

$$v^{\times}(z_1) \leq 0 \qquad \forall \in Z \tag{5.27}$$

where

$$v^{\times}(z_1) = Q^{\times}(z_1) - \frac{k}{w(z_1)} \tag{5.28}$$

where

$$Q^{\times}(z_1) = \frac{1}{2^{(k-2)}}(1,\, z_1,\, \pm1, \cdots, \pm1) \begin{pmatrix} \frac{S_{022}}{|S_0|} & -\frac{S_{012}}{|S_0|} & 0 & 0 & \cdots & 0 \\ -\frac{S_{021}}{|S_0|} & \frac{S_{011}}{|S_0|} & 0 & 0 & \cdots & 0 \\ 0 & 0 & \frac{1}{c} & 0 & \cdots & 0 \\ 0 & 0 & 0 & \frac{1}{c} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \frac{1}{c} \end{pmatrix} \begin{pmatrix} 1 \\ z_1 \\ \pm1 \\ \vdots \\ \pm1 \end{pmatrix}$$

$$= \frac{1}{2^{(k-2)}}\left[\frac{\mathcal{S}_{022} - z_1\mathcal{S}_{012}}{\mid \mathcal{S}_0 \mid},\ \frac{-\mathcal{S}_{012} + z_1\mathcal{S}_{011}}{\mid \mathcal{S}_0 \mid},\ \pm\frac{1}{c}, \cdots, \pm\frac{1}{c}\right] \begin{pmatrix} 1 \\ z_1 \\ \pm1 \\ \vdots \\ \pm1 \end{pmatrix}$$

$$= \frac{1}{2^{(k-2)}}\left[\frac{\mathcal{S}_{022} - 2z_1\mathcal{S}_{012} + z_1^2\mathcal{S}_{011}}{\mid \mathcal{S}_0 \mid} + \frac{1}{c} + \cdots + \frac{1}{c}\right]$$

$$= \frac{1}{2^{(k-2)}}\left[\frac{\mathcal{S}_{022} - 2z_1\mathcal{S}_{012} + z_1^2\mathcal{S}_{011}}{\mid \mathcal{S}_0 \mid} + \frac{(k-2)}{c}\right]. \tag{5.29}$$

Equivalently we must have

$$v(z_1) \le 0 \tag{5.30}$$

where $v(z_1) = Q(z_1) - \dfrac{2^{(k-2)}k}{w(z_1)}$ with $Q(z_1) = 2^{(k-2)}Q^{\times}(z_1)$. And $v(z_1)$ must be maximised at a, b over Z, a maximum of zero as $v(a) = v(b) = 0$. So, we need to consider the derivative of $v(z_1)$ at a, b. We have to explore the derivative of $v(z_1)$:

$$v'(z_1) = Q'(z_1) + \frac{k2^{(k-2)}w'(z_1)}{[w(z_1)]^2} \tag{5.31}$$

$$= L_k(z_1) - H_k(z_1) \tag{5.32}$$

where or $L_k(z_1) = Q'(z_1)$ and $H_k(z_1) = -\frac{k2^{(k-2)}w'(z_1)}{[w(z_1)]^2}$.

Now

$$\begin{aligned} L_k(z_1) &= Q'(z_1) \\ &= \frac{-2\mathcal{S}_{012} + 2z_1\mathcal{S}_{011}}{\mid \mathcal{S}_0 \mid} \end{aligned}$$

In particular,

$$\begin{aligned} L_k(a) &= \frac{2[p_a w(a) + p_b w(b)]a - 2[p_a w(a)a + p_b w(b)b]}{\mid \mathcal{S}_0 \mid} \\ &= \frac{2p_b w(b)(a-b)}{p_a p_b w(a) w(b)(a-b)^2} \\ &= \frac{-2}{(b-a)p_a w(a)} \end{aligned}$$

$$\begin{aligned} L_k(b) &= \frac{2[p_a w(a) + p_b w(b)]b - 2[p_a w(a)a + p_b w(b)b]}{\mid \mathcal{S}_0 \mid} \\ &= \frac{2p_a w(a)(b-a)}{p_a p_b w(a) w(b)(a-b)^2} \\ &= \frac{2}{(b-a)p_b w(b)} \end{aligned}$$

So

$$\begin{aligned} v'(a) &= L_k(a) - H_k(a) \\ &= L_k(a) + \frac{k2^{(k-2)}w'(a)}{[w(a)]^2} \\ &= \frac{-2}{p_a w(a)(b-a)} + \frac{k2^{(k-2)}w'(a)}{[w(a)]^2} \\ &= \frac{1}{w(a)}\Big[\frac{-2}{p_a(b-a)} + \frac{k2^{(k-2)}w'(a)}{w(a)}\Big] \\ &= \frac{k2^{(k-2)}w'(a)}{[w(a)]^2(b-a)}\Big[\frac{-2w(a)}{k2^{(k-2)}p_a w'(a)} + (b-a)\Big] \\ &= \frac{k2^{(k-2)}w'(a)}{[w(a)]^2(b-a)}\Big[b - [a + \frac{w(a)}{k2^{(k-3)}p_a w'(a)}]\Big] \\ &= \frac{k2^{(k-2)}w'(a)}{[w(a)]^2(b-a)}[b - h_b(a)] \end{aligned}$$

$$
\begin{aligned}
v'(b) &= L_3(b) - H_3(b) \\
&= L_3(b) + \frac{k2^{(k-2)}w'(b)}{[w(b)]^2} \\
&= \frac{2}{p_b w(b)(b-a)} + \frac{k2^{(k-2)}w'(b)}{[w(b)]^2} \\
&= \frac{1}{w(b)}\Big[\frac{2}{p_b(b-a)} + \frac{k2^{(k-2)}w'(b)}{w(b)}\Big] \\
&= \frac{k2^{(k-2)}w'(b)}{[w(b)]^2(b-a)}\Big[\frac{2w(b)}{k2^{(k-2)}p_b w'(b)} + (b-a)\Big] \\
&= \frac{k2^{(k-2)}w'(b)}{[w(b)]^2(b-a)}\Big[[b + \frac{w(b)}{k2^{(k-3)}p_b w'(b)}] - a\Big] \\
&= \frac{k2^{(k-2)}w'(b)}{[w(b)]^2(b-a)}[h_a(b) - a]
\end{aligned}
$$

Therefore

$$
\begin{aligned}
v'(a) &\propto w'(a)[b - h_b(a)] \\
v'(b) &\propto w'(b)[h_a(b) - a]
\end{aligned}
$$

So the sign of $v'(a)$ and $v'(b)$ depend on the signs of $w'(a)$ $[b - h_b(a)]$ and $w'(b)$ $[h_a(b) - a]$ respectively.

5.8 Proof of the Conjecture

The function $h_y(z)$ has exactly the same definition as in the **three** parameter case. Hence if it were increasing in z over $z < z_{max}$ and $z > z_{max}$ then the proof of the conjecture would be identical.

Condition (i) of section (3.2.6) does appear to hold for low dimensions $k \leq 6$, but for higher dimensions $h_y(z)$ can have two TP's : a maximal then a minimal one. Condition (ii) is satisfied for all dimensions. These assertions are evident from the plots of $h_y(z)$ in Figures (5.2), (5.3), (5.4), (5.5) (for different values of k, the number of parameters), (5.6), (5.7) (for different values of y (the end point-support point)) for the Logistic weight function.

What is also evident is that a weaker but still sufficient condition for proof of the conjecture is satisfied : namely that for $b > z_{max}$, $h_b(z) > b$ for $a^*(b) < z < z_{max}$, while for $a < z_{max}$, $h_a(z) < a$ for $z_{max} < z < b^*(a)$. There is in fact only one solution in each of the ranges $z < z_{max}$ and $z > z_{max}$ to the equation

$$h_y(z) \quad = \quad y.$$

5.9 Study of the function $h_y(z)$

In this section we will be looking at the function

$$h_y(z) \quad = \quad z + \frac{1}{k2^{(k-3)}} \frac{w(z)}{p_y(z)w'(z)}.$$

We note that there will be pairs of values (y_1, y_2) of y with $y_1 < z_{max} < y_2$ such that

$$h_{y_1}(z) \quad = \quad h_{y_2}(z) \qquad \forall \quad z.$$

This follows since $w(z)$ is unimodal. These values must satisfy $w(y_1) = w(y_2)$. If $w(z)$ is symmetric about zero then $y_1 = -y_2$. Otherwise numerical techniques will usually be needed to determine y_2 say, for given y_1. An exception to the above is of course $y = z_{max}$.

So a given function $h_y(z)$ could be labelled with two different y-values. For our purposes the most important labels are the higher y-values over $z \leq z_{max}$ and the lower y-values over $z \geq z_{max}$. So first of all this study leads us to the consideration of graphing the function $h_y(z)$. Recall that $h_y(z)$ is a function of z and of y and k.

We consider the dependence of $h_y(z)$ on z, y and k in turn.

We focus on $z < z_{max}$ throughout.

- **Dependence on z:**

  We have to remark that $h_y(z)$ is not monotonically increasing. [see Figure 5.6 and Figure 5.7]. Note that on these figures $\pm y$ values are also plotted. On both sides of the plot of $h_y(z)$ has local minimum and maximum TP's. Ideally we want to see $h_y(z)$ crossing $\pm y$ once. (Clearly, this is true if $h_y(z)$ is increasing in z.) The presence of TP's implies the possibility of three crossings. However the plots suggest that this does not happen. Fortunately the TP's lie between $\pm y$. Consequently, $h_y(z)$ crosses the value $\pm y$ once in $(-\infty, z_{max}]$ and once in $[z_{max}, \infty)$. As we can see from in Figure (5.1) it appears that:

  1. $z^*(-y) < z_1 < z_2 < z^*(+y)$ and $\underbrace{h'_y(z_1)}_{max\ TP} = \underbrace{h'_y(z_2)}_{min\ TP} = 0.$
  2. $-y \leq h_y(z_i) \leq y, \quad i = 1, 2$ where $h'(z_i) = 0 \quad i = 1, 2,$ $z^*(c)$ solves $h_y(z) = c.$

     Crucially $h_y(z_1) < y$ and $h_y(z_2) > -y$.

  Further the graphs of $h_y(z)$ for increasing k show that when the number of parameters increase the minimum TP approaches the point $(-y, -y)$ but never touches it, and this in turn gurantees that we will always end up having two support points.

  For further insights consider $d = d(z) = (k-1)(r-1)$. Substituting d in

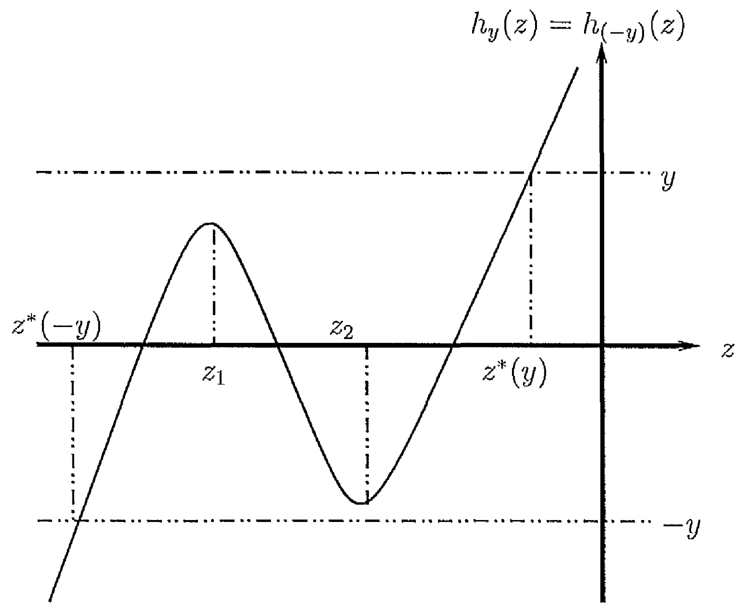


Figure 5.1: Plot of arbitrary $h_y(z)$ function.

$q(r)$ gives

$$q(r) = q^{\times}(d) = \frac{(k-1)\left\{d-2+\sqrt{d^2+\frac{4d}{k-1}+4}\right\}}{2k2^{(k-2)}d}.$$

Hence

$$h_y(z) = z + \frac{g(d)}{d'(z)} \tag{5.33}$$

where here $g(d) = \frac{1}{k}\left\{2-d+\sqrt{d^2+\frac{4d}{k-1}+4}\right\}\{d+(k-1)\}$.

As means of forming an impression about $h_y(z)$, we study $g(d)$. We note the following points :

1. $d \geq -(k-1)$ since $r \geq 0$. So $d \in [-(k-1), \infty)$.
2. $g(d) = 0$ at $d = -(k-1)$

   e.g. if $k = 6 \quad \Rightarrow \quad d = -5 \quad \Rightarrow g(d) = 0$
3. $g(d)$ is a positive function since $g(d) = \frac{w(z)}{p_y(z)}$.
4. $g(d) \longrightarrow \infty$ as $d \to +\infty$

5. In general, $g(d)$ has 2 TP's, first a local *maximum* and then a local *minimum*. However it seems that there are no TP's if $k \leq 5$. [see Figures (5.8), (5.9), (5.10)].

6. **The derivative of $g(d)$ with respect to d is**

$$\frac{\partial g(d)}{\partial d} = \frac{1}{k}\{2-2d-(k-1)+A+A'(d+(k-1))\}$$

with $A = \left(d^2 + \frac{4d}{(k-1)} + 4\right)^{\frac{1}{2}}$ and $A' = \frac{1}{2A}\left(2d + \frac{4}{k-1}\right)$.

So

$$\begin{aligned}
\frac{\partial g(d)}{\partial d} &= \frac{1}{k}\left\{2-2d-(k-1)+A+\frac{1}{2A}\left(2d+\frac{4}{k-1}\right)[d+(k-1)]\right\} \\
&= \frac{1}{k}\{2-2d-(k-1)\} \\
&+ \frac{1}{k}\left\{A + \frac{1}{A}(k-1)\left[\frac{d^2+d(k-1)+2}{k-1}\right] + \frac{2d}{A(k-1)}\right\} \\
&= \frac{1}{A(k-1)k}\{(k-1)[2-2d-(k-1)]A\} \\
&\quad + \frac{1}{A(k-1)k}\left\{A^2(k-1)+(k-1)\left[d^2+d(k-1)+2\right]+2d\right\} \\
&= \frac{1}{Ak}\left\{[(3-k)-2d]A + 2d^2 + \left[\frac{6+(k-1)^2}{k-1}\right]d + 6\right\}.
\end{aligned}$$

Now, solving $\frac{\partial g(d)}{\partial d} = 0$ will lead to

$$\underbrace{[(3-k)-2d]\sqrt{d^2+\frac{4d}{k-1}+4}}_{LHS} = \underbrace{\left\{-\left(2d^2+\left[\frac{6+(k-1)^2}{k-1}\right]d+6\right)\right\}}_{RHS}.$$

Now,

$$\begin{aligned}
RHS^2 &= \left\{-\left(2d^2+\left[\frac{6+(k-1)^2}{k-1}\right]d+6\right)\right\}^2 \\
&= 4[d]^4 + 4\left\{\frac{(k-1)^2+6}{k-1}\right\}[d]^3 \\
&\quad + \left\{24+\left[\frac{(k-1)^2+6}{k-1}\right]^2\right\}[d]^2 + 12\left\{\frac{(k-1)^2+6}{k-1}\right\}d + 36
\end{aligned}$$

and

$$\begin{aligned}
LHS^2 &= \{[(3-k)-2d]\,A\}^2 \qquad \text{where} \quad A = \sqrt{d^2 + \frac{4d}{k-1} + 4} \\
&= \left\{(3-k)^2 + 4d^2 + 4(3-k)d\right\}\left\{d^2 + \frac{4d}{k-1} + 4\right\} \\
&= 4d^4 + \frac{4}{(k-1)}\left\{k^2 - 4k + 7\right\}d^3 + \left\{(3-k)^2 + \frac{32}{k-1}(k-2)\right\}d^2 \\
&+ 4(k-3)\left\{\frac{(k-3)+4(k-1)}{(k-1)}\right\}d + 4(k-3)^2.
\end{aligned}$$

Hence,

$$\begin{aligned}
C(d) = RHS^2 - LHS^2 &= \frac{8k}{k-1}d^3 \\
&+ \left[24 + \frac{\left[(k-1)^2+6\right]^2 - (3-k)^2(k-1)^2}{(k-1)^2} - \frac{32(k-2)}{k-1}\right]d^2 \\
&- \frac{8k(k-8)}{k-1}d + 4k(6-k) \\
&= \frac{8k}{k-1}d^3 + \frac{4k}{(k-1)^2}\left\{k - 3k + 11\right\}d^2 \\
&- \frac{8k(k-8)}{k-1}d + 4k(6-k).
\end{aligned}$$

Stationary values of $g(d)$ will be a subset of the roots of the cubic equation

$$C(d) = 0.$$

Thus $C(d)$ can have at most three TP's and probably less. Consider $k = 6$. At this value the constant term of the cubic is zero and

$$\begin{aligned}
C(d) = 0 &\iff \frac{48}{5}d^3 + \frac{696}{25}d^2 + \frac{96}{5}d = 0 \\
&\iff \frac{12}{5^2}d\left[20d^2 + 58d + 40\right] = 0 \\
&\implies d_1(z) = 0 \quad \text{or} \quad d_2(z) = -1.13 \quad \text{or} \quad d_3(z) = -1.77.
\end{aligned}$$

We note that all the three roots are greater than $[-(k-1)] = -5$, the lower limit on d. However $d_3(z)$ is not a solution to $\frac{\partial g(d)}{\partial d} = 0$ while $d_1(z)$ and $d_2(z)$ do identify TP's of $g(d)$ (*minimal* and *maximal* TP, respectively). Thus $g(d)$ has two turning points. This appears to be the case for $k \geq 6$ in general while for $k \leq 5$, $g(d)$ is increasing.

7. In fact $\dfrac{\partial g(-(k-1))}{\partial d} > 0$, so there are only **two** TP's in $[-(k-1), \infty)$. [see Figure (5.11) (5.12) (5.13)].

8. It appears that $g(d)$ is an increasing function for $k \leq 5$ while for $k \geq 6$ it has two TPs as in the case $k = 6$. See Figures (5.8) (5.9) (5.9).

The above properties of $g(d)$ would appear to induce the same in $h_y(z)$ i.e. $h_y(z)$ increasing in z for $k \leq 5$, while developing two TP's for $k \geq 6$.

- **Dependence on y :**

For all $z_1 \leq z_{max}$, $h_y(z_1)$ decreases in y over $y \geq z_{max}$, and for all $z_1 \geq z_{max}$ $h_y(z_1)$ decreases in y over $y \leq z_{max}$. To prove these we write $h_y(z)$ in the following form:

$$h_y(z) = z + \frac{w(z)}{q(r)w'(z)} \qquad r = \frac{w(z)}{w(y)}$$

where $q(r)$ is the expression encountered in equation (5.8) and $r = \dfrac{w(z)}{w(y)}$. From the above expession of $h_y(z)$, proving that $q(r)$ is increasing in y would be sufficient to establish that $h_y(z)$ is decreasing in y. In Appendix B we prove that $q(r)$ is increasing in r.

- Hence $h_y(z)$ decreasing in r.
- Hence $r = \dfrac{w(z)}{w(y)}$ increases in y over $[z_{max}, \infty)$
- We note that $w(y)$ is decreasing in y over $[z_{max}, \infty)$,

Therefore, $h_y(z)$ decreases in y.

- **Dependence on k :**

  – Now we show analytically that the function $h_y(z)$ is decreasing in k. Substitute $d = (k-1)(r-1)$ in the function $g(d)$.

Therefore we will have,

$$h_y(z) \quad = \quad z + \frac{g(r,k)}{(k-1)r'(z)} \tag{5.34}$$

where

$$g(r,k) = \frac{r(k-1)}{k}\left\{2 - (k-1)(r-1) + \sqrt{[(k-1)(r-1)]^2 + 4r}\right\}.$$

Let

$$\begin{aligned} G(r,k) &= \frac{g(r,k)}{(k-1)} \\ &= \frac{r}{k}\left\{2 - (k-1)(r-1) + \sqrt{A}\right\} \end{aligned}$$

where $A = (k-1)^2(r-1)^2 + 4r$.

Therefore we will have,

$$h_y(z) \quad = \quad z + \frac{G(r,k)}{r'(z)}.$$

To show that $h_y(z)$ is decreasing in k we need to prove that $\frac{\partial h_y(z)}{\partial k} < 0$.

Now, let's take the derivative of $h_y(z)$ respect to k:

$$\frac{\partial h_y(z)}{\partial k} \quad = \quad \frac{\frac{\partial G(r,k)}{\partial k}}{r'(z)} - \frac{\frac{\partial r'(z)}{\partial k}G(r,k)}{[r'(z)]^2}.$$

Now $\frac{\partial r'(z)}{\partial k} = 0$, since r is independent of k. Therefore $\frac{\partial h_y(z)}{\partial k} = \frac{\frac{\partial G(r,k)}{\partial k}}{r'(z)}$. So, taking the derivative of $G(r,k)$ respect to k will be enough for us to see the behaviour of $h_y(z)$ with respect to k.

The derivative of $G(r,k)$ with respect to k is given by the following expression:

$$\begin{aligned}
\frac{\partial G(r,k)}{\partial k} &= \frac{r}{k}\left\{-(r-1)+A^{-\frac{1}{2}}(k-1)(r-1)^2\right\} \\
&+ \frac{(-r)}{k^2}\left\{2-(k-1)(r-1)+A^{\frac{1}{2}}\right\} \\
&= \frac{r(r-1)}{k}\left\{(k-1)(r-1)A^{-\frac{1}{2}}-1\right\} \\
&- \frac{r}{k^2}\left\{2-(k-1)(r-1)+A^{\frac{1}{2}}\right\} \\
&= \frac{r(k-1)(r-1)^2}{k\sqrt{A}}-\frac{r(r-1)}{k}-\frac{2r}{k^2} \\
&+ \frac{r(r-1)(k-1)}{k^2}-\frac{r\sqrt{A}}{k^2} \\
&= \frac{rk(k-1)(r-1)^2}{k^2\sqrt{A}}-\frac{rk(r-1)\sqrt{A}}{k^2\sqrt{A}} \\
&- \frac{2r\sqrt{A}}{k^2\sqrt{A}}+\frac{r(k-1)(r-1)\sqrt{A}}{k^2\sqrt{A}}-\frac{r\sqrt{A}\sqrt{A}}{k^2\sqrt{A}} \\
&= \frac{r}{k^2\sqrt{A}} \\
&\quad \left\{k(k-1)(r-1)^2-k(r-1)\sqrt{A}-2\sqrt{A}\right. \\
&\quad \left.+(k-1)(r-1)\sqrt{A}-A\right\} \\
&= \frac{r}{k^2\sqrt{A}}\left\{k(k-1)(r-1)^2-A+\sqrt{A}\right. \\
&\quad \left.[(k-1)(r-1)-k(r-1)-2]\right\} \\
&= \frac{r}{k^2\sqrt{A}} \\
&\quad \left\{k(k-1)(r-1)^2-\left[(k-1)^2(r-1)^2+4r\right]\right. \\
&\quad \left.-(r+1)\sqrt{A}\right\} \\
&= \frac{r}{k^2\sqrt{A}}\left\{-(k-1)(r-1)^2-4r-(r+1)\sqrt{A}\right\} \\
&= -\frac{r}{k^2\sqrt{A}}\left\{\left[(k-1)(r-1)^2+4r+(r+1)\sqrt{A}\right]\right\} \\
&= -T
\end{aligned}$$

Since $T > 0$, $\frac{\partial G(r,k)}{\partial k} = T < 0$. Therefore $G(r,k)$ is decreasing in k. As a consequence, $h_y(z)$ is decreasing in k (If $r'(z) > 0$ as is the case for $z \leq z_{max}$).

– **We now consider the *limit* of $h_y(z)$ as $k \to \infty$:**

Let $s = (r-1) \Longleftrightarrow r = (s+1)$.

and $l = (k-1) \Longleftrightarrow k = (l+1)$.

Substituting in $G(r,k)$ we have

$$h_y(z) = z + \frac{G_s(s,l)}{s'(z)} \tag{5.35}$$

where $G_s(s,l) = \frac{s+1}{l+1}\left\{-ls + 2 + \sqrt{l^2s^2 + 4(s+1)}\right\}$. Further,

$$\begin{aligned} G_s(s,l) &= \frac{s+1}{l+1}\left\{2 - ls + \sqrt{l^2s^2+4(s+1)}\right\} \\ &= \frac{s+1}{l+1} l \left\{\frac{2}{l} - s + \sqrt{s^2 + \frac{4(s+4)}{l^2}}\right\} \\ &= \frac{s+1}{1+\frac{1}{l}}\left\{\frac{2}{l} - s + \sqrt{s^2 + \frac{4(s+4)}{l^2}}\right\}. \end{aligned}$$

Hence,

$$\begin{aligned} \lim_{l\to\infty} G_s(s,l) &= (s+1)\{-s + \begin{cases} s & if \quad s > 0 \\ -s & if \quad s < 0, \quad (-1 < s < 0) \end{cases} \\ &= \begin{cases} 0 & if \quad s > 0 \\ -2s(s+1) & if \quad s < 0, \quad (-1 < s < 0) \end{cases} \end{aligned}$$

And $\lim_{l\to\infty} G_s(s,l)$ is always positive since $[-2s(s+1)]$ is positive on $-1 < s < 0$.

Therefore, the limit of $h_y(z)$ in equation 5.35 is :

$$\lim_{l\to\infty} h_y(z) = z + \begin{cases} 0 & if \quad s > 0 \\ \frac{-2s[s+1]}{s'(z)} & if \quad -1 < s < 0 \end{cases}.$$

Now

$$\begin{aligned} s = 0 \quad \Longleftrightarrow \quad r = 1 \quad &\Longleftrightarrow \quad w(z) = w(y) \\ &\Longleftrightarrow \quad z = y_l \end{aligned}$$

where, assuming $z < z_{max}$, $y_l < z_{max} < y$ and $w(y_l) = w(y)$.

So

$$\lim_{l\to\infty} h_y(z) = \begin{cases} z & if \quad s>0 \quad r>1 \quad i.e. \quad \frac{w(z)}{w(y)} > 1 \\ z - \frac{2s[s+1]}{s'(z)} & if \quad -1<s<0 \quad 0<r<1 \end{cases}$$

$$= \begin{cases} z & if \quad w(z) > w(y_l) \\ z - \frac{2s[s+1]}{s'(z)} & if \quad w(z) < w(y_l) \end{cases}$$

So

$$\lim_{l\to\infty} h_y(z) = \begin{cases} z + \frac{0}{s'(z)} & if \quad z > y_l \\ z - \frac{2s[s+1]}{s'(z)} & if \quad z < y_l \end{cases}$$

This suggests $h_y(y_l) = y_l$ is a minimum TP.

In particular, for the logistic weight function, $z_{max} = 0$ and $w(\cdot)$ is symmetric, so that $y_l = -y$. Hence,

$$\lim_{l\to\infty} h_y(z) = \begin{cases} z & if \quad z > -y \\ z - \frac{2w(z)[w(z)-w(0)]}{w'(z)} & if \quad z < -y \end{cases}$$

$$= \begin{cases} z & if \quad z > -y \\ z - \frac{2w(z)[w(z)-\frac{1}{4}]}{w'(z)} & if \quad z < -y \end{cases}$$

5.9.1 Explicit Solutions for some weight functions

We extend results of Torsney and Musrati(1993) for the Gamma, Beta, Normal weight functions.

We find explicit formulae for the D-optimal design weights for some weight functions

Case 1 : Symmetric Beta Weight Function

$w(b) = (1-b^2)^{\gamma-1}$, $\gamma > 1$, $-1 \leq b \leq 1$.

This weight function is symmetric about the origin, for all γ. Hence one optimal design must put equal weight on the 2^{k-1} points satisfying $z_1 = \pm b$, $z_j = \pm 1$, $j = 2, \cdots, l$ for some b, which can be determined by maximizing the determinant of the information matrix with respect to b; that is maximise

$$\psi(b) \quad = \quad 2\ln b + k\ln[(1-b^2)^{\gamma-1}].$$

Note that b can not assume the values 1 or -1, since $w(-1) = w(1) = 0$. Therefore the first order conditions for b is

$$\frac{\partial\psi(b)}{\partial b} \quad = \quad \frac{2}{b} + \frac{2b(\gamma-1)k}{(1-b^2)}$$

which implies

$$b \quad = \quad \pm\frac{1}{\sqrt{k(\gamma-1)+1}}. \tag{5.36}$$

For instance, if we let $\gamma = 3$ and $k = 3$ in equation (5.36), then the support points of the four-point design on $\mathcal{Z}_w = \{(z_1\, z_2) \,:\, -1 \leq z_j \leq 1 \quad j = 1, 2\}$ are $z_1 = \pm 0.378$, $z_2 = \pm 1$ with optimal weights $\frac{1}{4}$. These symmetric 2^{k-1} points design are globally D-optimal because they satisfy the necessary and sufficient condition of the equivalence theorem; that is they satisfy equation (5.1).

Case 2: Normal Weight Function $w(b) = e^{-b^2/2}$, $-\infty \leq b \leq \infty$.

This weight function is also symmetric about the origin. Hence one optimal design must have the same form as in Case 1. Now we have to maximise

$$\psi(b) = 2\ln b - \frac{kb^2}{2}.$$

Therefore first order conditions for b are

$$\frac{\partial\psi(b)}{\partial b} = \frac{2}{b} + kb$$

which implies

$$b = \pm\sqrt{\frac{2}{k}}. \tag{5.37}$$

For instance, if we let $k = 4$ in equation (5.37), then the support points of the eight-point design on

$Z_w = \{(z_1, z_2, z_3) : -\infty < z_1 < \infty, \ -1 \leq z_j \leq 1, \ j = 2, 3\}$ are

$z_1 = \pm 0.707107$, $z_j = \pm 1$, $j = 2, 3$ with optimal weights $\frac{1}{8}$. These symmetric 2^{k-1} points design are globally D-optimal because they satisfy the necessary and sufficient condition of the equivalence theorem; that is they satisfy equation (5.1).

Musrati(1992) and Torsney and Musrati (1993) reported these results for two parameter model.

5.9.2 Some Empirical Results for D-optimal designs

The general objective has been to find empirically D-optimal designs when $\mathcal{Z} = \{(z_1, \cdots, z_l) : a \leq z_1 \leq b, \quad -1 \leq z_j \leq 1 \quad j = 2, \cdots, l\}$ for all possible choices of a, b. In section (5.6) for the most weight functions we showed that two distinct values of z_1 produce the support points of the conjectured optimal designs of the various cases of $\mathcal{Z} = [a, b]$. Now we will show empirically that the equivalence theorem is satisfied by our conjectured optimal designs for all possible design intervals $[a, b]$. There are only two distinct value of z_1 and hence observations are taken at only two values of z_1.

Case 1 : $\mathcal{Z} = \mathcal{Z}_w = \{(z_1, \cdots, z_l) : -\infty \leq z_1 \leq \infty \; -1 \leq z_j \leq 1 \, j = 2, \cdots, l\}$

and $Supp(p^*) = \{-b^*, b^*\}$

In the case of **symmetric** weight functions $w(z_1)$, z_1- support points are $\pm b^*$ with $z_j = \pm 1$, $j = 2, \cdots, l$ and with equal weights where b^* maximizes $\{detM(p) = b^2[w(b)]^k\}$. We found the b^* value that maximizes $detM(p)$ for the **k** parameter case with the logistic, probit, normal and symmetric beta weight functions. Empirical D-optimal designs for five choices of $w(\cdot)$ are listed in the Table (5.1). We checked for optimality of this design, by checking the equivalence theorem for $z_1 = (-\infty, \infty)$, $z_j = \pm 1$, $j = 2, \cdots, 8$. Additionaly, Figure (5.14) represents the variance function for the Global D-optimal design on Z_w for the Logistic weight function, for the $k = 4$ parameter case. We consider further examples for this choice of weight function again with $k = 4$.

Case 2 : $\mathcal{Z} = \{(z_1, z_2, z_3) : a \leq z_1 \leq b, \quad -1 \leq z_j \leq 1, \quad j = 2, 3\}$

$a < a^*$, $b < b^*$ and $Supp(p^*) = \{max\{a, a^*(b)\}, b\}$

Results are very similar with next step. So we only show include emprical results for that.

Case 3 : $\mathcal{Z} = \{(z_1, z_2, z_3) : a \leq z_1 \leq b, \quad -1 \leq z_j \leq 1, \quad j = 2, 3\} \quad a > a^*$, $b > b^*$ and $Supp(p^*) = \{a, min\{b, b^*(a)\}$

For the 4 parameter logistic regression model we used as in Chapter 3, section 3.2.6, Case 3 **an alternating algorithm** to determine $b^*(a)$ for $k = 4$ and $a = -1.04$, -1.00, $-0.90 \cdots$ 1.00,1.04. The D-optimal support points and weights are summurized in Table 5.3. Figures 5.15 and Figures 5.16 show that the necessary and sufficient conditions of the equivalence theorem are satisfied.

Case 4 : $\mathcal{Z} = \{(z_1, z_2, z_2) : a \leq z_1 \leq b, \quad -1 \leq z_j \leq 1, \quad j = 2, 3\}$

$a > a^*$, $b < b^*$ and $Supp(p^*) = \{a, b\}$

For this z_1 interval the end points are the support points and the equivalence theorem is satisfied. See Figure 5.17.

5.9.3 Efficient Approximations

This section will be devoted to finding the efficiency of D-optimal designs based on Probit Regression Model and Normal regression models. To compare different designs, we will use a modification of the efficiency measure used by Atkinson and Donev (1992) , and proposed by Abdelbasit and Plackett(1983).

First we look at the ratio between probit regression model support points and normal regression model support points. As we can see from the Table5.1 the ratio of the probit model support points and normal model support points are approximately equal to $(1.15\sqrt{\frac{2}{k}})$.

We suggest that the design for the Normal regression model which, although not optimal for Probit response model, gives an efficient alternative to the optimal design for probit response model. To explore this we investigate the relevant D-**efficiency** which is based on the determinant of the information matrix. Let ξ^* be the optimal design for a $\boldsymbol{k}$ parameter Probit regression model, and $\xi^\dagger$ be the optimal design for the normal density weight function $\boldsymbol{k}$ parameter model. Determinant values under the probit regression model are:

$$\begin{aligned} detM(\xi^*) &= (b^*_{1(k)})^2[w_2(b^*_{1(k)})]^k \\ detM(\xi^\dagger) &= (b^*_{2(k)})^2[w_2(b^*_{2(k)})]^k \end{aligned}$$

where $b^*_{1(k)}$ identifies the global support points for the normal density weight function $\boldsymbol{k}$ parameter model, $b^*_{2(k)}$ identifies global support points for the probit regression $\boldsymbol{k}$ parameter model and w_2 represents the probit regression model weight function. Note, there is an expilicit solution for the normal density weight functions : $b^*_{1(k)} = \pm\sqrt{\frac{2}{k}}$.[See section 5.3.4, case 4].

Now we measure the efficiency of $\xi^\dagger$ relative to ξ^* as

$$D_{eff} = eff(\xi^\dagger, \xi^*) = \left\{ \frac{detM(\xi^\dagger, z_1)}{detM(\xi^*, z_1)} \right\}^{1/k} \tag{5.38}$$

$$= \left\{ \frac{(b^*_{2(k)})^2 [w_2(b^*_{2(k)})]^k}{(b^*_{1(k)})^2 [w_2(b^*_{1(k)})]^k} \right\}^{1/k}$$

$$= \left\{ \frac{b^*_{2(k)}}{b^*_{1(k)}} \right\}^{2/k} \left\{ \frac{w_2(b^*_{2(k)})}{w_2(b^*_{1(k)})} \right\}. \tag{5.39}$$

Note: If $D_{eff} = eff(\xi^\dagger, \xi^*)$ is high (e.g. around 90%) then $\xi^\dagger$ is an **approximately** optimal design for the probit weight function.

Results are given in Table5.2. According to these, the design which is optimal for a normal regression model is an efficient alternative to the optimal design for the probit regression model.

Global D-optimal upper z_1 support points for some Symmetric Weight Functions : Z_w[a]						
NUMBER OF PARAMETERS	LOGISTIC $\frac{e^{-z}}{(1+e^{-z})^2}$	PROBIT $\frac{1}{\sqrt{2\pi}}e^{\frac{-z^2}{2}}$	BETA γ=2 $(1-z^2)^{\gamma-1}\ \gamma > -1\ (-1,1)$	BETA γ=1.5 $(1-z^2)^{\gamma-1}\ \gamma > -1\ (-1,1)$	NORMAL e^{-z^2}	PROBIT/ NORMAL
2	1.54340	1.13810	0.57735	0.70711	1	1.14
3	1.22291	0.937563	0.50000	0.63245	0.816497	1.15
4	1.04363	0.815901	0.44721	0.57752	0.707107	1.15
5	0.925358	0.731994	0.40825	0.53452	0.632456	1.16
6	0.839882	0.669619	0.37796	0.50000	0.577350	1.20
7	0.774406	0.620896	0.35355	0.47140	0.534522	1.16
8	0.722180	0.581473	0.33333	0.44721	0.500000	1.16

Table 5.1: For Some Symmetric Weight Functions global D-optimal support points

[a] Z_w is widest possible design space

NUMBER OF PARAMETERS	SUPP OF PROBIT $\frac{1}{\sqrt{2\pi}}e^{\frac{-z^2}{2}}$	SUPP OF (NORMAL)^2 e^{-z^2}	SUPP OF NORMAL $e^{-z^2/2}$	D_{eff} Normal/Probit	D_{eff} $(normal)^2/probit$
2	1.13810	1	0.707107	0.984049	0.840160
3	0.937563	0.816497	0.577335	0.988044	0.874593
4	0.815901	0.707107	0.500000	0.990475	0.913689
5	0.731994	0.632456	0.447214	0.992103	0.929826
6	0.669619	0.577350	0.408248	0.993249	0.940869
7	0.620896	0.534522	0.377964	0.958106	0.948912
8	0.581473	0.500000	0.353553	0.994779	0.955029

Table 5.2: D-efficiency of normal weight function for probit weight function

Four parameter case: For Logistic weight Function, $z_1 = [A, B] = [b_1, \infty)$ **for fixed** b_1 **and** $b_1 > -b^*$ **the lower support point, optimal** b_2 p_1 **and** p_2 **value.**			
fixed b_1 **value**	$b_2^*(b_1)$	$p_{b_2}(b_1)$	$p_{b_1}(b_2)$
-1.04363	1.043625	0.125000	0.125000
-1.00000	1.074189	0.127208	0.122792
-0.90000	1.148056	0.132245	0.117755
-0.80000	1.226839	0.137156	0.112844
-0.70000	1.309847	0.141832	0.108168
-0.60000	1.396236	0.146186	0.103814
-0.50000	1.485121	0.150159	0.099841
-0.40000	1.575692	0.153728	0.096272
-0.30000	1.667281	0.156892	0.093108
-0.20000	1.759399	0.159673	0.090327
-0.10000	1.851721	0.162104	0.087896
0	1.944058	0.164222	0.085778
0.10000	2.036324	0.166066	0.083934
0.20000	2.128502	0.167670	0.082330
0.30000	2.220616	0.169068	0.080932
0.40000	2.312716	0.170288	0.079712
0.50000	2.404861	0.171354	0.078646
0.60000	2.497112	0.172289	0.077711
0.70000	2.589526	0.173110	0.076890
0.80000	2.682153	0.173832	0.076168
0.90000	2.775035	0.174469	0.075531
1.00000	-	-	-
1.04363	-1.043625	0.125000	0.125000

Table 5.3: For Logistic weight function D-optimal support points and weights.

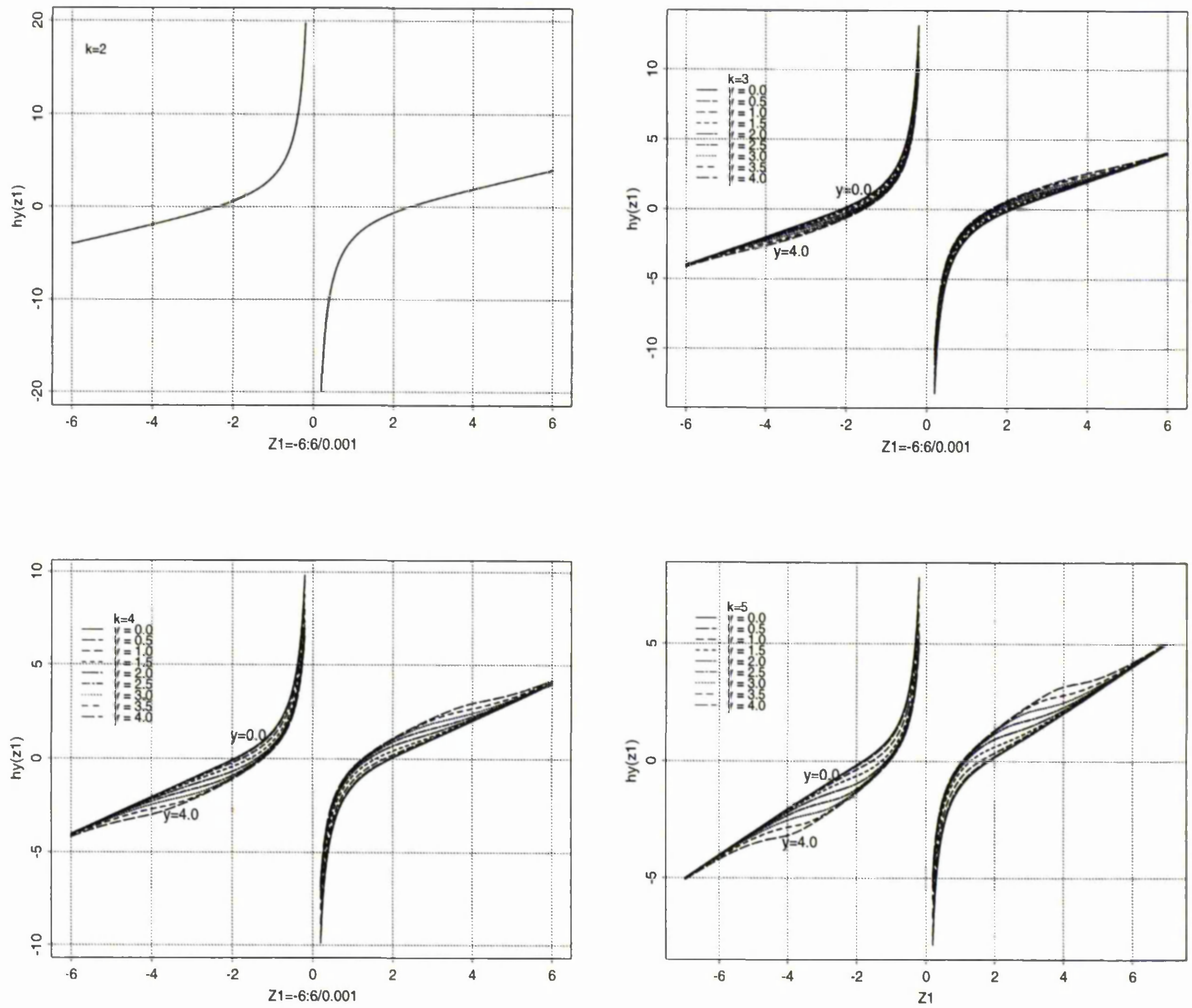


Figure 5.2: Plots of $h_y(z)$ for the Logistic Weight function, $y = 0.0, \cdots, 4.0$ for each of $k = 2, 3, 4, 5$.

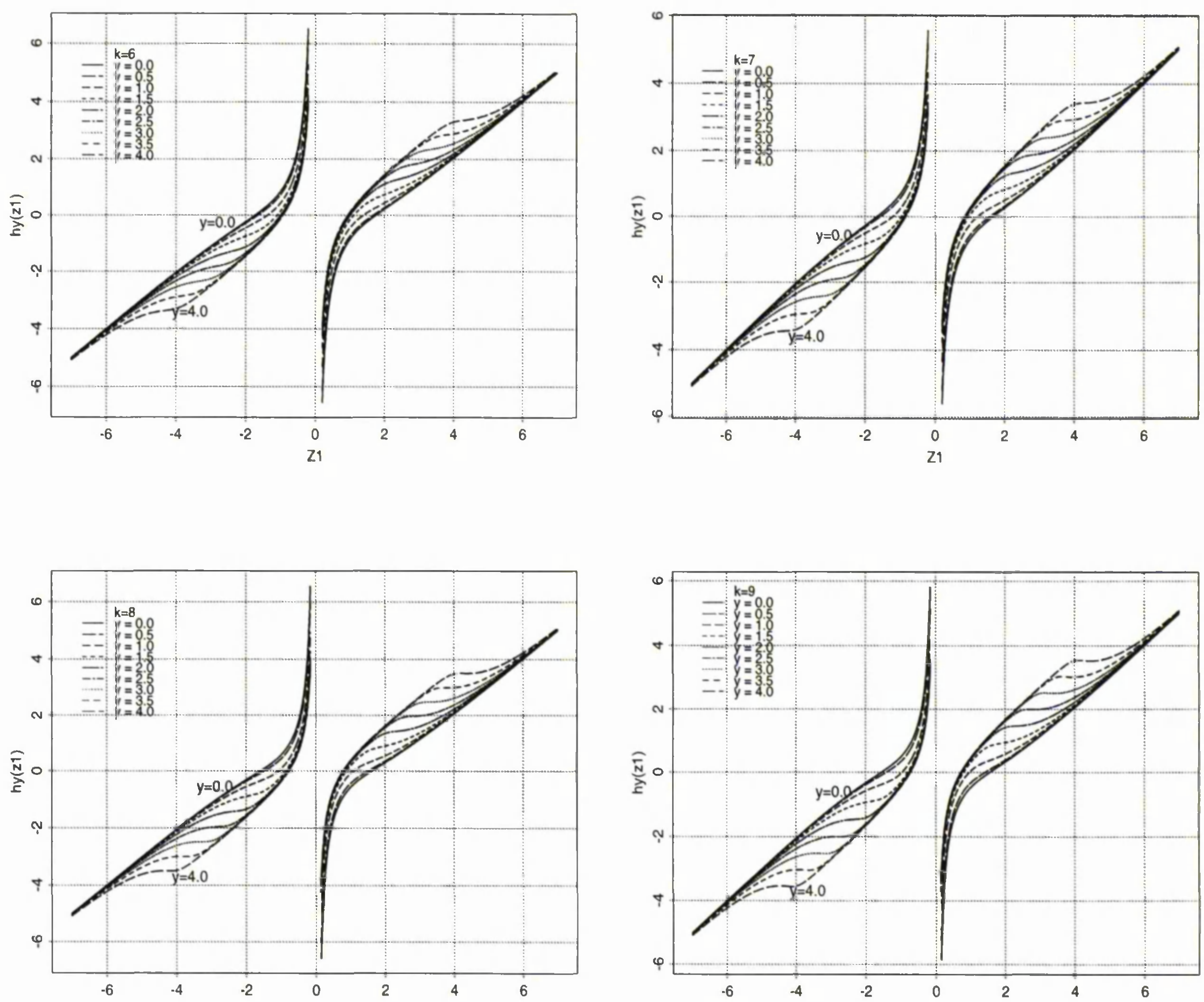


Figure 5.3: Plots of $h_y(z)$ for the Logistic Weight function, $y = 0.0, \cdots, 4.0$ for each of $k = 6, 7, 8, 9$.

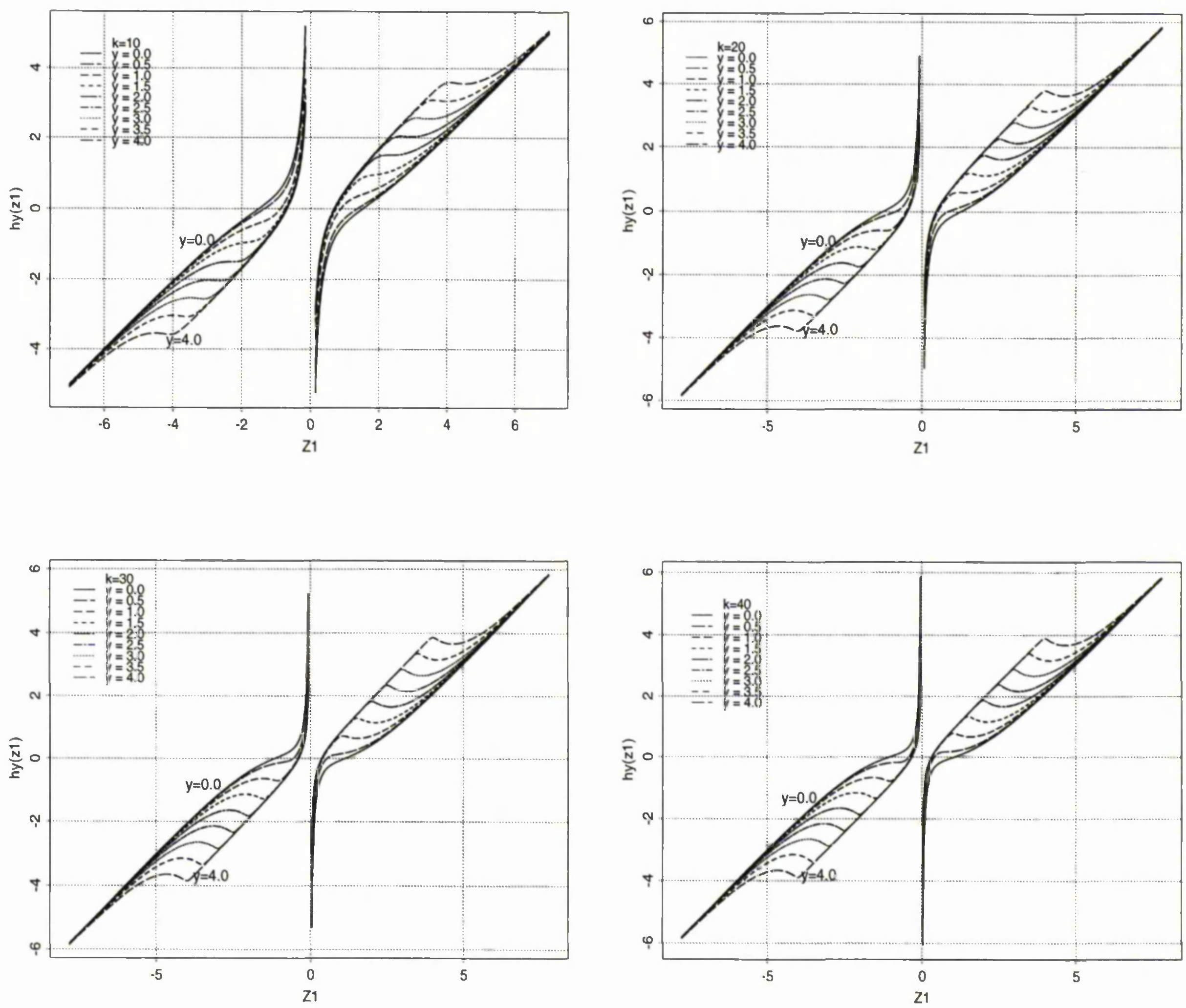


Figure 5.4: Plots of $h_y(z)$ for the Logistic Weight function, $y = 0.0, \cdots, 4.0$ for each of $k = 10, 20, 30, 40$.

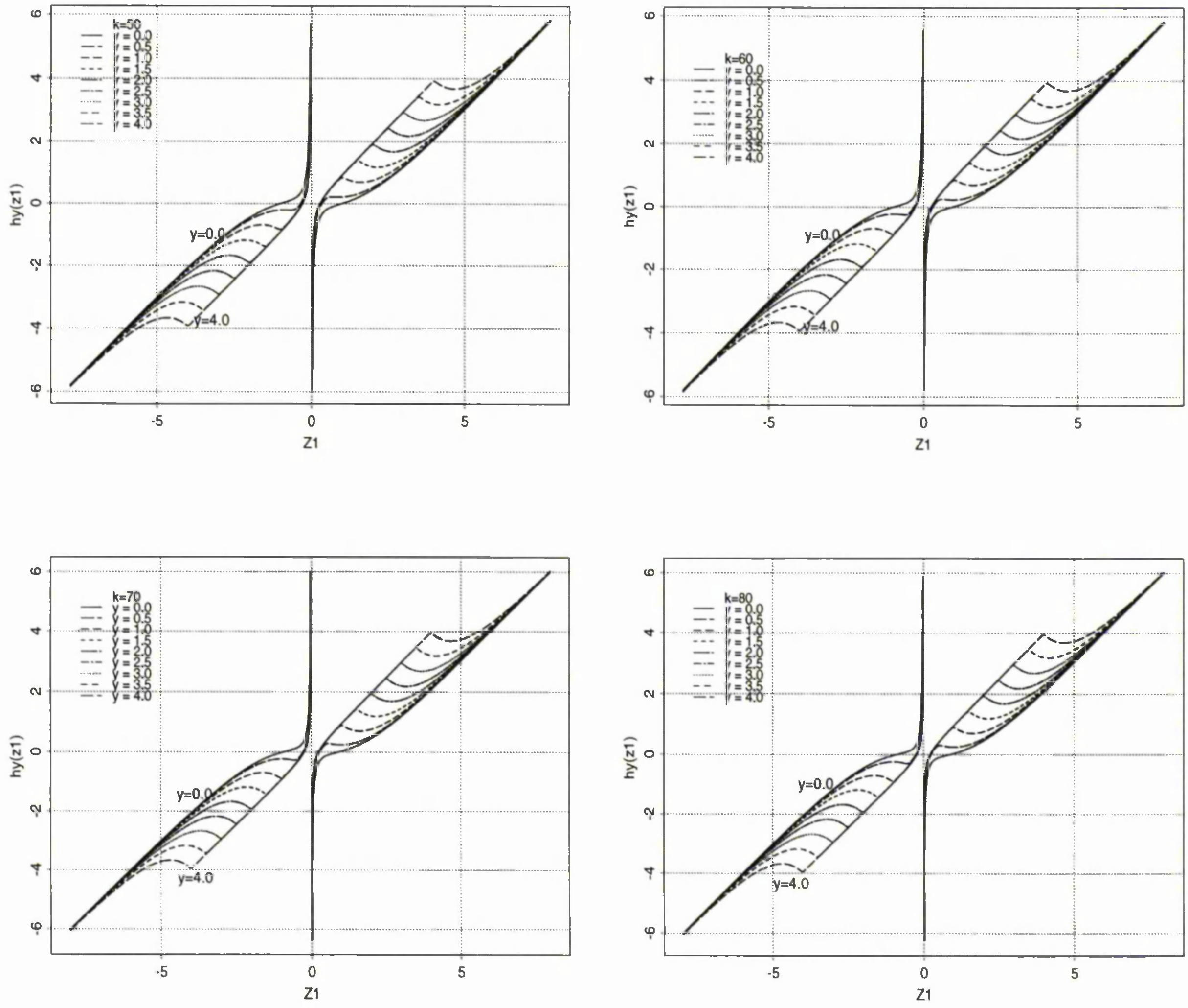


Figure 5.5: Plots of $h_y(z)$ for the Logistic Weight function, $y = 0.0, \cdots, 4.0$ for each of $k = 50, 60, 70, 80$.

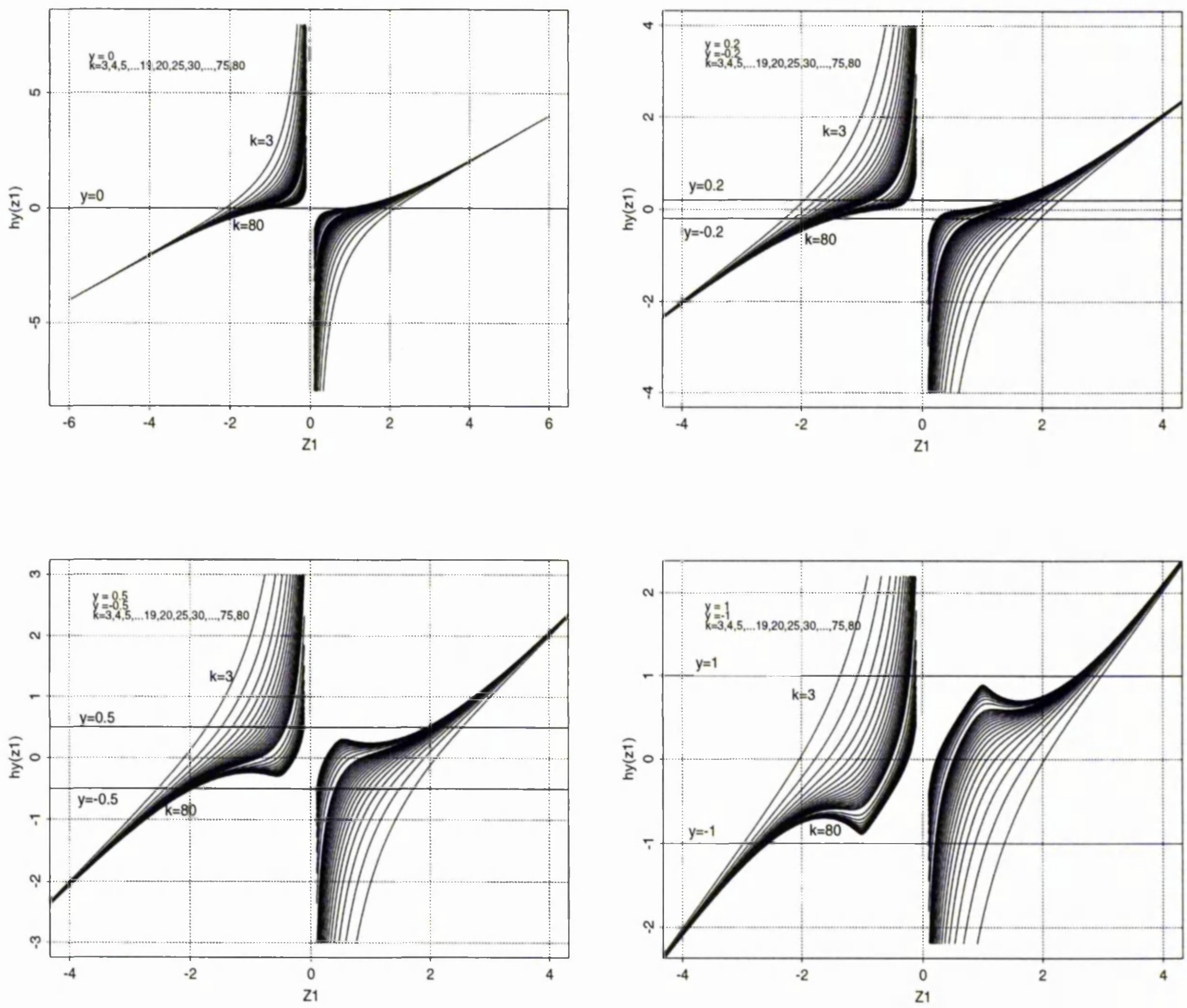


Figure 5.6: Plots of $h_y(z)$ for the Logistic Weight function, $k = 3, 4, 19, 20, 25, \cdots, 75, 80$ for each of $y = 0.0, 0.2, 0.5, 1.0$.

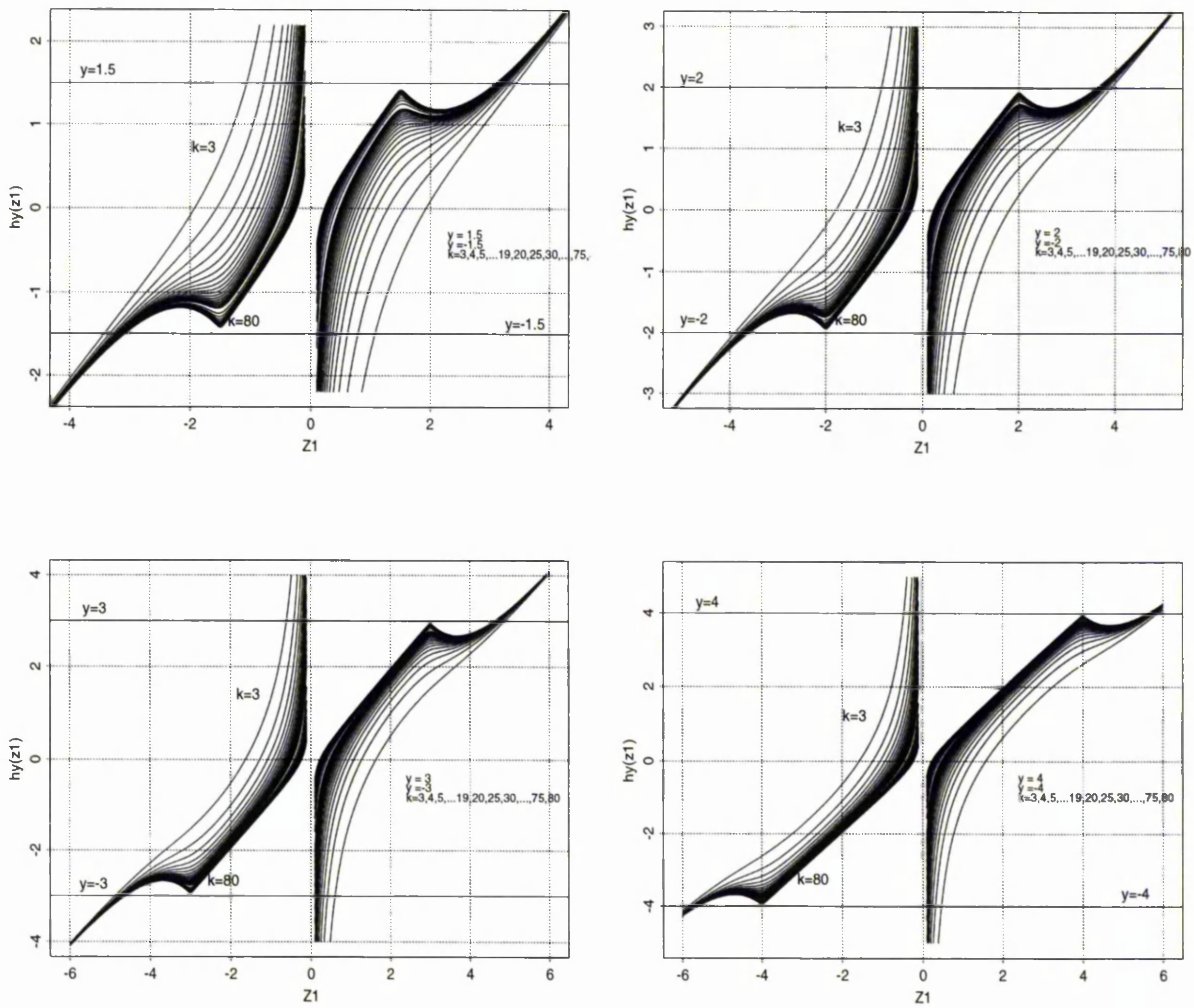


Figure 5.7: Plots of $h_y(z)$ for the Logistic Weight function, $k = 3, 4, 19, 20, 25, \cdots, 75, 80$ for each of $y = 1.5, 2.0, 3.0, 4.0$.

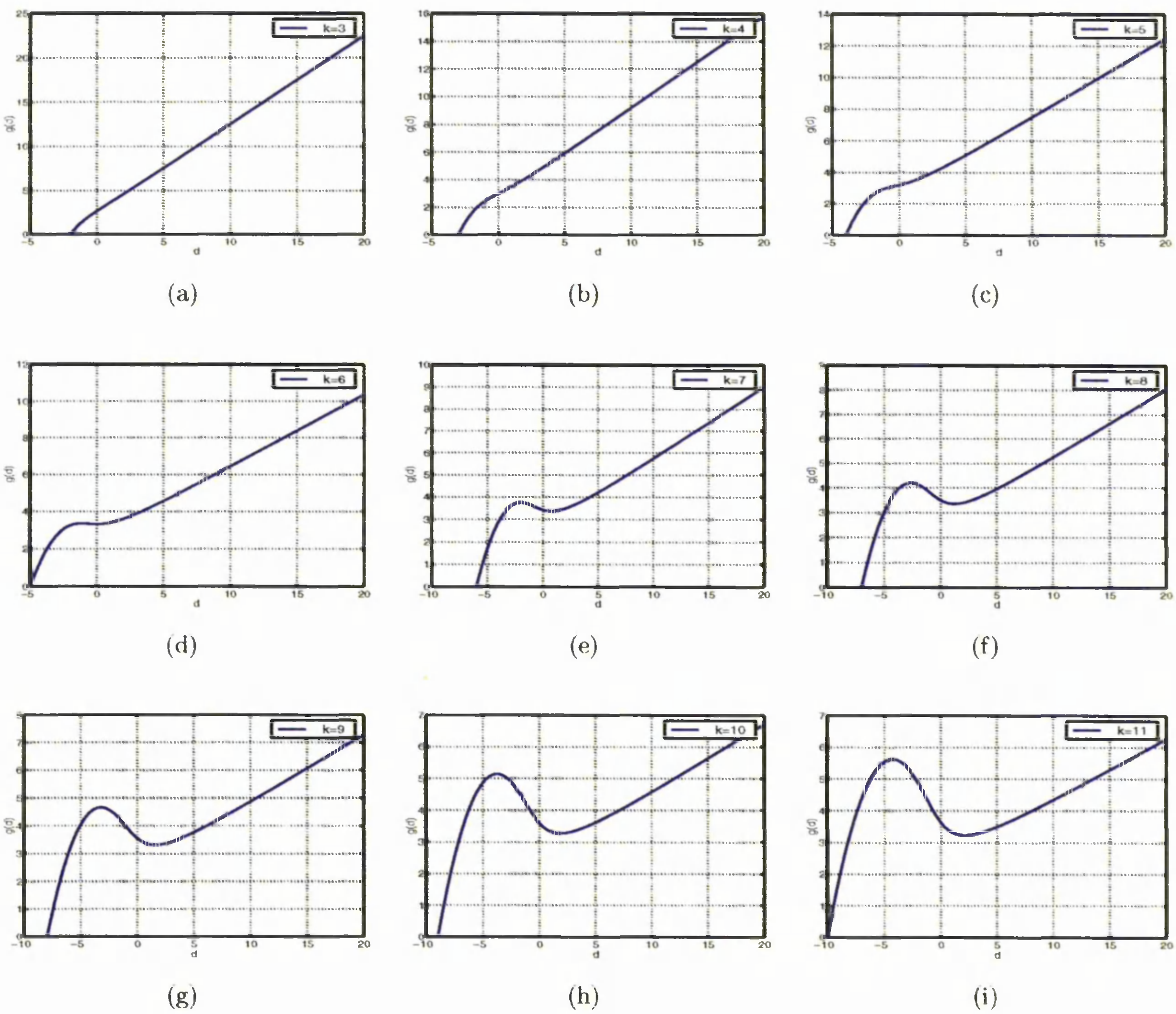


Figure 5.8: Plots of $g(d)$ for $k = 3, 4, 5, \cdots, 10, 11$ and $d \geq -(k-1)$.

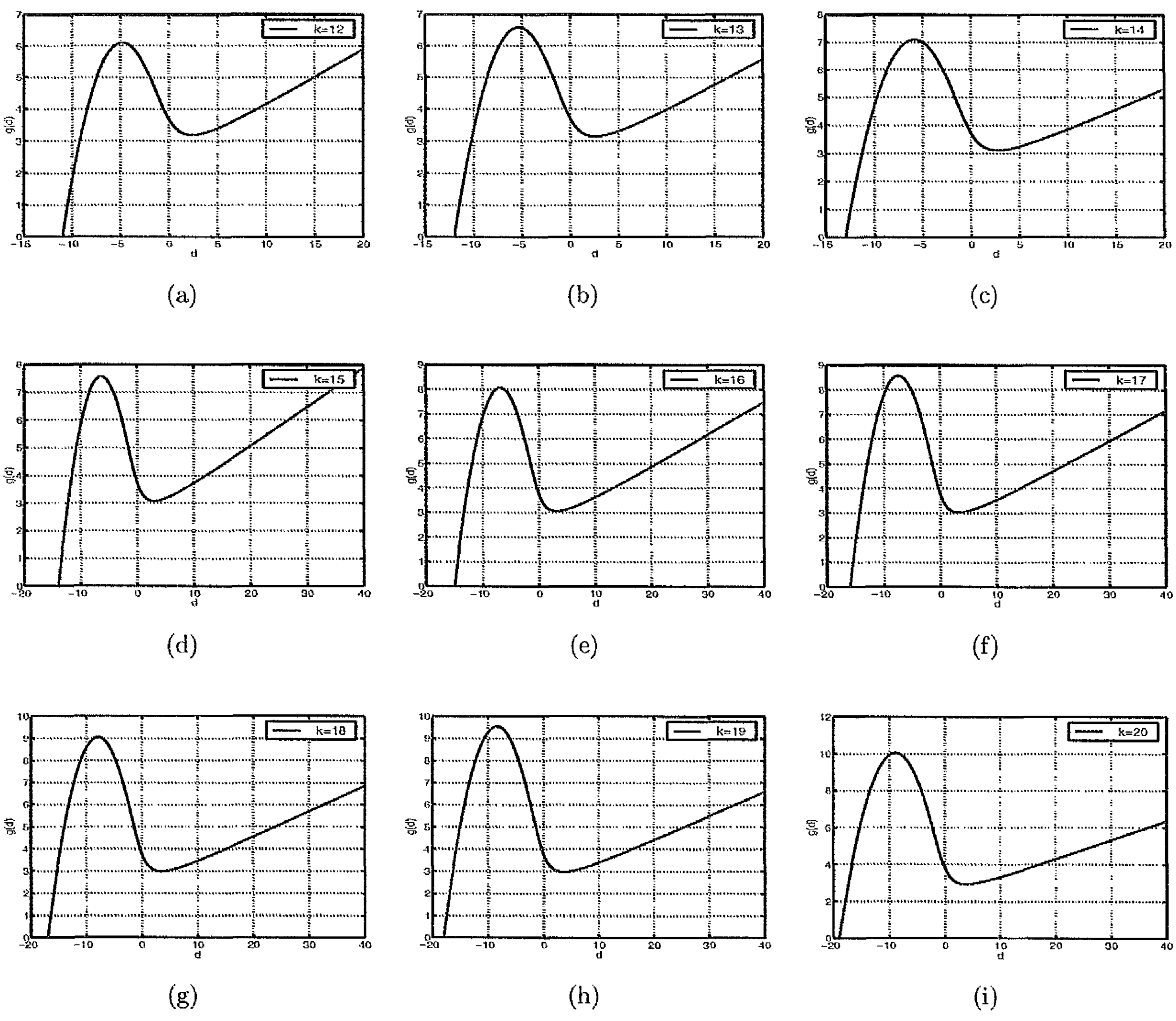


Figure 5.9: Plots of $g(d)$ for $k = 12, 13, \cdots, 19, 20$ and $d \geq -(k-1)$.

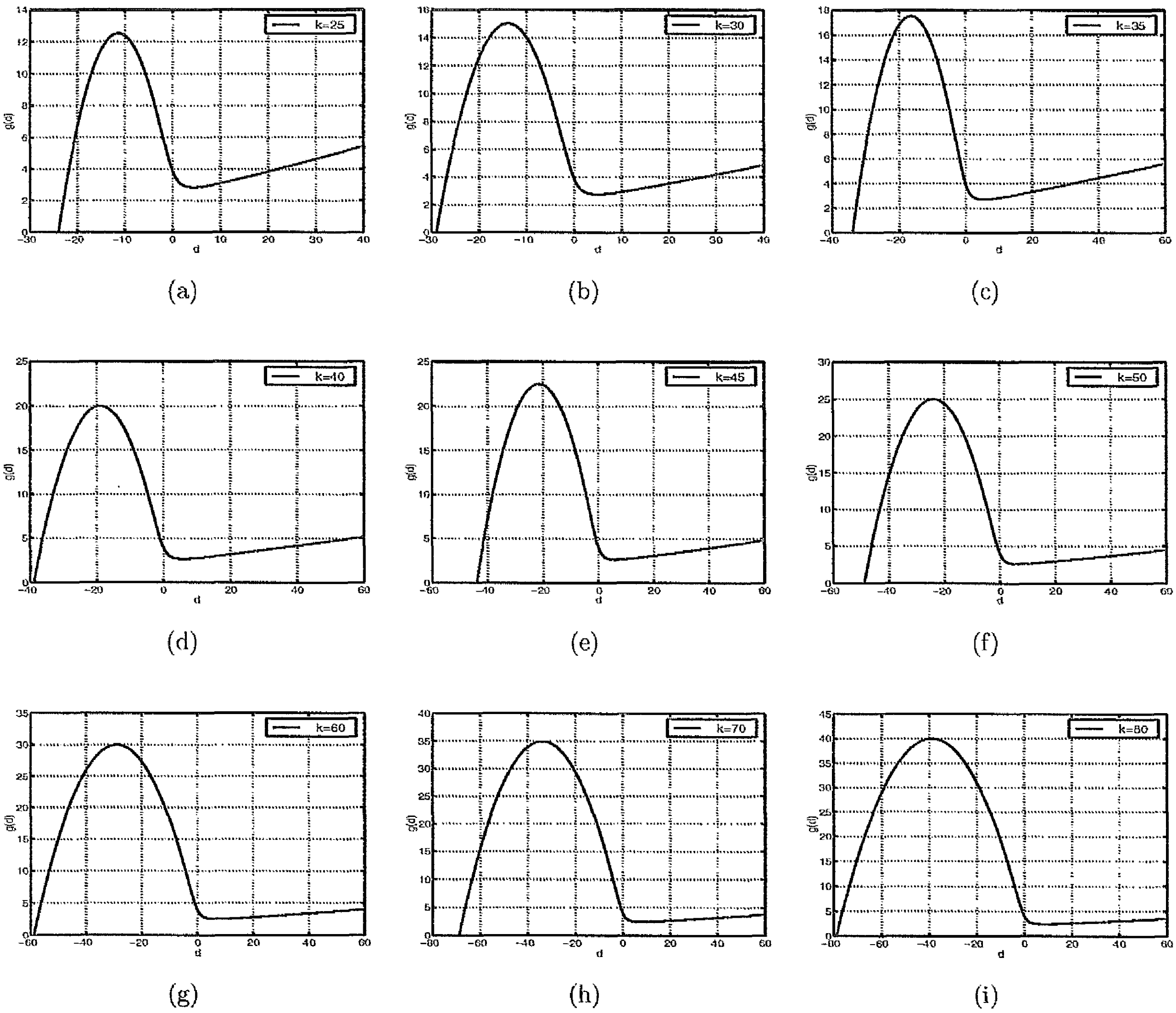


Figure 5.10: Plots of $g(d)$ for $k = 25, 30, \cdots, 75, 80$ and $d \geq -(k-1)$.

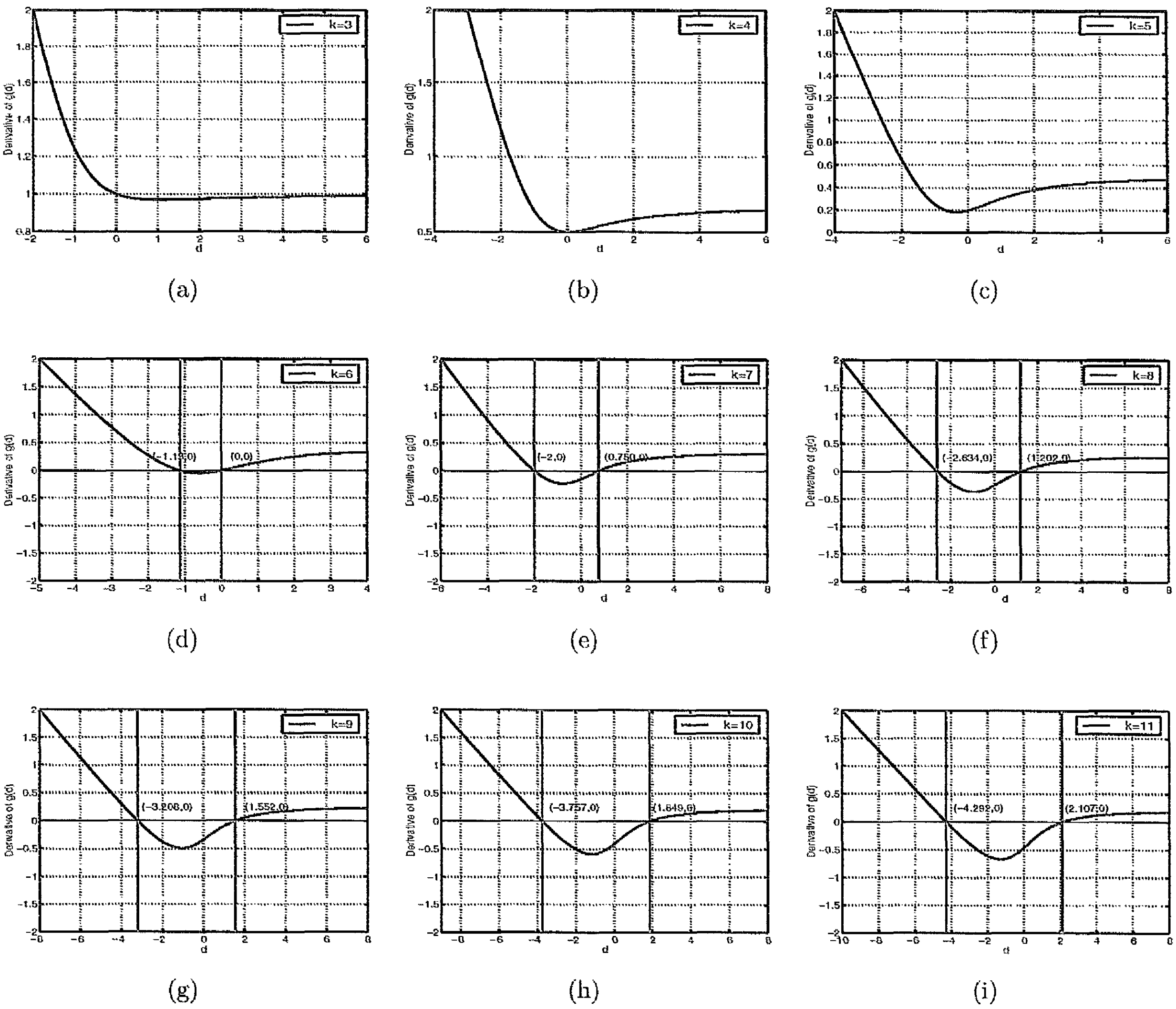


Figure 5.11: Plots of $\frac{\partial g(d)}{\partial d}$ for $k = 3, 4, 5, \cdots, 10, 11$ and $d \geq -(k-1)$.

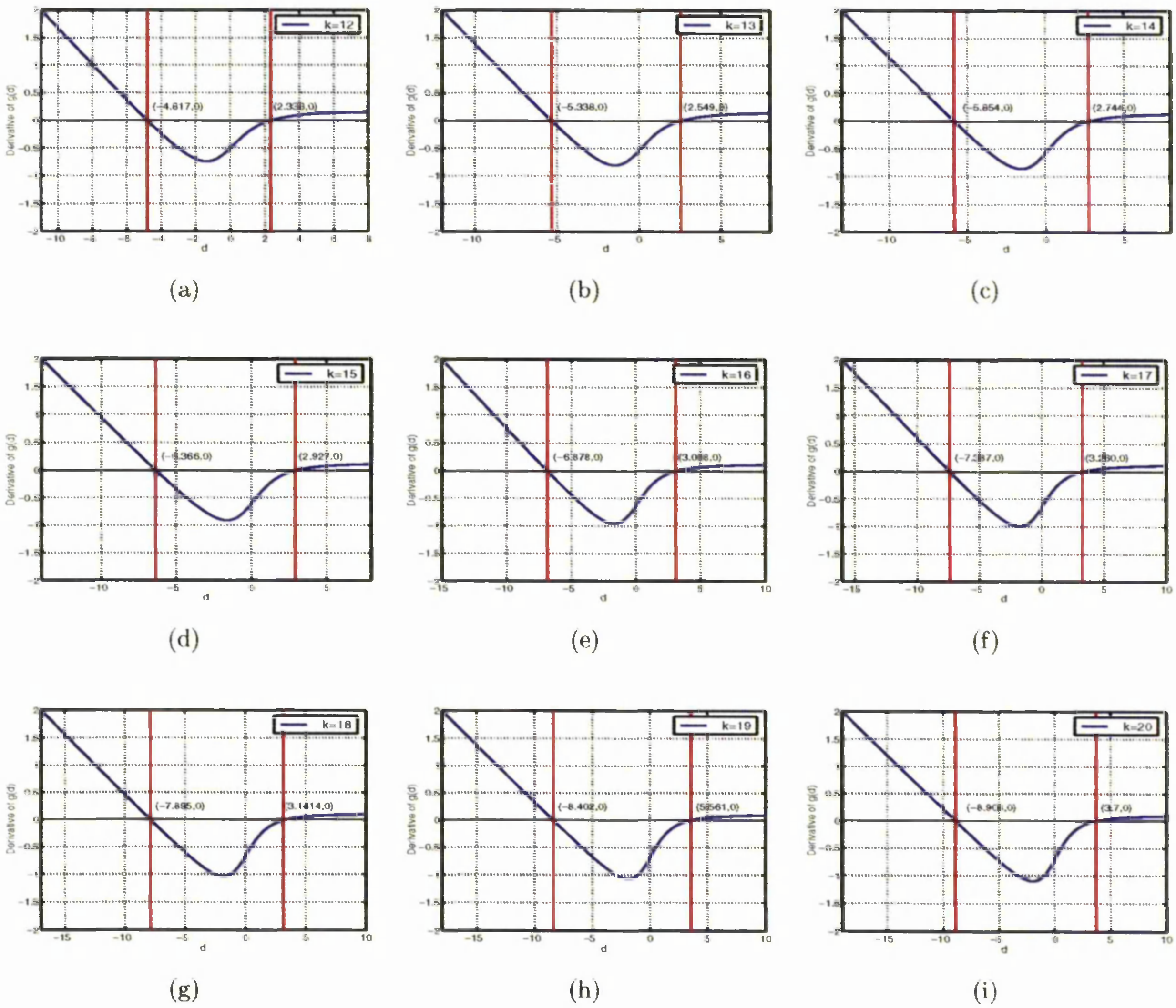


Figure 5.12: Plots of $\frac{\partial g(d)}{\partial d}$ for $k = 12, 13, \cdots, 19, 20$ and $d \geq -(k-1)$.

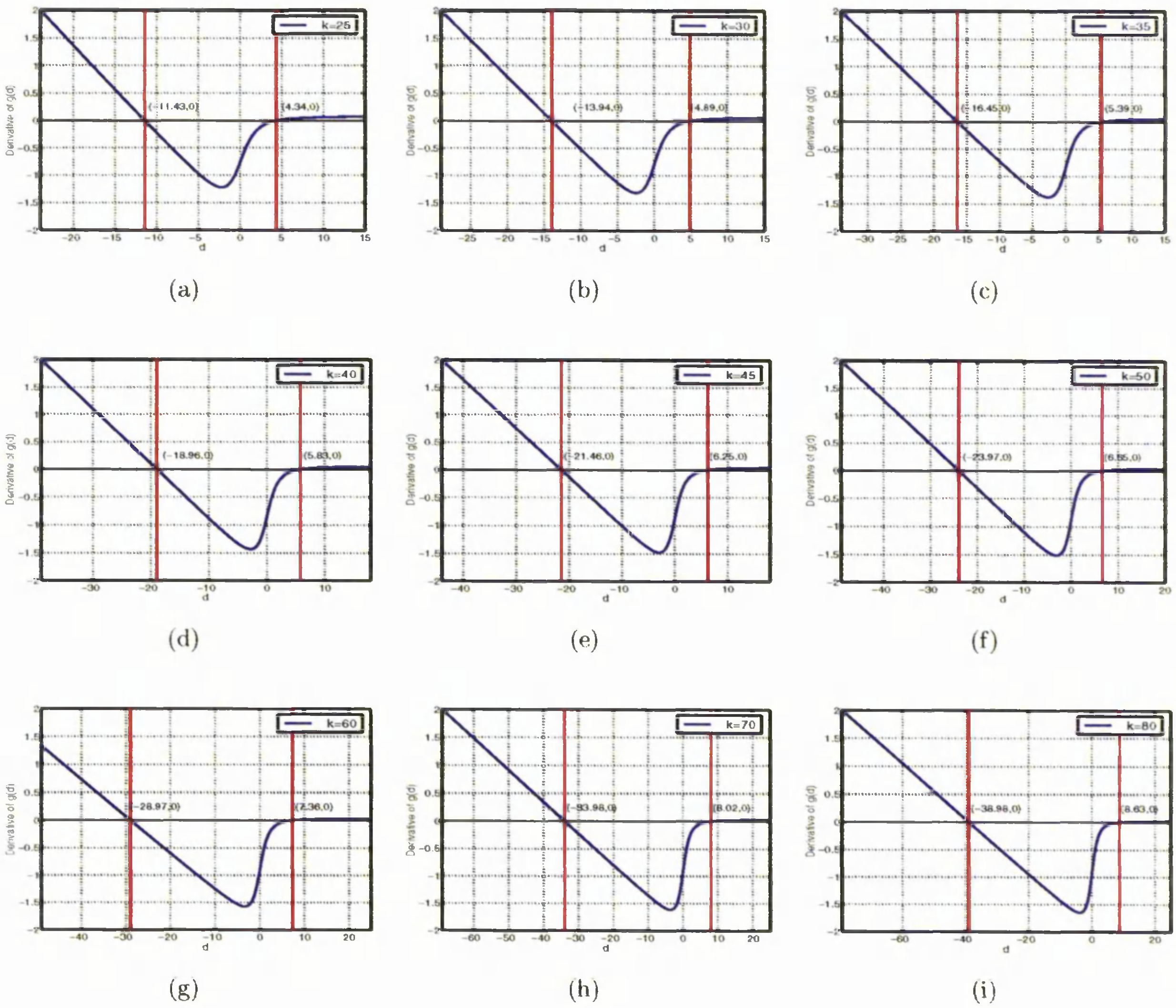
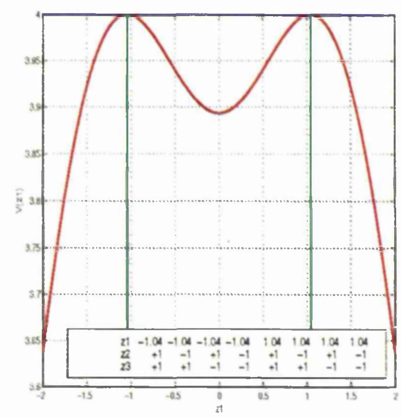


Figure 5.13: Plots of $\frac{\partial g(d)}{\partial d}$ for $k = 25, 30, \cdots, 75, 80$ and $d \geq -(k-1)$.



(a)

Figure 5.14: Plot of the variance function for the global symmetric D-optimal four-point design on $\mathcal{Z} = \mathcal{Z}_w = \{(z_1, z_2, z_3) : -\infty \leq z_1 \leq \infty \ -1 \leq z_j \leq 1,\ j = 2, 3\}$ for the logistic weight function($k = 4$).

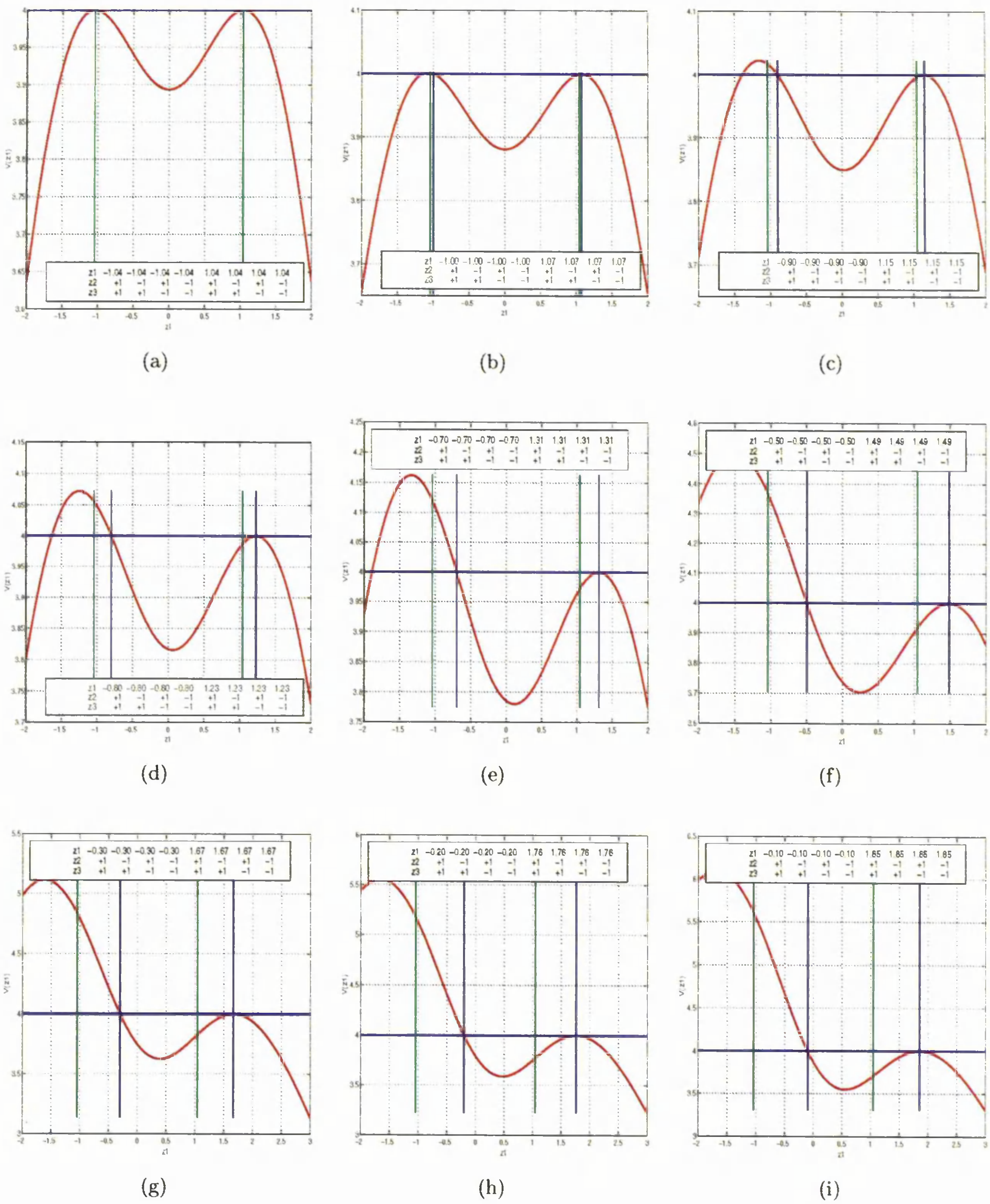
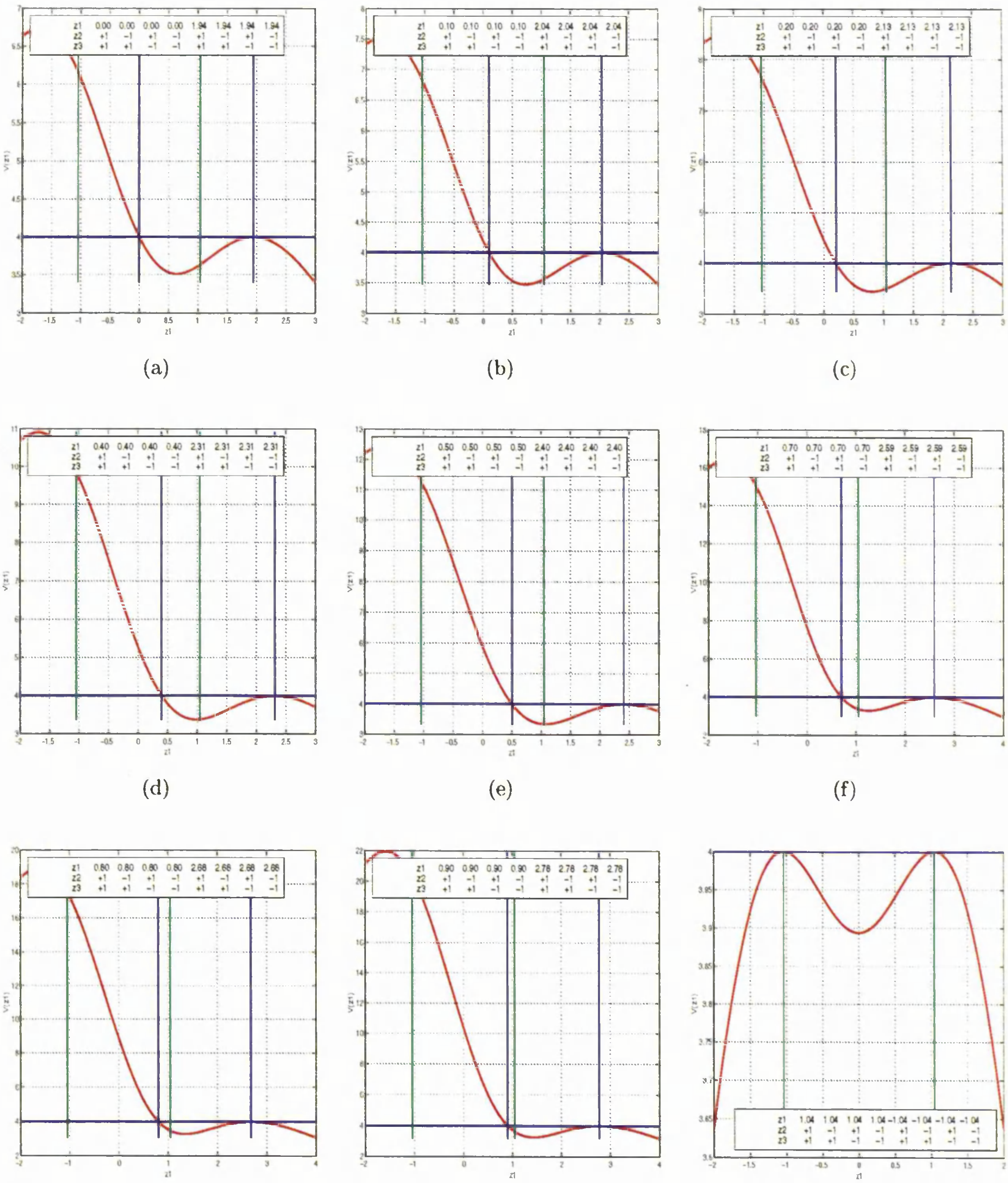


Figure 5.15: Some plots for the variance function $v(z_1)$ under an optimal design on $\mathcal{Z} = \{(z_1, z_2, z_3) : a \leq z_1 \leq b \ z_j = \pm 1, \ j = 2, 3\}$, $a > a^*$, $b > b^*(a)$ for the logistic weight function, $(k = 4)$.

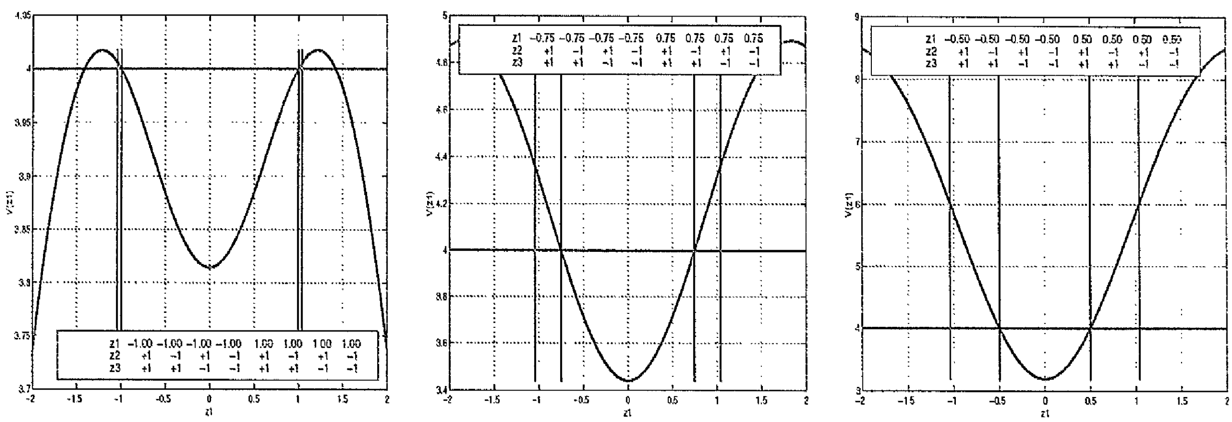


(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 5.16: Some plots for the variance function $v(z_1)$ under an optimal design on $\mathcal{Z} = \{(z_1, z_2, z_3) : a \leq z_1 \leq b \ \ z_j = \pm 1, \ j = 2, 3\}$, $a > a^*$, $b > b^*(a)$ for the logistic weight function, $(k = 4)$.



(a) (b) (c)

Figure 5.17: Some plots for the variance function $v(z_1)$ under an optimal design on $\mathcal{Z} = \{(z_1, z_2, z_3) : a \leq z_1 \leq b \ z_j = \pm 1, \ j = 2, 3\}$, $a > a^*$, $b < b^*$ for the logistic weight function, $(k = 4)$.

Chapter 6

Some Advances in Optimal Designs in Contingent Valuation Studies

6.1 Introduction

A Contingent Valuation Study is essentially a Sample Survey of a relevant population, the primary aim of which is to estimate that population's willingness to pay (WTP) for some new (or possible previously free) amenity or it might be to estimate what increase in charges the population is willing to pay for an established amenity. These amenities fall into the category of non-market goods or services. Examples of such studies arise in the areas of health or welfare e.g. payment for (provision) of medical programs see Donaldson (1993); recreation (e.g. payment for fishing permits); and the environment (e.g. payment for pollution reduction programs).

The first such study focussed on pollution in the Delaware River Basin, USA in 1947. A more recent example is seen in Hanley (1989) which reported a study into the Willingness to Pay of visitors to a part of the Queen Elizabeth Forest Park in Central Scotland. There was interest in four aspects; wildlife, landscape, recreation and all combined. Four WTP questions were asked. For the last category this was: 'Suppose the government was considering selling the Queen Elizabeth Forest Park to a private forestry company. This would mean people would no longer be able to visit it. If the only way to prevent this happening was for the Forestry Commission to raise revenue by selling day tickets to visitors, how much would you be willing to pay, per person per visit?' This kind of question is known as an open ended question.

6.2 Criticisms of CV Studies

An overriding criticism is that a CV question invites a hypothetical valuation, particularly if the respondent has not previously considered the issue in question. This could lead to biases of various kinds. Bishop and Heberlein (1979) construct a list of possibilities. Broadly speaking they fall into three categories: psychological, economic and statistical. Psychological biases include 'strategic bias' arising when a respondent 'overestimates' his/her WTP if (s)he suspects that payment will not become a reality, and 'free-riding' meaning WTP is 'underestimated' to keep real fees low. Economic biases include the 'embedding problem' arising if a respondent is unable to recognise other competing demands on a finite (recreational or environmental or health) budget. Finally CV studies are as subject to 'statistical' biases as any sample survey. For example length biased sampling is a potential problem with the study of Hanley (1989) since respondents were sampled on site. Two styles of enquiry were used : a self completion questionnaire and an interview. Those who stayed longer in the park were more likely to be interviewed.

6.3 Literature

These issues have given rise to a great deal of literature. Much of this appears in the following journals: Land Economics, American Journal of Agricultural Economics and The Journal of Environmental and Economic Management. This literature is rich in its use of statistical tools including regression, and binary regression methods since other potentially relevant questions are regularly included such as general questions on income and age. We shall also see the need for optimal regression designs. One apparent lack seems to be reference to methods of analyses developed in the survey methodology arena. The literature also includes a batch of 8 papers in Volume 34 of the Natural Resources Journal a legal publication, arising from contentious litigation concerning the use of CV studies in relation to the Exon Valdez Oil Spillage in Alaska.

6.4 Variations of the WTP question

With a view to resolving some of the criticisms of CV studies a blue ribbon panel was set up in the USA chaired by Arrow (the Nobel prize winning economist) and Solow. This produced a list of 15 recommendations for the conduct of CV studies. One of these stated that a dichotomous choice WTP question should be used, one of several alternatives to the open ended question which have evolved. These include:

1. Closed ended format (or payment card):

   The respondent is offered a list (possibly on a card) of possible payments and asked to identify the one closest to his/her maximum WTP. This variation was also used in the Hanley (1989) study. It produced higher mean WTP values than the open ended case for all four aspects of interest, 'significantly so' in the case of two of them.

2. Dichotomous Choice Format

   The respondent is offered a single payment or bid and simply responds Yes or No according to his/her willingness to pay this bid. This format is also known as a Discrete Choice or Single Bounded question.

3. Double Bounded Format

   Here the respondent is offered two dichotomous choice questions. If the answer to the first bid is YES, a higher bid is offered in the second question, otherwise a lower bid is offered. This is known as a Double Bounded question.

4. Iterative Bidding

   Here the respondent is offered a sequence of dichotomous choice questions, increasing or decreasing in bid-value offered according as the response to the first question is YES or NO respectively. The process stops when the response changes or the list of bids is exhausted.

One other variation on any WTP question is a WTA question which aims to identify a respondent's willingness to accept compensation for removal of a service or for foregoing a right to use an amenity. The Arrow panel recommended the use of a WTP question and that this be of dichotomous choice format.

Clearly the binary responses to such a question require binary data techniques; in particular binary regression methods. Also the bids offered must be chosen. A distribution of bids across respondents is required; i.e. a binary regression design is needed.

6.5 Design of Single Bid Dichotomous Choice CV Studies

In a single bid dichotomous choice CV study a bid value, say x, must be chosen for a respondent. This is a design variable. An axiomatic assumption is that his/her response will be YES if $WTP \geq x$ and NO otherwise, where WTP is the respondent's true willingness to pay. In order to apply the design theory, we need to make assumptions about the distribution of WTP across the population of interest. Common assumptions in the CV literature have been that WTP or $lnWTP$ has a logistic or normal distribution. The logistic distribution is given a utility function theory justification [See Alberini (1995), Kanninen (1993) Nyquist (1992)]. Let $G(\cdot)$ denote the cumulative distribution function of WTP, so that $G(x) = \mathbf{Pr}(WTP \leq x)$. It is convenient to assume that

$$G(x) = G_0\left(\frac{h(x)-\mu}{\sigma}\right),$$

where $h(x)$ is an increasing function, μ and σ can be interpreted as a location and a scale parameter of $h(WTP)$, and $G_0(\cdot)$ is a standardised distribution.
It is natural to focus on modelling the probability of a YES response.

$$\begin{aligned}
\mathbf{Pr}(YES|BID = x) &= Pr(WTP \geq x) \\
&= 1 - G(x) \\
&= 1 - G_0\left(\frac{h(x)-\mu}{\sigma}\right) \\
&= F\left(\frac{\mu - h(x)}{\sigma}\right) \\
&= F(\alpha + \beta h(x)) \\
&= F(z)
\end{aligned}$$

where

$$\begin{aligned} z &= \alpha + \beta h(x) \\ \alpha &= \mu/\sigma, \qquad \beta = (-1/\sigma) \\ F(z) &= 1 - G_0(-z). \end{aligned}$$

Note that the function $F(z)$ satisfies the properties of a cumulative distribution function. Also β should be negative. Thus we have formulated a binary regression model, with the variable z representing a standardised design variable, like that of Chapter 2.

Thus our CV design problem, under the parameter dependent linear transformation $z = \alpha + \beta x$, can be transformed to a D-optimal or a c-optimal design problem for a weighted linear design problem with weight function

$$w(z) = \frac{f^2(z)}{F(z)[1 - F(z)]}.$$

Optimal designing here means choosing bids for respondents (possible values of WTP). A criterion needs to be chosen. D-optimality is a possibility if both parameters are of interest, in effect μ and σ. We focus on this. However estimation of mean WTP is usually of primary interest. We should want then to minimise the (asymptotic) variance of $\hat{\mu}$. This is an example of the c-optimal criterion.

The c-optimal criterion aims to minimise the asymptotic value of $\mathbb{V}(\underline{\boldsymbol{c}}^T \hat{\underline{\boldsymbol{\lambda}}})$ where $\underline{c}^T \underline{\lambda}$ is a known linear combination of the unknown parameters $\underline{\lambda} = (\alpha, \beta)^T$. Under the transformation $z = \alpha + \beta x$, this transforms to another c-optimal criterion. The geometrical characterisation of a c-optimal design, due to Elfving (1952) (see also Chernoff (1979)) is based on identifying the boundary of the convex hull of $G \cup \{-G\}$. The vector $\underline{c}$ extended if necessary will cut this at a point which is a convex combination of points in G and $\{-G\}$. These points are the support points and the convex weights are the optimal weights. Algebraic solutions for these weights are given in Kitsos, Titterington and Torsney (1988). In the case of two parameter models there must exist a design with 1 or 2 support points. Ford *et al.* (1992) derived c-optimal designs for all vectors $\underline{c}$ and for all choices of $\mathcal{Z} = [a, b]$, identifying, in particular, changes from one to two-point designs. It is clear from the convexity of the plots of G that there will be one-point designs for many choices of $\underline{c}$. Wu (1988) extended this work to percentile estimation as noted above. A particular conclusion is that for the choices of $F(\cdot)$ considered, the optimal design for estimating the median is a one-point design taking all observations at the (current provisional estimate) of the median. This transforms to $z = 0$ in the case of the normal and the logistic choices of $F(\cdot)$.

If the criterion is good estimation of the median by minimising the asymptotic variance of its estimate the design is to take all observations at the currently believed value of the median. If however we wish good estimation of both parameters of the model the D-optimal criterion could be optimised.

To completely define our problem we need to be clear about the design interval for z. Clearly WTP and hence x is positive. Hence a design interval for x must be positive. This in turn could impose restrictions on z, unless the function $h(x)$ of section (6.5) is unrestricted. For example $h(x) = ln(x)$. In this case distributions such as the standardised normal or logistic are feasible choices of $F(\cdot)$. This of course implies that WTP is log-normal or log-logistic. However in the early CV literature raw WTP has been assumed to be normal or logistic. This corresponds to $h(x) = x$.

Then $z = \alpha + \beta x \leq \alpha$, if $\beta < 0$. So a largest design interval for z is $\mathcal{Z} = (-\infty, b]$, $b = \alpha$. Possibly this should be further restricted to a finite interval $[a, b]$ a transformation of limits $\boldsymbol{c}$ and $\mathbf{d}$ on $\boldsymbol{x}$ (whatever is $h(x)$, where c is a minimum viable charge and d is a maximum politically acceptable price.

However this issue has been ignored in the CV literature on design. These have effectively assumed $\mathcal{Z} = (-\infty, \infty)$, in which case the standardised support points quoted may not transform back to positive WTP values. Kanninen (1993) for the logistic and Alberini (1995) for the normal report three (classes of) such designs namely: a D-optimal design, the design for minimising the asymptotic variance of the estimate of the median (c-optimality) and designs (which depend on the sample size) for minimising the width of a fiducial interval estimate of the median. The c-optimal design is the one point design placing all weight at $z = 0$ and hence at the (currently known) median WTP-value. The others are symmetric designs in z placing equal weight at values $\pm z^*$, where z^* maximises the relevant criterion over such symmetric designs. We have already reported such values for the D-criterion in Chapter 2.
These same authors went on to consider Double Bounded CV studies to which we now turn.

6.6 On Design of Double Bounded CV Studies

Recall that a Double Bounded CV study presents each respondent with two bids, the second being higher or lower than the first according as the response to the first bid is Yes or No. Kanninen (1993) and Alberini (1995) report constrained optimal designs for these bids under which the first bid is set equal to the currently known median i.e $z = 0$ and the second bid is $+z^*$ or $-z^*$ according as the answer to the first bid is Yes or No, z^* being chosen to optimise the relevant criterion. They report the values of z^* for the criteria for which they reported single bid designs, and for the same distribution. A crucial further assumption was that this distribution was assumed to be the same at both bids. To distinguish this approach from the following we call it the Univariate Approach.

Alberini (1995) relax this assumption. They consider the notion that a respondent has two WTP-values, WTP_1 and WTP_2 at the two bids respectively. Hence we call this the Bivariate Approach. There would be a justification for this if there was a time lag between offering the two bids thereby allowing for a change in opinions. Alternatively some argue that the respondent may react to the first bid, resulting in a revision of their opinions. The authors assume a bivariate normal distribution for $(lnWTP_1, lnWTP_2)$ with a common mean μ, a common standard deviation σ and a correlation ρ. Thus $\exp(\boldsymbol{\mu})$ is median WTP. Let x_1 and x_2 be the two bids to be offered. Then standardised design variables (w.r.t. μ, σ) are $z_i = \frac{[(\ln x_i) - \mu]}{\sigma}$. For fixed or known ρ then, this is a two-parameter model. Alberini (1995) determine constrained c-optimal designs under which $z_1 = 0$ for all respondents and for YES responses to this bid

$$z_2 = \begin{cases} +z^* & \text{with probability } \lambda^* \\ -z^* & \text{with probability } 1 - \lambda^*, \end{cases}$$

while for NO responses

$$z_2 = \begin{cases} -z^* & \text{with probability } 1 - \lambda^* \\ +z^* & \text{with probability } \lambda^*. \end{cases}$$

with z^* and λ^* being chosen optimally. Their values depend on ρ. Note that an implication is that some of those who respond YES to the first bid may be offered a lower bid and vice versa.

Table 6.1 shows that for low ρ, the second bids are extremely close to the median. If the two WTP variables are uncorrelated, the design problem reduces to that of finding the optimal design for the median of each of the two single-bounded

Table 6.1: Optimal Variance Minimizing Designs for the Bivariate Probit Model[a]. Alberini (1995)

ρ	0.1	0.2	0.4	0.5	0.7	0.9	0.95	0.9999
z^*	0.0054	0.0230	0.1038	0.1812	0. 4956	0.9529	0.9803	0.9816
λ^*	0.5319	0.5641	0.6310	0.6666	0.7468	0.8564	0.8989	0.9955

[a]The first bid value is always $c = exp(\mu)$; the second bid value are $c^{UP} = exp(z^*_{yes}\sigma + \mu)$ and $c^{DN} = exp(-z^*_{yes}\sigma + \mu)$ with probability λ^* or $(1 - \lambda^*)$ depending on the answer to the first WTP question.

models associated with the two payment questions. The single-bounded model would be applied in this situation. As ρ increases to one, the design tends to the double-bounded variance-minimizing design : virtually all of those who answered "yes" are offered a bid value equal to $exp(0.9816\sigma + \mu)$, and all of those who answered "no" to the first question are offered a bid value equal to $exp(-0.9816\sigma + \mu)$. [1]

6.7 Designs for Second Bids

The rationale of the approach we now advocate is that a design for the second bid of a double bounded CV question should, wherever possible, be based on the conditional distribution of WTP given the response, YES or NO, at the first bid; that is the c.d.f. $F(\cdot)$ and p.d.f. $f(\cdot)$ above should take the relevant (standardised) conditional forms. We consider both the Univariate and Bivariate approaches.

[1]The Optimal design suggested in this section defines four groups of respondents, which are described by ("yes" to the first question, second bid lower than the first), ("yes" to the first question, higher second bid), ("no" to the first question, lower second bid), and ("no" to the first question, higher second bid). Clearly, if the correlation between the first and the second WTP values is high but the sample size is small, the second and the forth of those groups may be composed of very few respondents (or none at all), and the emprical frequency of one type of response in those two groups may be one. In that case the MLE estimator may not be defined. These problems should be absent if the sample size is sufficiently large.

6.7.1 Univariate Approach

Suppose that the first bid is x. Then our standardised design variable is $z = \alpha + \beta h(x)$. Let $c = z$. Suppose that the answer to the first bid is NO, so that $WTP < x$. Hence the random variable $U = \alpha + \beta h(WTP) > c \quad if \quad \beta < 0$. A relevant standardised conditional distribution function is

$$\begin{aligned} P(U \leq z | U \geq c) &= \frac{[F(z) - F(c)]}{[1 - F(c)]}, \quad z \geq c \\ &= F(z | z \geq c) \ \text{ say,} \end{aligned}$$

since $F(\cdot)$ is the marginal c.d.f. of U. The corresponding conditional p.d.f. is

$$f(z|z \geq c) = \frac{f(z)}{[1 - F(c)]}, \quad z \geq c$$

Thus the design problem for the second bid transforms to a weighted linear regression design problem with weight function

$$\begin{aligned} w(z) &= \frac{f^2(z|z \geq c)}{F(z|z \geq c)[1 - F(z|z \geq c)} \\ &= \frac{f^2(z)}{[F(z) - F(c)][1 - F(z)]}, \quad z \geq c \end{aligned} \tag{6.1}$$

If the response to the first bid is YES the corresponding c.d.f, p.d.f. and weight function are

$$\begin{aligned} F(z|z \leq c) &= \frac{F(z)}{F(c)} \\ f(z|z \leq c) &= \frac{f(z)}{F(c)} \\ w(z) &= \frac{f^2(z)}{F(z)[F(c) - F(z)]} \end{aligned} \tag{6.2}$$

We note that both equations (6.1) and (6.2) are unbounded at $z = c$. For the case of equation (6.1) see plots of $\sqrt{w(z)}$ in Figures (6.2) (6.3) (6.4) (6.5) (b) (e) (h) (k). $z = c$ is not a permissable member of the design interval.

6.7.2 Bivariate Approach

Here we have the concept of potentially different WTP values, WTP_1, WTP_2 at the two bids. We consider the following general scenario under which for some increasing function $h(\cdot)$, the pair $h(WTP_1)$, $h(WTP_2)$ have a joint distribution indexed by a common location parameter μ, a common scale parameter σ, and a third parameter, say ρ, measuring correlation or possibly some other form of dependence. Further we assume that the standardised variables $Z_i = (h(WTP_i) - \mu)/\sigma$, $i = 1, 2$ have a joint distribution indexed only by ρ.

In these terms the relevant standardised distributions for the second bid given the response YES or NO, at the first bid are those of Z_2 conditional on $Z_1 \geq c$ or $Z_1 \leq c$, where c represents a standardised initial bid.

Let $X = Z_1$, $Z = Z_2$. Assume that these have joint c.d.f. $F_{xz}(x, z)$, marginal c.d.f.'s $F(x)$, $F(z)$ and respective marginal p.d.f's $f(x)$, $f(z)$.

We want to determine the conditional distribution of Z given $X \geq c$ or $X \leq c$. The respective c.d.f.s, p.d.f.s and weight functions are:

$$
\begin{aligned}
F(z|z \geq c) &= \frac{[F(z) - F_{xz}(c, z)]}{[1 - F_x(c)]}, \\
f(z|z \geq c) &= \frac{[f(z) - \partial F_{xz}(c, z)/\partial z]}{[1 - F_x(c)]}, \\
w(z) &= \frac{[f(z) - \partial F_{xz}(c, z)/\partial z]^2}{[F(z) - F_{xz}(c, z)][1 - F_x(c) - F(z) + F_{xz}(c, z)]}, \\
& \qquad -\infty < z < \infty
\end{aligned}
\tag{6.3}
$$

and

$$
\begin{aligned}
F(z|z \leq c) &= \frac{[F_{xz}(x, z)]}{F_x(c)}, \\
f(z|z \leq c) &= \frac{\partial F_{xz}(c, z)/\partial z}{F_x(c)}, \\
w(z) &= \frac{[\partial F_{xz}(c, z)/\partial z]^2}{F_{xz}(c, z)[1 - F_{xz}(c, z)]}, \quad -\infty < z < \infty
\end{aligned}
\tag{6.4}
$$

A clear added dimension to weight functions at Equations (6.3) and (6.4) are that they require calculation of a joint c.d.f. and one of its first partial derivatives. By current standards these are not available 'explicitly' for the standard bearer of joint distributions -bivariate normality although there are published programs for calculating the joint c.d.f.. This was the distribution used by Alberini (1995) but they did not need to calculate these terms. As to the logistic there is no standard choice of bivariate extension.Various classes have been proposed, including one based on copulas. This is a joint c.d.f. defined as follows:

$$F_{xz}(x,z) = H\{F_x(x), F_z(z)\},$$

where $H(u,v)$ is a joint c.d.f on $[0,1]^2$ with uniform marginals, i.e. $H(u,1) = u$, $H(1,v) = v$. The function $H(u,v)$ is known as a copula. It is a tool for generating joint distributions with given marginals; see Hutchinson and Lai (1991) Chapter 10. One example is Plackett's distribution for which

$$H = H(u,v) = \frac{[1 + (\psi - 1)(u+v)] - \sqrt{[1 + (\psi - 1)(u+v)]^2 - 4\psi(\psi-1)uv}}{2(\psi - 1)}.$$

The parameter ψ is a constant global cross ratio since

$$\psi = \frac{H(1-u-v+H)}{[(u-H)(v-H)]}.$$

It is a measure of dependence, taking the value 1 when the underlying uniform random variables are independent. In the results we report below we adopt this particular copula and assume $F_x(z) = F_z(z) = F(z) = exp^z/(1 + exp^z)$; i.e. a common logistic marginal. This results in the following simplifications of (6.3) and (6.4) respectively

$$w(z) = \frac{f^2(z)[1 - \partial H(u,v)/\partial v]^2}{[F(z) - H(u,v)][1 - F(c) - F(z) + H(u,v)]}, \quad (6.5)$$

$$w(z) = \frac{f^2(z)[\partial H(u,v)/\partial v]^2}{H(u,v)[1 - H(u,v)]}, \quad (6.6)$$

where $u = F(c)$, $v = F(z)$. For the logistic $F'(z) = F(z)[1 - F(z)]$. See Figures (6.1) (b) and (d) for the case (6.3) and Figures (6.1) (f) and (h) for the case (6.4).

6.7.3 Result for the Bivariate Approach

It is again illuminating to study plots of the set G; that is of $\underline{g}(z) = \sqrt{w(z)}\ (1, z)^T$ for the appropriate set of z-values. We consider the Bivariate Approach first. Figure (6.1) (a) and (c), depicts a plot of G for case (6.5) of $w(z)$ (see Figure (6.1) (b) and (d) for the case of (6.6)) with $F(z)$ the logistic $\psi = 1.6$ $c = -5$ and $-\infty < z < \infty$. This is typical of the other values of c and ψ. The shape is similar to that of Figure (2.1) for unconditional binary weight functions. Namely it appears to be a **closed convex curve** in R^2 for the widest choices of $\mathcal{Z}$. In terms of Silvey's minimal ellipsoid argument we have the same conclusion. The minimal central ellipsoid containing G can only touch it twice in which case the D-optimal design has two support points. This is indeed the case, the support points being -1.544, 1.558. The same conclusion is reasonable for the section of G corresponding to the interval $a \leq z \leq b$ i.e on the design interval $[a, b]$ for z. Moreover the solution should be the same as that of the conjecture of Chapter 2. [2].

The structure of c-optimal designs should also be similar to those derived in Ford et al.(1992) for arbitrary design intervals $[a, b]$. These are either one point designs or two point designs which may comprise both endpoints or include only one of them or neither according to rules similar to those for D-optimality. In this case the values a^{**}, b^{**} are the support points on $(-\infty, \infty)$ if, for the vector $\underline{c}$ defining the c-optimal criterion, two points are needed. They are independent of $\underline{c}$. It is likely that for estimating the median, which should correspond to $\underline{c} = (1, 0)$ if $z = 0$ is a standardised median, the optimal design will be the one point design

[2]Let $a^* = -1.544$, $b^* = 1.55$ so that these are the support points on $[a, b] = (-\infty, \infty)$. They are therefore also the support points on $[a, b]$, where $a \leq a^*$ and $b \geq b^*$. Consider $a \geq a^*$ and let $b^*(a)$ denote the value of z which maximises the D-optimal criterion over two-point designs with support points a and z subject to $z > a$. The points a and $b^*(a)$ should be the support points on the design interval $[a, \infty)$ and hence on $[a, b]$ if $b \geq b^*(a)$. Similarly consider $b \leq b^*$ and let $a^*(b)$ denote the values of z which maximises the D-optimal criterion over two point designs with support points z and b, subject to $z < b$. The points $a^*(b)$ and b should be the support points on the design interval $(-\infty, b]$ and hence on $[a, b]$ if $a \leq a^*(b)$. Otherwise the support points should be the endpoints a and b. In particular this should be the case if $a \geq a^*$ and $b \leq b^*$

taking observations at $z = 0$.

6.7.4 Result for the Univariate Approach

We turn now to the Univariate Approach which we have studied more extensively in respect of D-optimality. Plots of the $\sqrt{w(z)}$ of Equation (6.1), and of $z\sqrt{w(z)}$ and the corresponding G are shown in Figure (6.2) , Figure (6.3), Figure (6.4), and Figure (6.5), with $F(z)$ the logistic for a range of values of the standardised initial bid c. These illustrate various points:

1. First G is no longer bounded, (since $w(z)$ is not bounded : $w(z) = \infty$ when $z = c$) at least in the first component of $\underline{g}(z)$ as z approaches c from above since $w(z)$ is infinite at c. Thus we must impose an arbitrary, lower bound a on z satisfying $a > c$ and **'cut away'** that part of G corresponding to $c < z < a$. We focus on D-optimal designs on $[a, b]$ for $b = \infty$.

2. Second the shape of G changes with c.

i. $c > 0$

For positive c, $\underline{g}(z) \to (\infty, \infty)^T$ as $z \to c$ from above. In general G has the shape of an increasing curve. [see Figures (6.4) and (6.5) (a), (d), (g), (h).]

ii. $c = 0$

For $c = 0$ it rises to a maximal turning point, and thereafter $g_2(z) = z\sqrt{w(z)} \to 0$ as $z \to c = 0$ from above. [see Figure (6.3) (j).]

iii. $c < 0$

For negative c, $g_1(z) \to +\infty$, $g_2(z) \to -\infty$ as $z \to c$ from above. For large enough negative c however G initially begins to exhibit something of the 'closed' convex shapes seen above, before **'turning'** to proceed to the above limits thereby forming what we call a **'tail'**. [see Figures

(6.4) and (6.5)] Denote by c^* the critical value of c for which this first happens.

We considered the case where the answer to the first bid is NO. For that reason we look at plots of the weight function of Equation (6.1) : For each value of c, this weight function is different from those of Chapter 2 which were unimodal. But here there can be either no **turning point** or two. From Figures (6.2) (6.3) (6.4) (6.5) (b), (e), (h), (k) it appears that $w(z)$ is decreasing for large c, but below some critical value of c^* it possesses two TP's.

Critical value of c^*.

Denote by $z^\dagger$ the value of z at which the maximal turning point occurs. We note that there must therefore be a value of $a < z^\dagger$ such that $w(a) = w(z^\dagger)$. The critical value c^* is the value of c such that $F(c) = 1/9$ as we now show.

Since $f(z) = F(z)[1 - F(z)]$ for the logistic then

$$\begin{aligned} w(z) &= \frac{F^2[1-F]^2}{\{[F-F(c)][1-F]\}}, \qquad F = F(z) > F(c) \\ &= \frac{F^2[1-F]}{[F-F(c)]} \end{aligned}$$

Hence $\ell(F) = \ln(w(z)) = 2\ln(F) + \ln(1-F) - \ln(F - F(c))$. $w(z)$ has TP's if $l(F)$ has TP's. The solution to $\ell'(F) = 0$ are solutions to a quadratic equation in F, if they exist. The discriminant of this quadratic function is in turn a quadratic function in $F(c)$, which is positive only for $0 < F(c) < 1/9$. Calculation of the critical c value is summurized in Appendix C.

6.7.5 Conjecture for the support points.

CASE 1 : We conjecture that for $c > c^*$, the D-optimal designs on $[a, \infty)$ have, for all a, two support points a and $b^*(a)$ as defined above. This will be the

D-optimal design on $[a, b]$ for $b > b^*(a)$. For $b < b^*(a)$ we conjecture the support points to be a and b.

Support points $Supp(p^*)$	**Design Interval**
$\{a, b^*(a)\}$	$\mathcal{Z} = [a, b], \quad a > c, \quad b = \infty.$
$\{a, b^*(a)\}$	$\mathcal{Z} = [a, b], \quad a > c, \quad b > b^*(a).$
$\{a, b\}$	$\mathcal{Z} = [a, b], \quad a > c, \quad b < b^*(a).$

CASE 2 : For $c < c^*$ the situation is more complicated. Sometimes there are two support points which may or may not include a and sometimes there are three support points including a. For the design interval $[a, \infty)$, we believe that the solution can be summarised in terms three critical values of a, say $a(L)$, $a(M)$, $a(U)$ as follows.

For $a \leq a(L)$ and $a \geq a(U)$ there are two support points a and $b^*(a)$. For $a(L) \leq a \leq a(M)$ there are three support points including a. The other two points can be found by maximising the D-optimal criterion subject to a being a support and using the explicit formulae for the three weights stated Chapter 2, section (2.2.5). For $a(M) \leq a \leq a(U)$ there is a fixed two-point support consisting of $a(U)$ and $b^*\{a(U)\}$. For $c = -5$ values are $(a(L)\ a(M)\ a(U))$ = (-4.73, -4.60, -1.586).

Range of Value	**Support points** $Supp(p^*)$	**Design Interval**
$-c < a < a(L)$	$\{a, b^*(a)\}$	$\mathcal{Z} = [a, b],\ a > c,\ b = \infty$
$a(L) < a < a(M)$	$\{a, \quad z_1^*(a), \quad z_1^*(a)\}$	$\mathcal{Z} = [a, b],\ a > c,\ b = \infty$
$a(M) < a < a(U)$	$\{a_U, \quad b^*(a(U))\}$	$\mathcal{Z} = [a, b],\ a > c,\ b = \infty$
$a > a(U)$	$\{a, \quad b^*(a)\}$	$\mathcal{Z} = [a, b],\ a > c,\ b = \infty$

6.7.6 Geometrical explanation of the optimal design

A rationale for the above solution can be found by considering Silvey's minimal ellipsoid argument.

- First we study G for all $z \geq c$. We denote this G by G_c. Each point in G_c is defined by a unique value of z. One can describe G_c as a locus which starts at the origin and follows an almost closed convex smoothly changing path, which, for large negative **c** will almost come back to the origin but at some point it turns away from the origin developing a 'tail' in convergence to $(\infty, -\infty)$ as $z \longrightarrow c$ from above. One exception to this is the case $c = -\infty$ when $G_{-\infty}$ will come back to the origin as $z \longrightarrow c$. In fact this is the G of Figure (6.2) a,d,g,j for ordinary logistic regression. For large negative finite c the 'almost closed convex ' part of G_c must be closely approximated by this logistic regression case.

- Now consider the case $\mathcal{Z} = [a, \infty)$, $G = \{\underline{g}(z) \in G_c : z \geq a\}$. For sufficiently large values of a, the tail of G extends well out towards $(\infty, -\infty)$. So that $\underline{g}(a)$ is to the right of and below $\underline{g}(z^\dagger)$. Intuitively the minimal ellipsoid touches G at $\underline{g}(a)$ and at one other point above and to the left of $\underline{g}(z^\dagger)$, that is at a point corresponding to a value of z say $b^*(a)$, above $z^\dagger$.

  Now think of a increasing so that we are 'cutting away' more of G. It seems plausible that the above solution remains valid at least until the value of a such that $w(a) = w(z^\dagger)$. Thereafter from some value $a(L)$ onwards there is clearly the potential in the case $b = \infty$ for the minimal ellipsoid to touch at $\underline{g}(a)$ and at two points corresponding to values of **z** on either side of $z^\dagger$.

  A justification for the fixed two-point design over $a(M) \leq a \leq a(U)$ can be drawn from the approximation, noted above, between $G_{-\infty}$ and the almost closed convex part of G_c for large negative c. The optimal design for this part of G_c and for $G_{-\infty}$ must be approximately the same. This would mean that the D-optimal design for the 'almost closed convex' part of G_c has two non-extreme support points say a^*,b^* so that the minimal central ellipsoid containing this part of G_c, contains strictly within it that section of this 'almost closed convex' part of G_c corresponding to $z < a^*$. It will also contain part of the tail of G_c but only part of it since clearly the tail must cross any bounded set. The value of z at which this crossing occurs identifies the value of $a(M)$, while $a(U) = a^*$.

This value $a(M)$ like the value $a(L)$ is a value of a at which there is a change from two support points to three or vice versa. Strictly speaking for these two values there are two active support points (with equal weight therefore) but in addition there is a 'sleeping' support point with zero weight. In the case of $a = a(L)$ the active support points are $a(L)$ and $b^*\{a(L)\}$ and the sleeping point is a value say z_u above these. In the case of $a(M)$ the support points are $\mathbf{a}(\mathbf{U})$ and $b^*\{a(U)\}$ while $a(M)$ is the sleeping support point. As a increases from $a(L)$ to $a(M)$ the weight at a decreases from 1/2 to 0. A set of equations for identifying these values can be derived from the fact that if there is a three point design on $[a, \infty)$ with a as a support point then the variance function $w(z)(1, z)(M^*)^{-1}(1, z)^T$ (where M^* is the optimal design matrix) must have turning points at the two higher support points, say z_1, and z_2. The triplet $(a(L), b^*\{a(L)\}, z_u)$ must be the values of (a, z_1, z_2) which satisfy the two zero derivative equations plus the equation setting the (explicit formula for the) weight at $\mathbf{z_2}$ to zero. The triplet $(a(M), a(U), b^*\{a(U)\})$ instead satisfies the zero weight at $a(M)$.

These conjectures for $c < c^*$ cover the design interval $\mathcal{Z} = [\mathbf{a}, \infty)$ for all values of a. For the case of a finite $\mathcal{Z} = [a, b]$ we have limited comments. If b is greater than the upper support point of the design on $[a, \infty)$ then that design must also be optimal for $\mathcal{Z} = [a, b]$. Otherwise the design must differ.

It is fairly likely that b will be a support point (e.g. when $a > a^*(U)$) but this may not always be so. In general there will be designs with 2 or 3 support points which may include both a and b as support points or only one of them.

Remark 6.1. *Plots of the function $H(z)$ for various values of c are given in Figure (6.7). These depict different shapes according to the value of c.*

- $c > c^*$: *We have shapes similar to those in chapter 2, section (2.5.1). $H(z)$ is convex over (c, ∞) and convex increasing if $c > 0$. See plots in Figures (6.7) (d), (e), (f), (g), (h), (i), (j), (k), (l). An upward sloping line with a negative intercept can cross $H(z)$ at most twice in (c, ∞).*
- $c < c^*$: *$H(z)$ has a different shape now; a reflection of the weight function, $w(z)$. $H(z)$ is convex up to the some point concave increasing and again convex increasing.*

Remark 6.2. *So in the case $c > c^*$ the equivalence theorem is satisfied by our conjectured optimal designs for all possible design interval [a,b] if the function $h(z)$ is increasing over $z \leq z_{max}$ and over $z \geq z_{max}$. This seems to be the case from the plots of $h(z)$ in Figure (6.6). In fact $h(z)$ seems to be increasing for all c. So the best two-point D-optimal design on $[a, b]$, $a > c$ is possibly given by the conjecture if $c < c^*$.*

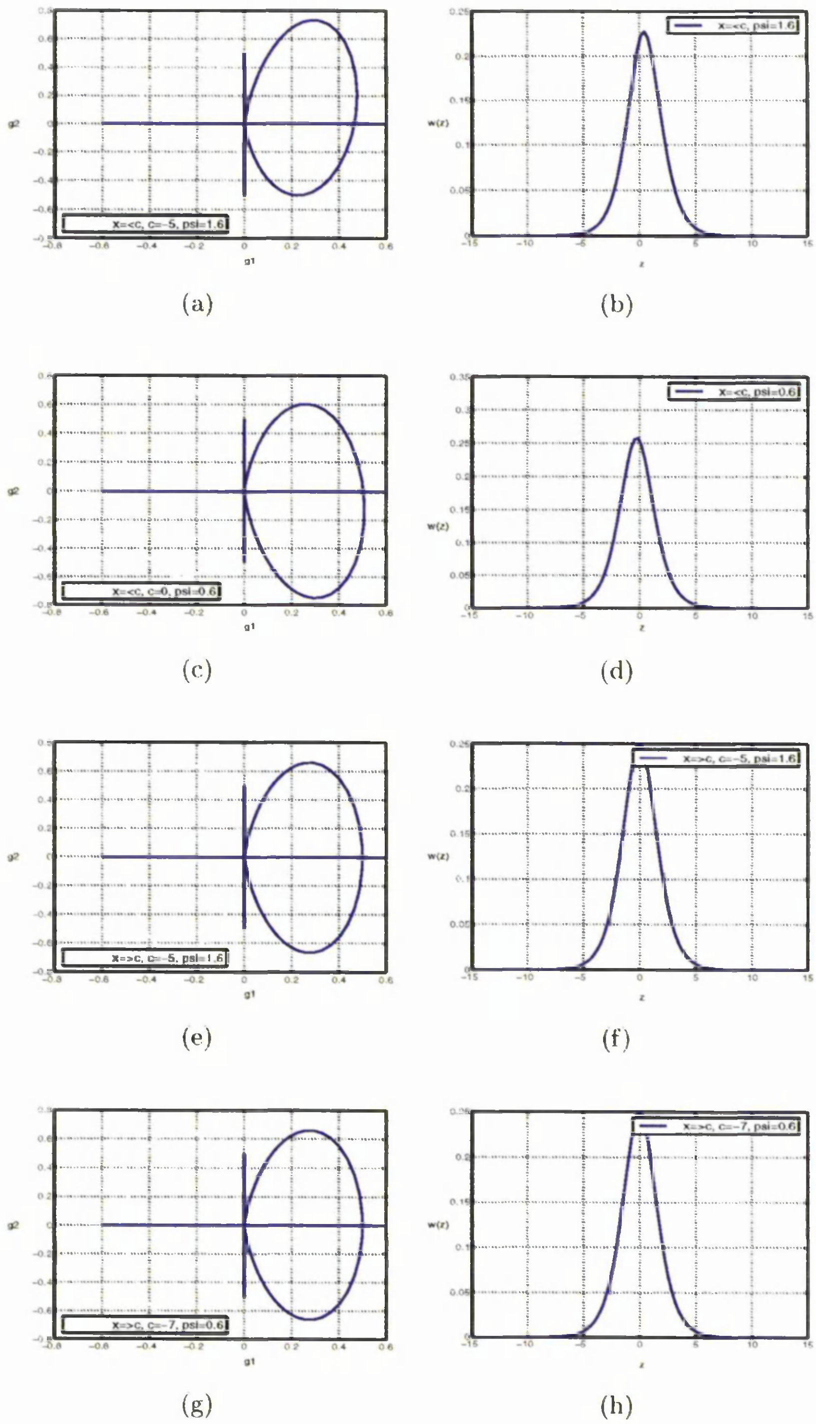


Figure 6.1: Plots of G and $w(\cdot)$ for Bivariate Approach : Plackett's Distribution with Logistic Marginals for various "c" and "ψ".

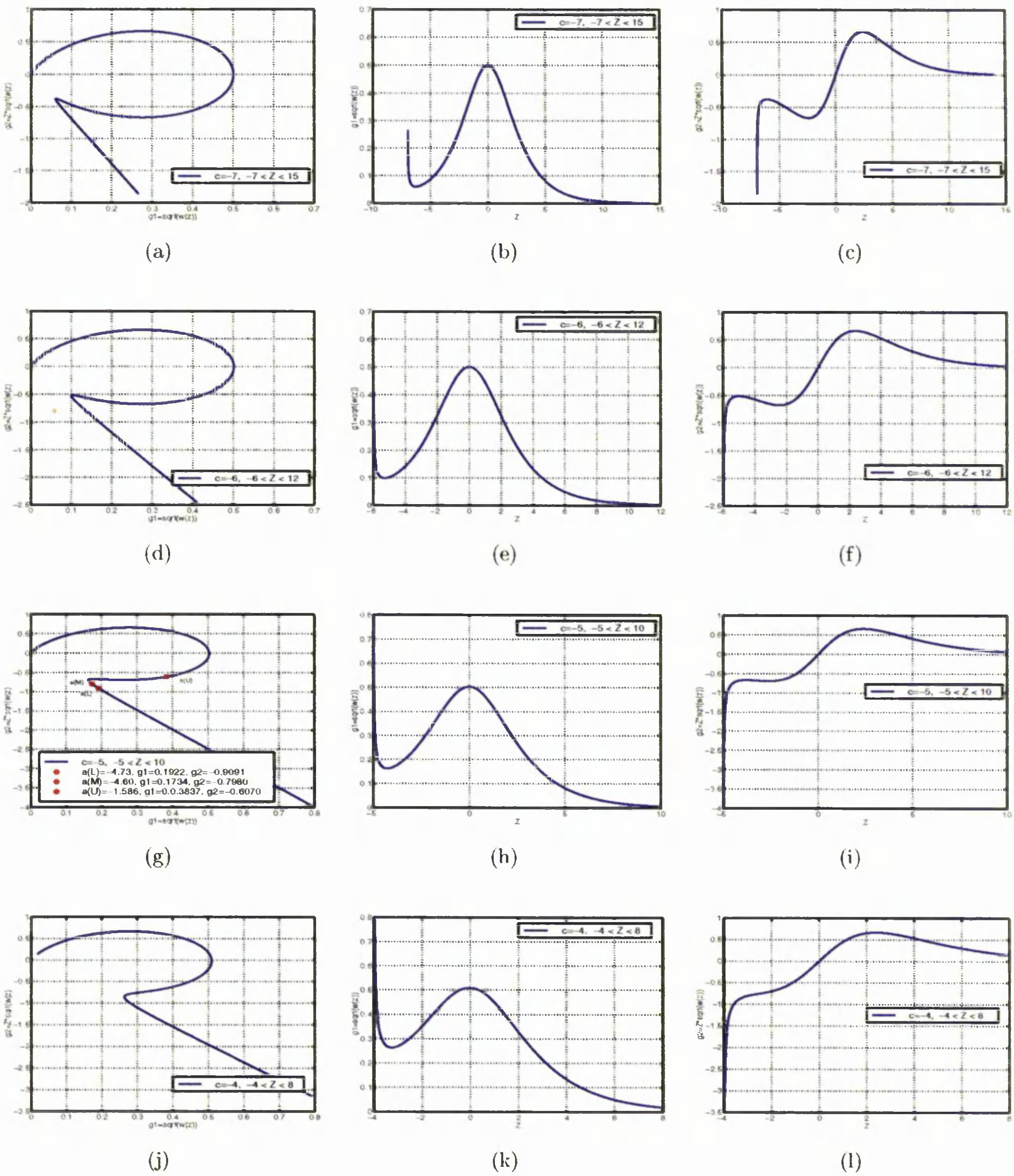


Figure 6.2: Plots of G, $g_1 = \sqrt{w(z)}$ and $g_2 = z\sqrt{w(z)}$ for Univariate Approach Logistic Function for various "c" values; $z \geq c$, $c = -7, -6, -5, -4$.

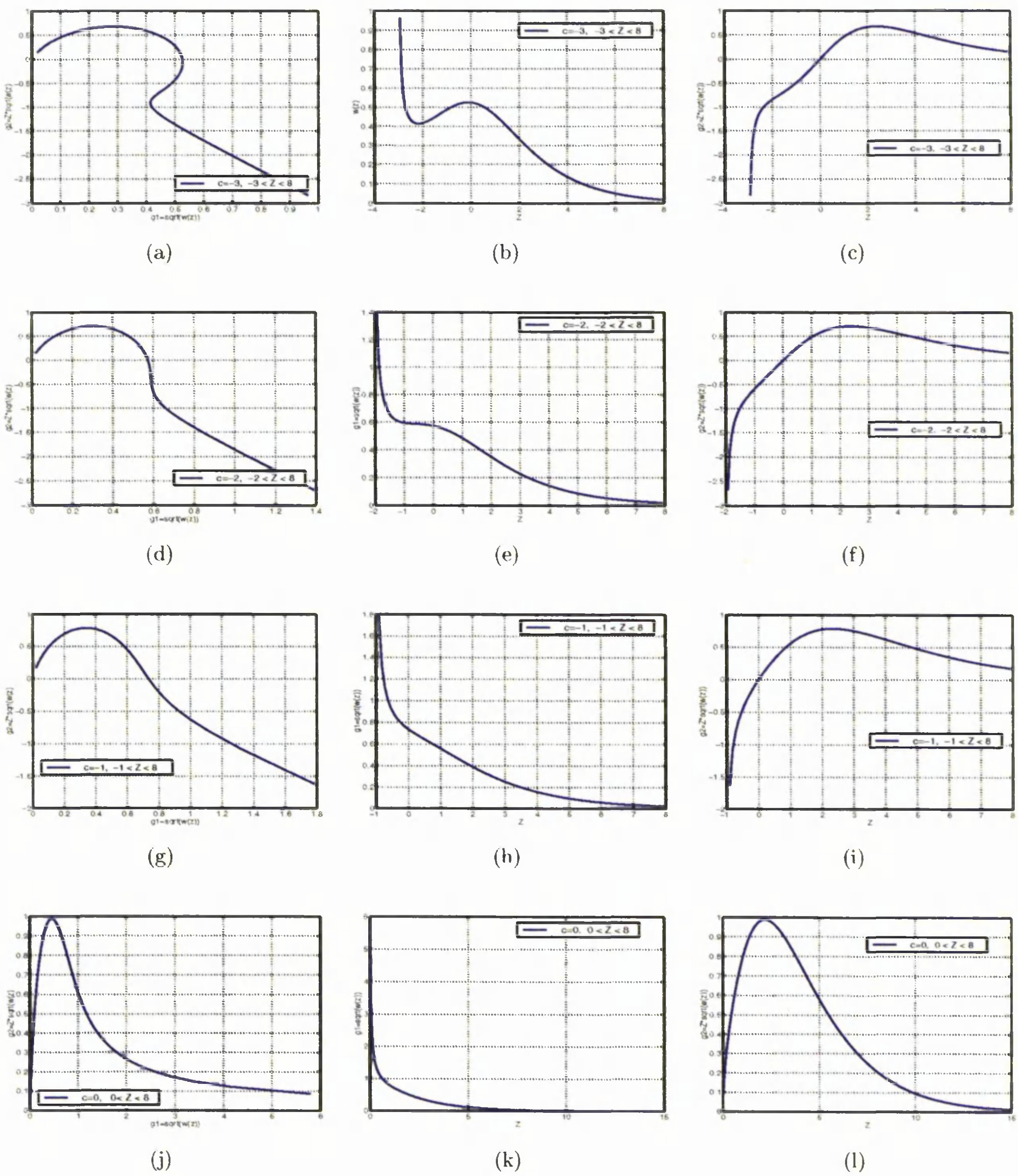


Figure 6.3: Plots of G, $g_1 = \sqrt{w(z)}$ and $g_2 = z\sqrt{w(z)}$ for Univariate Approach Logistic Function for various "c" values; $z \geq c$ $c = -3, -2, -1, 0$.

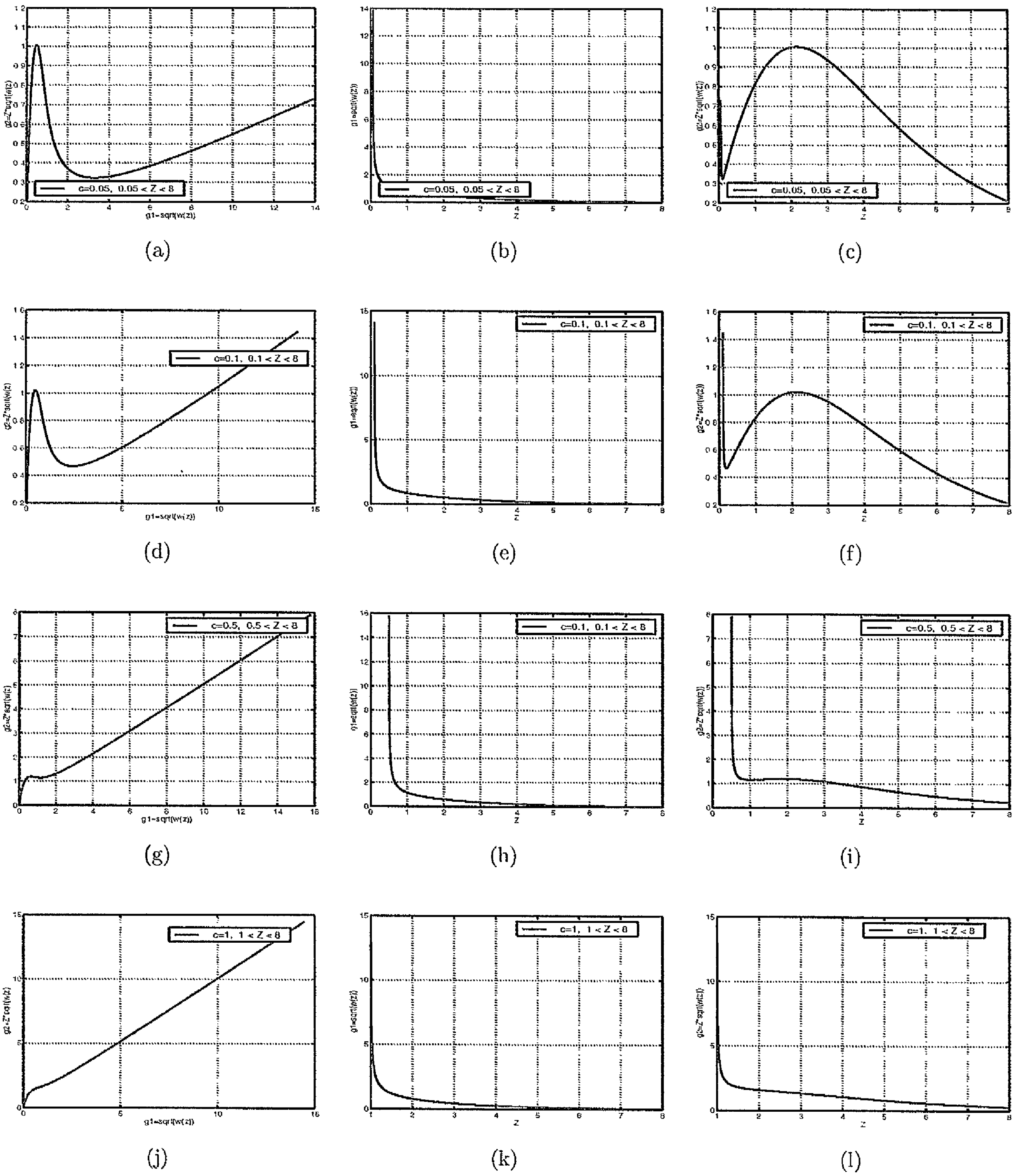


Figure 6.4: Plots of G, $g_1 = \sqrt{w(z)}$ and $g_2 = z\sqrt{w(z)}$ for Univariate Approach Logistic Function for various "c" values; $z \geq c$, $c = 0.05, 0.01, 0.1, 1$.

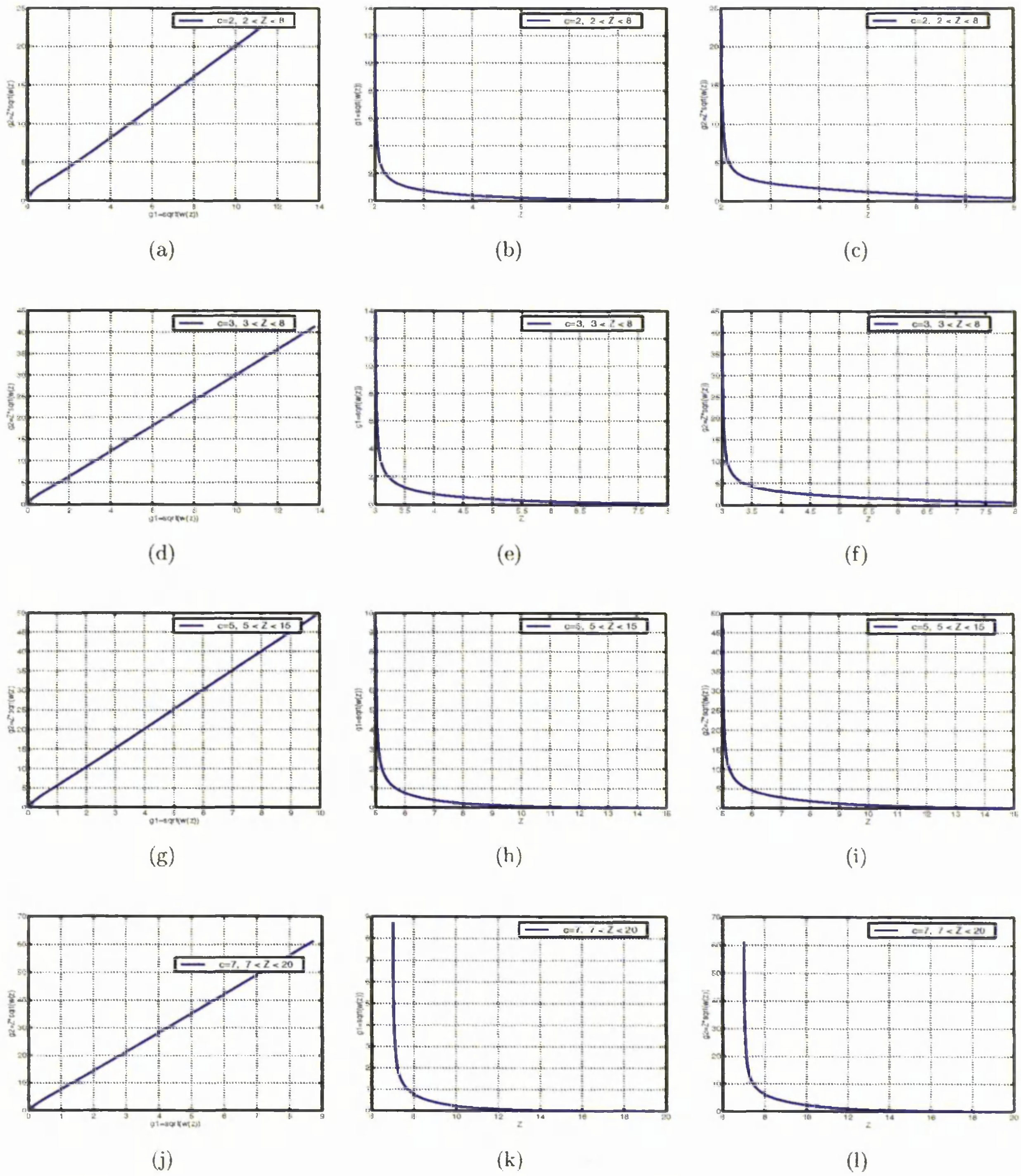


Figure 6.5: Plots of G, $g_1 = \sqrt{w(z)}$ and $g_2 = z\sqrt{w(z)}$ for Univariate Approach Logistic Function for various "c" values; $z \geq c$, $c = 2, 3, 5, 7$.

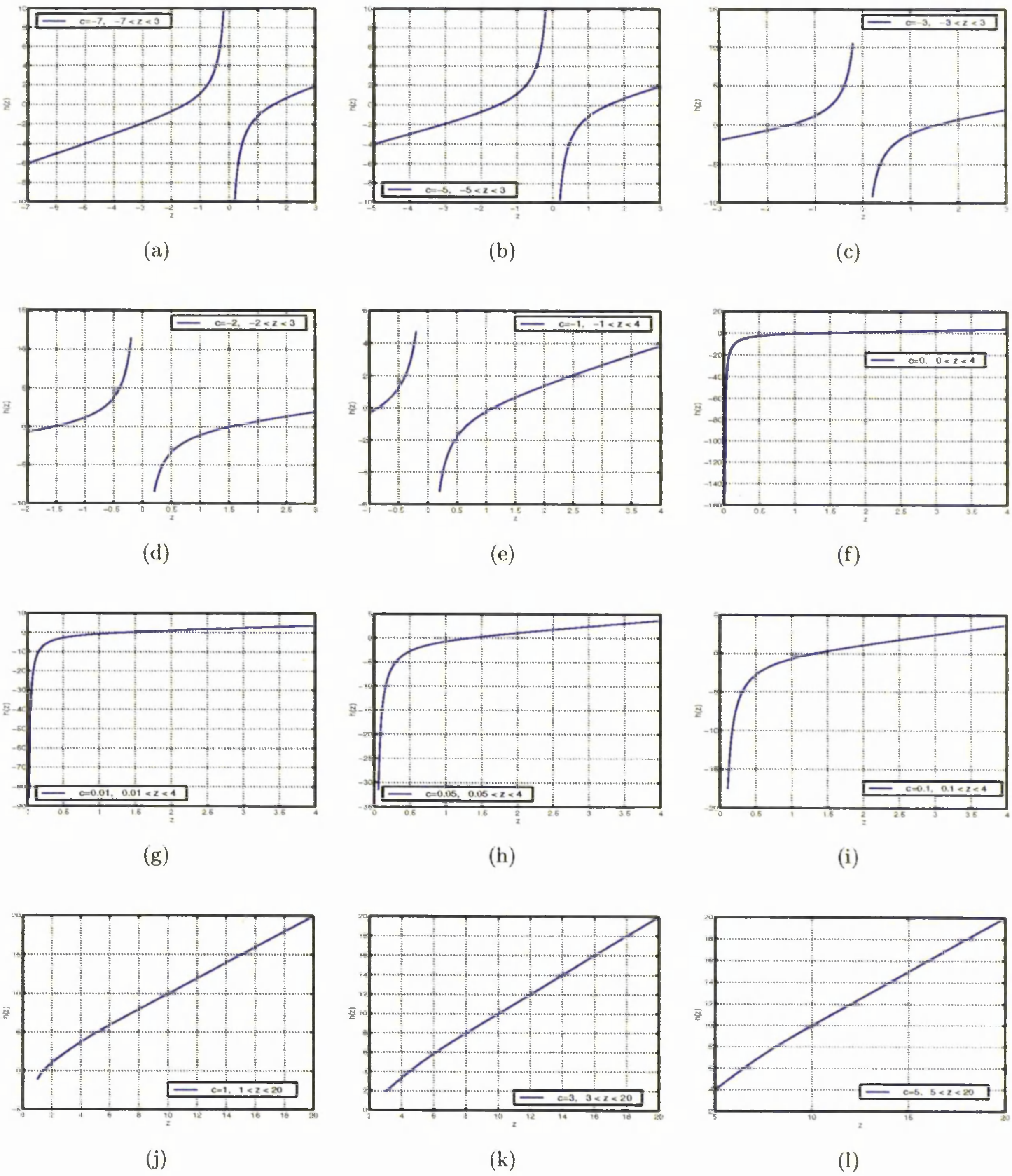


Figure 6.6: Plots of $h(z)$ for Univariate Approach Logistic Binary Weight Function, $c = -7, -5, -3, -2, -1, 0, 0.01, 0.05, 0.1, 1, 3, 5$.

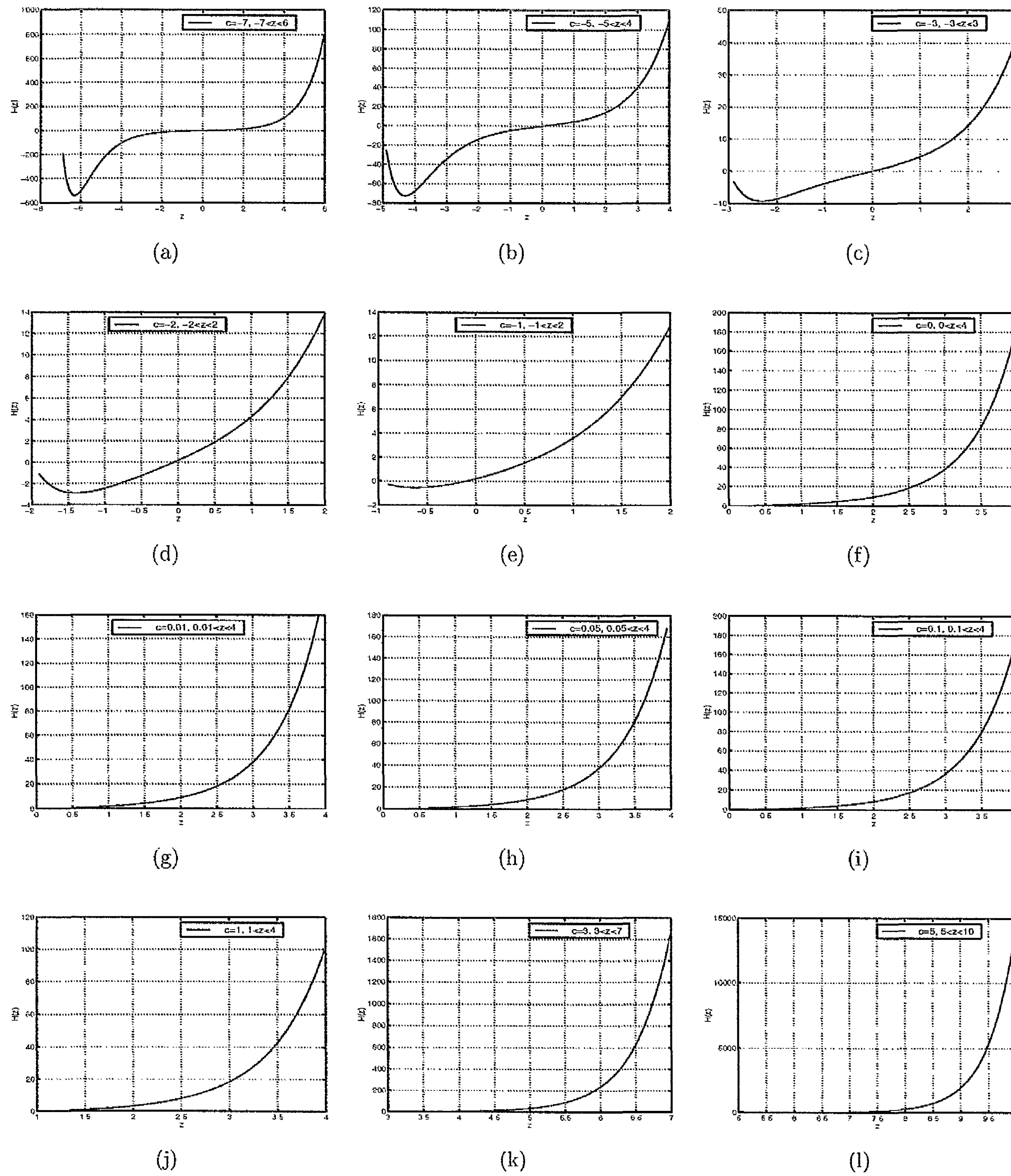


Figure 6.7: Plots of $H(z)$ for Univariate Approach Logistic Binary Weight Function, $c = -7, -5, -3, -2, -1, 0, 0.01, 0.05, 0.1, 1, 3, 5$.

Chapter 7

Conclusion

7.1 Discussion of results

We have derived locally D-optimal designs for various binary regression problems. The results discussed so far show that for the binary regression model, locally D-optimal designs can be sensible designs if provisional information on the true values of the parameters is available from pilot studies.

Our exploration of D-optimality made extensive use of weighted linear regression, and this led us to exploiting and applying well established results on optimal designs in linear models.

First of all, we considered D-optimal designs for binary response models with one design variable, and we transformed the design problem to one for a weighted linear regression model, the weight function being :

$$w(z) = \frac{f^2(z)}{F(z)[1 - F(z)]}$$

where $f(z) = F'(z)$ is the density of $F(z)$, and the design interval being $Z = [a, b]$. We also considered various other (non-binary) weight functions.

We established that for many weight functions the optimal design is a two point design. The support points for the z values are :

$$\begin{aligned} Supp(\xi^*) &= \{a^*, b^*\} \qquad a < a^*, b > b^* \\ Supp(\xi^*) &= \{max\{a,\ a^*(b)\}, b\} \qquad a < a^*, b < b^* \\ Supp(\xi^*) &= \{a, min\{b,\ b^*(a)\}\} \qquad a > a^*, b > b^* \\ Supp(\xi^*) &= \{a, b\} \qquad a > a^*, b < b^* \end{aligned}$$

where a^*, b^*, $a^*(b)$, $b^*(a)$ maximise the determinant over relevant intervals.

These results follow if

(i) the function $H(z) = \dfrac{-w'(z)}{[w(z)]^2}$ is first concave increasing then convex increasing,

(ii) the function $h(z) = z + \dfrac{2w(z)}{w'(z)}$ is increasing (this also guarantees that $G(\mathcal{Z}_w)$ is closed convex). In some cases the ratio $w(z)/w'(z)$ is also increasing. (Note: $G(\mathcal{Z})$ shows that induced design space and $\mathcal{Z}_w$ the widest possible design space.

Secondly, we studied the more general situation of multiple design variables. Multiparameter design problems also tranformed to weighted regression design problems in design variables $z_1, \cdots, z_l$ with rectangular design spaces. Such that $z_1 \in [a, b]$ and $-1 < z_j < 1$, $j = 2, \cdots, l$. For many of our weight functions optimal designs consist of taking observations at two values of z_1, the two values satisfying the above conjecture. One such design consists of dividing the total weight at each of these values equally across all combinations of $z_j = \pm 1$, $j = 2, \cdots, l$).

We also considered some bounded design spaces. We found optimal designs for 2 design variables z_1 z_2 as above when their design space is a polygon. Some of the above results extended to this. In particular for many weight functions at most two observations can be taken along any edge. We note that Sitter and Fainaru (1997) considered the case $z_2 = z_1^2$.

Possible problems for future consideration are :

- to extend the work of chapter 4 on polygonal design spaces to higher dimensions,
- to establish necessary and sufficient conditions on $w(z)$ for guaranteeing the above conjecture.

Finally, results from optimal design theory have been used uncritically in the CV literature in respect of various design criteria. We made improvements. We also reported new results which focus on optimal designs for the second bid of a double bounded study given the response at the first bid. These will be useful when there is a time gap between offering the two bids.
We reviewed with some minor criticisms the use of optimal designs in the case of CV studies with dichotomous choice questions and have offered new designs for the 2nd question of a double bounded dichotomous choice question conditional on the response to the first bid.
There remains much to do including deriving designs, in respect of 'D-optimal' and 'median' oriented criteria, of the following kinds :

- designs for both stages of a double bounded CV study-univariate and bivariate cases.This will involve theory of designs for multivariate responses.
- Designs for the bivariate approach when the dependence parameter is treated as unknown. This is a consideration for both the univariate and bivariate approach.

- Designs for the bivariate approach to double bounded CV studies when different location and scale parameters are assumed at the two bids. This would seem a natural extension of the common location/scale case although the choice of criterion is possibly unclear. Good estimation of the parameters of the second bid may be of greater importance.

- Designs when finite limits are imposed on bids for both the double bounded and single bounded cases.

- Designs for optimal bids when other explanatory variables are included in a model for WTP.

References

Abdelbasit, K. and Plackett, R. (1983). Experimental design for binary data. *Journal of the American Statistical Association*, 78(381):90–98.

Alahmadi, A. M. (1993). *Algorithms for the Construction of Constrained and Unconstrained Optimal Designs.* PhD thesis, Department of Statistics, University of Glasgow.

Alberini, A. (1995). Optimal designs for discrete choice contingent valuation surveys: Single bound, double bound, and bivariate models. *Journal of Environmental Economics and Management*, 28:287–306.

Atkinson, A. and Donev, A. (1992). *Optimum Experimental Designs.* Oxford Statistical Science Series . 8. Oxford Science Publications.

Bishop, R. and Heberlein, T. (1979). Measuring values of extramarket goods:are indirect measures biased? *Amer.J.Agri.Econ.*, 61:927–930.

Box, G. and Lucas, H. (1959). Design of experiments in nonlinear situations. *Biometrika*, 46:77–90.

Chernoff, H. (1979). Sequential analysis and optimal design. Revised ed. Philadelphia : SIAM.

Donaldson, C. (1993). Theory and practice of willingness to pay for healt care. HERU Discussion Paper 01/93, Dept. Pub. Health, Univ of Aberdeen.

Elfving, G. (1952). Optimum allocation in linear regression theory. *Ann.Math.Statist.*, 23:255–262.

Farrell, R. H., Kiefer, J., and Walbran, A. (1967). Optimum multivariate designs. In Atkinson and Donev, 1992, page 132.

Fedorov, V. (1972). *Theory of Optimal Experiments - Originally Published in Russian under the title THEORIYA OPTIMAL'NOGO EKSPERIMENTA by izdatel'stvo Moskovskogo Universiteta,1969.* Academic Press, first edition.

Ford, I., Torsney, B., and Wu, C. (1992). The use of a canonical form in the construction of locally optimal designs for non-linear problems. *J.R. Statist. Soc. Ser*, 54(2):569–58.

Hanley, N. (1989). Valving rural recreation benefits: an empirical comparison of two approaches. *J.Agri.Econ.*, 40:361–374.

Hutchinson, T. and Lai, C. (1991). *The Engineer Statistician's Guide to Continous Bivariate Distributions*, chapter 10-11, pages 157–233. Rumsby Scientific Publishing, first edition.

Kanninen, B. (1993). Optimal experimental design for double-bounded dichotomous choice contingent valuation. *Land Economics*, 69(2):138–146.

Karlin, S. and Studden, W. (1966). Optimal experimental designs. *Ann.Math.Stat.*, 37:783–815.

Kiefer, J. (1959). Optimum experimental designs. *JRSSB*, 21:272–319.

Kiefer, J. (1961). Optimum designs in regression problems ii. In Atkinson and Donev, 1992, page 132.

Kiefer, J. and Wolfowitz, J. (1961). Optimum designs in regression problems. In Atkinson and Donev, 1992, page 132.

Kiefer, J. and Wolftiz, J. (1960). The equivalence of two extremum problems. *Can.J.Math.*, 12:363–6.

Kitsos, C., Titterington, D., and Torsney, B. (1988). An optimal design problem in rhythmometry. *Biometrics*, 44:657–671.

Musrati, A. (1992). Locally optimal designs for binary and weighted regression models. Master's thesis, Department of Statistics, University of Glasgow.

Nyquist, H. (1992). Optimal designs of discrete response experiments in contingent valuation studie. *The Review of Economics and Statistics*, 74:559–563.

Pazman, A. (1986). *Foundations of Optimum Experimental Design.* In Atkinson and Donev, 1992.

Sebastiani, P. and Settimi, R. (1997). A note on d-optimal designs for a logistic regression model. *Journal of statistical planning and inference*, 59:359–368.

Sibson, R. (1972). Discussion on 'results in the theory and construction of d-optimal experimental designs. *J. Roy. Stat. Ser. B*, 34:181–183.

Silvey, S. (1972). Discussion on 'results in the theory and construction of d-optimal experimental designs'. *J. Roy. Stat. Ser. B*, 34:174–175.

Silvey, S. (1980). *Optimal Design. An Introduction to the Theory for Parameter Estimation.* Chapman and Hall.

Sitter, R. and Fainaru, I. (1997). Optimal designs for the logit and probit models for binary data. *Canadian Journal of Statistics,*, 25:175–189.

Sitter, R. and Torsney, B. (1995a). D-optimal designs for generalized linear models. In Kitsos, C. and Muller, W., editors, *MODA4-Advances in Model-Oriented Data Analysis,*, pages 87–102. Physica-Verlag. Refereed Proceedings of the 4th international workshop in Spetses, Greece, June 5-9, 1995.

Sitter, R. and Torsney, B. (1995b). Optimal designs for binary response experiments with two design variables. *Statistica Sinica*, 5:405–419.

Torsney, B. (1983). A moment inequality and monotonicity of an algorithm. In Kortanek, K. O. and Fiacco, A. V., editors, *Proceedings, International Symposium on Semi-Infinite Programming and Applications at University* *of Texas, Austin, Lecture Notes in Economics and Mathematical Systems.*, volume 215, pages 249–260.

Torsney, B. and Alahmadi, A. M. (1992). Further development of algorithms for constructing optimizing distributions. In Fedorov, V., Muller, W., and Vuchkov, I., editors, *Proc. 2nd IIASA Model Oriented Data Workshop in St. Kyrik, Bulgaria*, pages 121–129, Physica-Verlag GmbH Co. P.O. Box 10 52 80 69042 Heildelberg, FRG. Physica-Verlag.

Torsney, B. and Musrati, A. (1993). On the construction of optimal designs with applications to binary response and to weighted regression models. In Muller, W., Wynn, H., and Zhigliausky, A., editors, *Proc. 3rd Model Oriented Data Workshop in Petrodvorets,Russia*, volume 3, pages 37–52, Physica-Verlag GmbH Co. P.O. Box 10 52 80 69042 Heildelberg, FRG. Physica-Verlag.

van Berkum, E. (1995). Optimal designs for regression models. Nederlands.

Whittle, P. (1973). Some general points in the theory of optimal experimental designs. *J. Roy. Statist. Soc. B*, 35:123–130.

Wu, C. (1988). Optimal design for percentile estimation of a quantal response curve. In Y.Dodge, V.Fedorov, H., editor, *Optimal Design and Analysis of Experiments*, pages 213–22. Elsevier Science Publishers B.V.(North Holland).

Wynn, H. (1972). Results in the theory and construction of *d*-optimal experimemtal designs. *J.Roy.Stat.Ser.B*, 34:133–147,170–186. (with discussion).

Appendix A

Monotonocity of $q_3(r)$

In this appendix, we prove that $q_3(r)$ is an increasing function of r.

$$q_3(r) = \frac{(r-2)+\sqrt{A}}{6(r-1)}$$

where $A = (r-2)^2 + 3(r-1)$.

To prove that $q_3(r)$ is increasing, it suffices to prove that $\dfrac{\partial q_3(r)}{\partial r} > 0$. Now let's consider the following expression of the derivative of $q_3(r)$ with respect to r:

$$\begin{aligned}
\frac{\partial q_3(r)}{\partial r} &= \frac{\left[1+\frac{1}{2}A^{\frac{-1}{2}}\left[2(r-2)+3\right]\right](6(r-1)) - 6\left[(r-2)+\sqrt{A}\right]}{36(r-1)^2} \\
&= \frac{6(r-1)}{36(r-1)^2} + \frac{\left[2(r-2)+3)\right]\left[6(r-1)\right]}{72(r-1)^2\sqrt{A}} - \frac{6(r-2)}{36(r-1)^2} - \frac{6\sqrt{A}}{36(r-1)^2} \\
&= \frac{1}{72(r-1)^2\sqrt{A}} \\
&\quad \Big\{12\sqrt{A}(r-1) + 12(r-1)(r-2) + 18(r-1) - 12(r-2)\sqrt{A} \\
&\quad -12\left[(r-2)^2+3(r-1)\right]\Big\} \\
&= \frac{1}{72(r-1)^2\sqrt{A}}\Big\{12\sqrt{A} + 12(r-1)(r-2) + 18(r-1) - 12(r-2)^2 \\
&\quad -36(r-1)\Big\} \\
&= \frac{1}{72(r-1)^2\sqrt{A}}\left\{12\sqrt{A} + 6(r-1)\left[2(r-2)-3\right] - 12(r-2)^2\right\}
\end{aligned}$$

$$= \frac{1}{72(r-1)^2\sqrt{A}}\left\{12\sqrt{A}+6(r-1)(2r-7)-12(r-2)^2\right\}$$
$$= \frac{1}{72(r-1)^2\sqrt{A}}\left\{12\sqrt{A}+12r^2-54r+42-12r^2+48r-48\right\}$$
$$= \frac{1}{12(r-1)^2\sqrt{A}}\left\{\underbrace{2\sqrt{A}}_{LHS}-\underbrace{(r+1)}_{RHS}\right\}$$

From the above expression of $\frac{\partial q_3(r)}{\partial r}$, all we need to do is to prove that $LHS > RHS$ since $\frac{1}{36(r-1)^2\sqrt{A}} > 0$. Now, because $LHS > 0$ and $RHS > 0$, proving that $(LHS)^2 > (RHS)^2$ is equivalent to proving that $(LHS) > (RHS)$. Let $f(r) = LHS^2 - RHS^2$. If we expand the expression of $f(r)$, then we get the following:

$$\begin{aligned}
f(r) &= LHS^2 - RHS^2 \\
&= \left[2\sqrt{(r-2)^2+3(r-1)}\right]^2 - (r+1)^2 \\
&= 4\left[(r-2)^2+3(r-1)\right] - (r+1)^2 \\
&= 4\left[r^2-4r+4+3r-3\right] - \left[r^2+2r+1\right] \\
&= 4r^2-16r+16+12r-12-r^2-2r-1 \\
&= 3r^2-6r+3 \\
&= 3(r^2-2r+1) \\
&= 3(r-1)^2
\end{aligned}$$

The above expression of $f(r)$ clearly shows that $f(r) > 0$ which is equivalent to $LHS^2 > RHS^2$.

□

Appendix B

Study of function $h_y(z)$

B.1 Monotonocity of $q(r)$.

In this appendix, we prove $q(r)$ is an increasing function with respect to r.

$$q(r(z)) = \frac{(k-1)(r(z)-1) - 2 + \sqrt{A}}{2k2^{(k-2)}(r(z)-1)}$$

where $A = (k-1)^2(r(z)-1)^2 + 4r(z)$.

To prove that $q(r)$ is increasing, it suffices to prove that $\frac{\partial q(r)}{\partial r} > 0$. Now let's consider the following expression of the derivative of $q(r)$ with respect to r:

$$\begin{aligned}
\frac{\partial q(r)}{\partial r} &= \frac{1}{2k2^{(k-2)}}\left\{\frac{(k-1)(r-1) - [(k-1)(r-1)-2]}{(r-1)^2}\right\} \\
&+ \frac{1}{2k2^{(k-2)}}\left\{\frac{\frac{1}{2}\left[(k-1)^2(r-1)^2 + 4r\right]^{\frac{-1}{2}}\left[(k-1)^2 2(r-1) + 4\right](r-1)}{(r-1)^2}\right\} \\
&- \frac{1}{2k2^{(k-2)}}\left\{\frac{\sqrt{A}}{(r-1)^2}\right\} \\
&= \frac{1}{2k2^{(k-2)}}\left\{\frac{2}{(r-1)^2} + \frac{2\left[(r-1)(k-1)^2 + 2\right]}{2(r-1)\sqrt{A}} - \frac{\sqrt{A}}{(r-1)^2}\right\} \\
&= \frac{1}{2k2^{(k-2)}}\left\{\frac{2}{(r-1)^2} + \frac{(r-1)(k-1)^2 + 2}{(r-1)\sqrt{A}} - \frac{\sqrt{A}}{(r-1)^2}\right\} \\
&= \frac{1}{2k2^{(k-2)}}\left\{\frac{2}{(r-1)^2} + \frac{(k-1)^2}{\sqrt{A}} + \frac{2}{(r-1)\sqrt{A}} - \frac{\sqrt{A}}{(r-1)^2}\right\}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2k2^{(k-2)}(r(z)-1)^2\sqrt{A}} \\
&\quad \left\{2\sqrt{A} + (k-1)^2(r-1)^2 + 2(r-1) - \left(\sqrt{A}\right)^2\right\} \\
&= \frac{1}{2k2^{(k-2)}(r(z)-1)^2\sqrt{A}} \\
&\quad \left\{2\sqrt{A} + (k-1)^2(r-1)^2 + 2(r-1) - (k-1)^2(r-1)^2 - 4r\right\} \\
&= \frac{1}{2k2^{(k-2)}(r(z)-1)^2\sqrt{A}}\left\{2\sqrt{A} - 2r - 2\right\} \\
&= \frac{1}{k2^{(k-2)}(r(z)-1)^2\sqrt{A}}\left\{\underbrace{\sqrt{A}}_{LHS} - \underbrace{(r+1)}_{RHS}\right\}
\end{aligned}
$$

From the above expression of $\frac{\partial q(r)}{\partial r}$, all we need to do is to prove that $LHS > RHS$ since $\frac{1}{k2^{(k-2)}(r(z)-1)^2\sqrt{A}} > 0$. Now, because $LHS > 0$ and $RHS > 0$, proving that $(LHS)^2 > (RHS)^2$ is equivalent to proving that $(LHS) > (RHS)$. Let $f(r) = (LHS)^2 - (RHS)^2$. If we expand the expression of $f(r)$, then we get the following:

$$
\begin{aligned}
f(r) &= LHS^2 - RHS^2 \\
&= \left\{\sqrt{A}\right\}^2 - \{(r+1)\}^2 \\
&= A - (r+1)^2 \\
&= A - [(r-1)+2]^2 \\
&= A - (r-1)^2 - 4r + 4 - 4 \\
&= (k-1)^2(r-1)^2 + 4r - (r-1)^2 - 4r \\
&= (r(z)-1)^2\left\{(k-1)^2 - 1\right\}
\end{aligned}
$$

The above expression of $f(r)$ clearly shows that $f(r) > 0$ which is equivalent to $(LHS)^2 > (RHS)^2$.

□

Appendix C

Conditional Contingent Valuation.

C.1 Critical value of c.

We have shown that the Double Bounded Dichotomus choice Conditional Univariate Model Weight function to be

$$w_c(z) = \frac{f^2(z)}{[1 - F(z)]\,[F(z) - F(c)]} \qquad z > c.$$

In the case of Logistic Distribution $f(z) = F(z)\,[1 - F(z)]$. Because of that we can rewrite $w_c(z)$ as follows:

$$w_c(z) = \frac{\{F(z)\,[1 - F(z)]\}^2}{[1 - F(z)]\,[F(z) - F(c)]} \qquad z > c.$$

To explore the critical value of c, we compute the derivative of the $w_c(z)$ with respect to z[1]. However, before taking the derivative of the $w_c(z)$ we can make some simplifications on the formulae of $w_c(z)$:

$$w_c(z) = \frac{F^2(z)\,[1 - F(z)]}{[F(z) - F(c)]} \qquad z > c. \tag{C.1}$$

[1] As we can see plots of $\sqrt{w_c(z)}$ at Figures (6.2) (6.3) (6.4) (6.5) (b), (e), (k), (h), for negative value of c ($c < -2$) $w_c(z)$ changes from being decreasing to having two TP's

Equation C.1 can be solved as follows:

$$\frac{1}{w_c(z)} = \frac{[F(z)-F(c)]}{F^2(z)\,[1-F(z)]}.$$

Hence

$$\begin{aligned}\ln\left[\frac{1}{w_c(z)}\right] &= \ln[F(z)-F(c)] - 2\ln F(z) - \ln[1-F(z)],\\ -\ln\{w_c(z)\} &= \ln[F(z)-F(c)] - 2\ln F(z) - \ln[1-F(z)].\end{aligned}$$

So

$$\begin{aligned}\frac{\partial \ln[w_c(z)]}{\partial z} &= \frac{f(z)}{[F(z)-F(c)]} - \frac{2f(z)}{F(z)} + \frac{f(z)}{1-F(z)}\\ &= f(z)\left\{\frac{1}{[F(z)-F(c)]} - \frac{2}{F(z)} + \frac{1}{1-F(z)}\right\}\\ &= f(z)\left\{\frac{F(z)\,[1-F(z)] - 2\,[F(z)-F(c)]\,[1-F(z)] + F(z)\,[F(z)-F(c)]}{F(z)\,[1-F(z)]\,[F(z)-F(c)]}\right\}\\ &= f(z)\left\{\frac{[F(z)-F^2(z)] + [F(z)-F(c)]\,[F(z)-2(1-F(z))]}{F(z)\,[1-F(z)]\,[F(z)-F(c)]}\right\}\\ &= f(z)\left\{\frac{[F(z)-F^2(z)] + [F(z)-F(c)]\,[3F(z)-2]}{F(z)\,[1-F(z)]\,[F(z)-F(c)]}\right\}\\ &= f(z)\left\{\frac{F(z)-F^2(z)+3F(z)^2-[2+3F(z)]\,F(z)+2F(c)}{F(z)\,[1-F(z)]\,[F(z)-F(c)]}\right]\\ &= f(z)\left\{\frac{\overbrace{2F^2(z)-[1+3F(c)]\,F(z)+2F(c)}^{Q(z)}}{F(z)\,[1-F(z)]\,[F(z)-F(c)]}\right\}.\end{aligned}$$

Therefore $\frac{\partial \ln[w_c(z)]}{\partial z} = 0$ if $Q(z) = 0$ where $Q(z) = 2F^2(z)-[1+3F(c)]\,F(z)+2F(c)$ **if** $Q(z)$ **has roots.**

Roots of $Q(z)$ can be written as follows:

$$\begin{aligned}F(z)_{1,2} &= \frac{(1+3F(c)) \pm \sqrt{(1+3F(c))^2-16F(c)}}{4}\\ &= \frac{(1+3F(c)) \pm \sqrt{1+6F(c)+9F^2(c)-16F(c)}}{4}\\ &= \frac{(1+3F(c)) \pm \sqrt{1-10F(c)+9F^2(c)}}{4}\\ &= \frac{(1+3F(c)) \pm \sqrt{(1-F(c))(1-9F(c))}}{4}\end{aligned}$$

There are three possibilities regarding the discriminant:

$$[1 - F(c)]\,[(1 - 9F(c)] = \begin{cases} = 0 & if \quad F(c) = 1 or F(c) = \frac{1}{9} \\ > 0 & if \quad 0 \leq F(c) \leq \frac{1}{9} \\ < 0 & if \quad F(c) > \frac{1}{9} \end{cases}$$

Thus $Q(z)$ has roots iff $0 \leq F(c) \leq \dfrac{1}{9}$ i.e. for $c \leq F^{-1}(1/9)$. For the logistic $c = \ln(1/8) = -2.07944$.