OPTIMUM DESIGN 2000

OPTIMUM DESIGN 2000

Editors

Anthony Atkinson
London School of Economics

Barbara Bogacka
Queen Mary & Westfield College

Anatoly Zhigljavsky
Cardiff University

The volume is dedicated to Valeri Fedorov
on the occasion of his 60th birthday

Preface

This volume contains many of the papers presented at the conference “Optimum Design 2000: Prospects for the New Millennium" held in Cardiff, UK on April 12th – 14th, 2000. The majority of the papers consider aspects of optimum experimental design from the theoretical to applications. Many of the optimisation problems arising in the optimum design theory in general and the articles in this volume in particular, fall into the category of nonconvex, including global, optimization.

The papers are organised in two sections. Since we are at the beginning of a new millennium the first paper starts by looking back at the beginnings of optimum experimental design in 1917 and sketches the development up to Kiefer’s paper read to the Royal Statistical Society in 1959. This is the first in a group of papers which we have labelled "Theory", which cover the more general aspects, such as the properties and methods of construction of designs. In the "Applications" section there are papers on sequential design problems arising in the pharmaceutical industry and on the designs with discrete factors which occur in agriculture. There are also papers on training neural networks, on the efficient design of sampling methods, on design problems in mixtures, arising from a study of glass manufacturing, and on the resistance of Brazilian weeds to herbicides.

The number attending the meeting was deliberately limited to 50 in order to encourage a workshop atmosphere and the exchange of ideas. The international nature of the conference is shown by the presence of scientists from 15 countries. The attendance of people from so many countries, especially those of Central Europe, was made possible by the generosity of our sponsor SmithKline Beecham. The organisers are most grateful to this company and especially to Dr. Darryl Downing for the financial support. We are also grateful to the Royal Statistical Society for endorsing our conference, which was organised under their auspices. The University of Cardiff and the capital city of Wales provided an excellent environment not only for the formal sessions, but also for those informal discussions which are such an important part of any scientific conference. Photographs at the end of the book provide a glimpse of two evenings of technical and social interaction. We are grateful to Patrick Laycock of UMIST for making his photographs available.

The contributions to this volume have been selected by the editors from those submitted for publication. All have been refereed and, where necessary,

Englished. Four papers have discussions. We thank the referees and discussants for their contributions to ensuring the quality of these proceedings.

We trust that this volume will be of interest to statisticians and the wider scientific community and will serve to provide them with a view of the breadth and excitement of work in what is currently a very active area.

Anthony Atkinson
Barbara Bogacka
Anatoly Zhigljavsky

Contents

Part II. Applications

PART ONE

T H E O R Y

Chapter 1

SOME HISTORY LEADING TO DESIGN CRITERIA FOR BAYESIAN PREDICTION

Anthony C. Atkinson
Department of Statistics, London School of Economics
London WC2A 2AE, UK
a.c.atkinson@lse.ac.uk

V. V. Fedorov
SmithKline Beecham Pharmaceuticals
Valeri V Fedorov@sbphrd.com

Abstract After a short history of optimum design we develop design criteria for Bayesian prediction in which a combined forecast is used

Keywords: D-optimality, discrimination between models

Introduction

The standard theory of optimum experimental design provides a simple, clear and elegant approach for a well-defined problem with a single model. We first review the history of the development of the theory for one model, before recalling work on model discrimination and Bayesian design. We then derive design criteria for Bayesian prediction in which the response of interest is a weighted combination of predictions from fitted models.

At the beginning of a new millennium it is appropriate to glance backwards at the history of optimum design, a subject which goes slightly further back into the previous millennium than Kiefer (1959), the paper which gave the subject its current structure and focus. Some of the earlier history can be gleaned from Kiefer's paper, who however, perhaps perversely, starts with a history of topics he is not going to discuss. A history of statistical work on designed experiments is presented by Atkinson and Bailey (2001) in a paper written to celebrate the centenary of *Biometrika*. They treat designs both for agricultural

experiments, broadly those in which the factors are discrete, and those for industrial experiments, in which the factors are continuous. We concentrate on the latter and give a brief history design for regression models leading, via the Equivalence Theorem, to algorithms for the construction of designs for linear and nonlinear models, in which the work of George Box and coauthors plays a role. The alphabetic nomenclature for optimum designs was introduced by Kiefer (1958). We tend to use it anachronistically.

The history starts with Smith (1918), an amazing paper which was 40 years before its time. It gives G-optimum designs for polynomials of order up to six in a single variable x. The design region X is known, with x between -1 and 1. Finally the observational errors are independent, identically distributed and additive, so that ordinary least squares is the appropriate estimation method. For this well specified situation the author obtained the designs which minimise the maximum variance of prediction for each polynomial over the design region. Not only were the optimum designs found, but they were proved to be optimum. Other designs were considered, for example uniform designs, and the effect of non-constant variance investigated. Kirstine Smith's paper thus had all the components needed to specify an optimum design: a model, a design region, a design criterion and specification of the errors, components we have to specify for our Bayesian prediction problem.

Smith's paper seems not to have had much immediate impact. Wald (1943) suggested the comparison of designs using the values of non-centrality parameters for tests of hypotheses about all parameters. This criterion is impracticable, since it depends on the values of the unknown parameters. As a practical replacement he was led to the generalized variance, equivalent to comparing the determinants of the information matrix, and so introduced D-optimality. His examples are of the efficiency of Latin and higher squares. For the designs for one variable polynomials considered by Smith, de la Garza (1954), in a paper without references, shows that it is only necessary to consider the information matrices of designs with trials at $p + 1$ points, where p is the number of parameters in the model, these points including the coded ends of the design region ± 1. Kiefer (1959) states that this is wrong, the result applying only to continuous designs defined by a probability measure. Guest (1958) shows that, for minimax variance (G-optimality), these design points are roots of an equation in the derivatives of a Legendre polynomial. Hoel (1958) obtains the same designs as minimising the generalized variance, and comments that his designs are the same as those of Guest. Two years later Kiefer and Wolfowitz (1960) proved the equivalence of G- and D-optimality and so showed that the designs do indeed have to be the same.

Much of the structure of optimum designs for regression models was elucidated by Elfving (1952) who considered regression in two variables without a constant. He showed that, if both parameters are of interest, the design is

supported on at most two or three points and gave geometrical arguments as to why this is so. His design criterion was what became known as A–optimality. He also comments that the points of the optimum design lie on an ellipsoid which contains the design region, an insight further developed by Silvey and Titterington (1973). Elfving's results were almost immediately extended by Chernoff (1953) who, following Wald, considered D–optimality (and also D_s–optimality) for models which could be linear or asymptotically linear after a Taylor series expansion.

Box and Lucas (1959) use these results of Chernoff, Elfving and Wald to find locally D-optimum designs for the single response of nonlinear models arising in chemical kinetics and give a geometric interpretation of D–optimality which follows from that of Elfving. For designs in which the number of observations is equal to the number of parameters, which is usually the case with nonlinear models, the optimum design maximizes the area of the simplex in the design locus formed by the design points and the origin. Here the design locus is the space of the parameter sensitivities, that is the partial derivatives of the response with respect to the parameters of the model.

The designs of Box and Lucas are locally optimum. They therefore may change as the parameter estimates change. Box and Hunter (1965) develop a sequential strategy for updating the design one trial at a time as observations become available. They give a Bayesian justification for D-optimality and provide an analytical derivation of the best conditions for the next trial. They prove that trials should be added at the point where the variance of prediction using the linearized model is a maximum, that is so suggestive of the equivalence of G- and D-optimality. It is a short step to consider the same sequential design generation for linear models, into which the parameter values do not enter, and so to obtain the algorithm for the sequential construction of D–optimum designs given by Fedorov (1972), a translation of a 1969 book in Russian, and by Wynn (1970). An early survey in English is that of Fedorov and Malyutov (1972).

1. SEVERAL MODELS

The standard theory is concerned with a single model. Two kinds of extension to design with more than one model are technically much in the same vein:

1 If several models, or aspects of models, are of interest, the standard criteria can be extended to compound criteria which are positively weighted linear combinations of standard criteria, to which an equivalence theorem applies (Atkinson and Cox, 1974; Läuter, 1974);

2 If interest is solely in choosing one model out of two, the $T-$optimum designs of Atkinson and Fedorov (1975a) are appropriate. For more than two models the more complicated sequential search of Atkinson and Fe-

dorov (1975b) is available. Equivalence between certain model selection problems and parameter estimation problems was initially discussed in Wald (1943) and, later, in Fedorov and Khabarov (1986), who proved equivalence for a range of different criteria.

We now consider two Bayesian developments. The first leads to adaptive designs:

1 Box and Hill (1967) describe a Bayesian procedure for discriminating between models which leads to the sequential updating of the prior probabilities of the models. Fedorov and Pázman (1968) developed an adaptive procedure for simultaneous discrimination and parameter estimation. There are however difficulties in the calculation of the probabilities of models of differing dimension, which are described by O'Hagan (1994, Ch. 7);

2 Chaloner and Larntz (1989) derive Bayesian optimum designs in which the locally optimum design criterion for linearised models is replaced by its expectation over the prior distribution of the parameters. There are several examples of the use of such criteria in this volume. Chaloner and Verdinelli (1995) review Bayesian experimental design, using decision theory to unify the development of many criteria.

In this paper we try to be completely within the framework of Bayesian prediction as in Hoeting et al. (1999), where all models are used to predict the behaviour of the response.

2. BAYESIAN PREDICTION OF THE RESPONSE

Let $y^T = (y_1, \ldots, y_N)$ be a vector of N independent observations, $\eta^T = (\eta(x_1), \ldots, \eta(x_N))$ and let

$$p(y|\eta) \;=\; \prod_{i=1}^{N} p(y_i|\eta(x_i)) \;, \quad E\{y|\eta\} = \eta \;.$$

In general the function $\eta(x)$ may coincide with one of K functions $\eta_1(x), \ldots, \eta_K(x)$. For the moment we consider the case of two possible regression models and assume that

$$\eta(x) = \eta_1^a(x)\eta_2^{1-a}(x), \tag{1.1}$$

where the prior distribution of a is Be(π), with $\eta_1(x) = \theta^T f(x)$ and $\eta_2(x) = \gamma^T g(x)$. The form of (1.1) assumes that the model generating the observations is selected from two possible models accordingly to a Bernoulli rule. The conditional distribution of the observations can then be written

$$p(y|\theta, \gamma, a) = p^a(y|\theta^T F)p^{1-a}(y|\gamma^T G) \;,$$

where $F = \{f(x_i)\}_1^N$ and $G = \{g(x_i)\}_1^N$ are the matrices of carriers in the two models.

Not only do we have a mixture of models, we also have prior distributions of the parameters. Let the joint prior distribution be

$$p(\theta, \gamma, a) = \pi^a (1-\pi)^{1-a} p^a(\theta) p^{1-a}(\gamma),$$

when the marginal distribution of the observations is

$$p(y) = \pi \int d\theta \; p(y|\theta^T F) p(\theta) + (1-\pi) \int d\gamma \; p(y|\gamma^T G) p(\gamma)$$

$$= \pi_N p_1(y) + (1-\pi_N) p_2(y),$$

where

$$p_1(y) = \int d\theta \; p(y|\theta^T F) p(\theta), \;\; p_2(y) = \int d\gamma \; p(y|\gamma^T G) p(\gamma),$$

$$\pi_N = \pi p_1(y)/p(y).$$

A discussion of the care that needs to be taken in assigning prior distributions to regression models and their parameters is given by O'Hagan (1994, Ch. 7).

There are two possible motivations for Model (1.1). In the first it is assumed that there are two populations (for instance, males and females) of sizes proportional to π and $1-\pi$ correspondingly. The observations y are collected from only one subject sampled at random from the mixture of these two populations. Another motivation is the "real" Bayesian approach, in which the probability π is understood as a measure of prior belief in Model 1.

If, as in O'Hagan (1994, Ch. 1), we use quadratic loss as a measure of the precision of prediction then the best predictor (or the Bayes estimator with quadratic loss) of the response at the point x is

$$\bar{\eta}(x) = E\{\eta(x)|y\} = \pi_N \bar{\eta}_1(x) + (1-\pi_N)\bar{\eta}_2(x), \tag{1.2}$$

where

$$\bar{\eta}_1(x) = \int d\theta \; \theta^T f(x) p(\theta|y), \; p(\theta|y) = \frac{p(y|\theta^T F) p(\theta)}{p_1(y)} \; ,$$

with a similar definition for $\bar{\eta}_2(x)$. To derive (1.2) we use the Bayes formula

$$p(\theta, \gamma, a|y) = \frac{p(y|\theta, \gamma, a) p(\theta, \gamma, a)}{p(y)} \; .$$

The variance of this prediction is

$$\mathrm{Var}\{\eta(x)|y\} = E\{[\eta_1^a(x)\eta_2^{1-a}(x) - \bar{\eta}(x)]^2|y\}$$

$= \pi_N(1-\pi_N)[\bar{\eta}_1(x)-\bar{\eta}_2(x)]^2 + \pi_N \mathrm{Var}_1\{\eta_1(x)|y\} + (1-\pi_N)\mathrm{Var}_2\{\eta_2(x)|y\},$

which extends straightforwardly in the case of K populations to

$$\mathrm{Var}\{\eta(x)|y\} = \sum_{j,k}^{K} \pi_{Nj}\pi_{N_k}(\bar{\eta}_j(x) - \eta_k(x))^2 + \sum_{j}^{K} \pi_{Nj}\mathrm{Var}_j\{\eta_j(x)|y\}$$

$$= \sum_{j}^{K} \pi_{Nj}[(\bar{\eta}_j(x) - \bar{\eta}(x))^2 + \mathrm{Var}_j\{\eta_j(x)|y\}]. \tag{1.3}$$

A similar expression may be derived for

$$E[\eta(x) - \bar{\eta}(x)][\eta(x') - \bar{\eta}(x')]. \tag{1.4}$$

All criteria used in this paper are based on (1.3) although derivation of more complicated criteria based on (1.4) is relatively straightforward.

3. DESIGN OBJECTIVES AND OPTIMALITY CRITERIA

Let us assume that a practitioner is interested in the response function over the region X^*. Then it is natural for a pessimist to describe the precision measure of the predicted response by

$$\max_{x \in X^*} \mathrm{Var}\{\eta(x)|y\}, \tag{1.5}$$

while a more optimistic practitioner would probably be satisfied if the quantity

$$\mathcal{D}(y, x_1, \ldots, x_N) = \int_{X^*} dx \mathrm{Var}\{\eta(x)|y\}$$

$$= \sum_{j=1}^{K} \pi_{Nj} \int_{X^*} dx\ [(\bar{\eta}_j(x) - \bar{\eta}(x))^2 + \mathrm{Var}_j\{\eta_j(x)|y\}]. \tag{1.6}$$

were small enough.

Both of the preceding criteria depend upon the vector y, that is obviously unknown before the experiment. So, to proceed with optimum design, we have either to use some sequential adaptive methods or to replace (1.5) and (1.6) by quantities which do not include y. In what follows we use only the second approach. To operate within that approach an assumption on linearity with respect to unknown parameters will be needed. In addition we assume the normality of all prior distributions and distributions $p(y_i|\eta(x_i))$ and we also assume the independence of the y_i.

Introducing

$$E\ldots = \int \ldots p(y)dy$$

we may replace $\mathrm{Var}\{\eta(x)|y\}$ by $E\{\mathrm{Var}\{\eta(x)|y\}\}$. The latter does not depend on the observations y and the optimization problem

$$\min_{\xi} E\{\mathrm{Var}\{\eta(x)|y\}\} \tag{1.7}$$

can be considered. Unfortunately, integrating out y is not a trivial problem for any practically reasonable collection of models. One may verify that

$$\mathrm{Var}_j\{\eta_j(x)|y\} = f_j^T(x)[NM_j(\xi) + D_j^{-1}]^{-1} f_j(x), \tag{1.8}$$

where

$$M_j(\xi) = \sigma^{-2} \int_X \xi(dx)\ f_j(x) f_j^T(x), \ \ U_j = \int_{X^*} dx f_j(x) f_j^T(x)\ .$$

Observing that

$$\sum_j^K \pi_{Nj}(\bar{\eta}_j(x) - \bar{\eta}(x))^2 \le \sum_j^K \pi_{Nj}(\bar{\eta}_j(x) - \bar{\bar{\eta}}(x))^2\ ,$$

where $\bar{\bar{\eta}}(x) = \sum_{j=1}^K \pi_j \bar{\bar{\eta}}_j(x), \ \ \bar{\bar{\eta}}_j(x) = \int d\theta\ \theta^T f(x) p(\theta|\eta),$ and $\eta = E\{y\}$, we conclude that

$$E\{\mathrm{Var}\{\eta(x)|y\}\} = E\{\sum_j^K \pi_{Nj}[(\bar{\eta}_j(x) - \bar{\eta}(x))^2 + \mathrm{Var}_j\{\eta_j(x)|y\}]\}$$

$$\le \sum_j^K \pi_j[(\bar{\bar{\eta}}_j(x) - \bar{\bar{\eta}}(x))^2 + f_j^T(x)[NM_j(\xi) + D_j^{-1}]^{-1} f_j(x)]\ . \tag{1.9}$$

Correspondingly

$$\Omega(\xi) = \mathrm{tr} U_j[NM_j(\xi) + D_j^{-1}]^{-1} \le E\{\mathcal{D}\} = E\{\int_{X^*} dx\ \mathrm{Var}_j\{\eta_j|y\}\}$$

$$\le \int_{X^*} dx \sum_j^K \pi_j[(\bar{\bar{\eta}}_j(x) - \bar{\bar{\eta}}(x))^2 + \mathrm{tr} U_j[NM_j(\xi) + D_j^{-1}]^{-1} = \Pi(\xi)\ . \tag{1.10}$$

Thus, we have found lower and upper bounds for the criterion

$$E\{\int_{X^*} dx\ \mathrm{Var}_j\{\eta_j|y\}\}$$

which do not depend upon the observations and can be minimized with respect to ξ prior to an experiment.

The term $\sum_j^K \pi_j(\bar{\bar{\eta}}_j(x) - \bar{\bar{\eta}}(x))^2$ may be viewed as the average variance between models while the second term, $\sum_j^K \pi_j \mathrm{tr} U_j[NM_j(\xi) + D_j^{-1}]^{-1}$, is the average variance within models. The latter is a special case of a compound linear criterion. It can be shown that both $\sum_j^K \pi_j(\bar{\bar{\eta}}_j(x) - \bar{\bar{\eta}}(x))^2$ and $\sum_j^K \pi_j \mathrm{tr} U_j[NM_j(\xi) + D_j^{-1}]^{-1}$ are convex functions of ξ and therefore all the results of optimal design theory can be applied: see Atkinson and Donev (1992), Fedorov and Hackl (1996) and Pukelsheim (1993).

Although the upper bound $\Pi(\xi)$ can be improved the resulting formulae are quite complicated and do not significantly change the approach. Theoretically it is expedient to construct designs minimizing $\Omega(\xi)$ and $\Pi(\xi)$ and then to look for some compromise design. For practical needs it is probably sufficient to find only $\xi^* = \arg\min_\xi \Pi(\xi)$ to guarantee that the average variance of prediction is of a reasonable magnitude.

Averaging is more difficult for criterion (1.5). From the convexity of the "max"-function, Jensen's inequality, (1.3) and (1.8), it follows that

$$E\{\max_{x \in X^*} \mathrm{Var}\{\eta(x)|y\}\} \geq \sum_j^K \pi_j f_j^T(x)[NM_j(\xi) + D_j^{-1}]^{-1} f_j(x) \ . \quad (1.11)$$

We have not been able to find a upper bound for $E\{\max_{x \in X^*} \mathrm{Var}\{\eta(x)|y\}\}$. As a conclusion to this section we recommend minimization of (1.11) which, at least, assures a small value of the maximum of the "within" model variance over X^* .

4. SHANNON INFORMATION

The criteria introduced in the previous section may be considered as generalizations of criteria related to different functions of the variance of the response function. If interest is in the parameters defining the response, the D-criterion is a popular choice in the single model case. In the Bayesian approach this criterion is closely related to the Shannon information of the posterior distribution of the parameters. A similar criterion may be proposed for several competing models. This idea was first explored by Fedorov and Pazman (1968) and, recently, by Sebastiani and Wynn (2000) in another setting.

From the definition of the Shannon information (e.g. Cover and Thomas, 1991) manipulations similar to those in the previous section, combined with the Bayesian calculus, lead to

$$\Upsilon(y) = \sum_{j=1}^K \pi_{Nj} \Upsilon_j(y) + \sum_{j=1}^K \pi_{Nj} \ln \pi_{Nj} \ , \quad (1.12)$$

where

$$\Upsilon_j(y) = \int d\theta \; p(\theta_j|y) \ln p(\theta_j|y) \; .$$

Under the assumptions of normality and linearity

$$\Upsilon_j(y) = \frac{1}{2} \ln |NM_j(\xi) + D_j^{-1}| + \text{const} \tag{1.13}$$

and

$$\begin{aligned} \pi_{Nj} &\sim \frac{\exp(-s_{Nj}^2/2)}{(2\pi)^{(N+m_j)/2}|\Sigma_N + F_j^T D^j F_j|} \\ &\sim \frac{\exp(-s_{Nj}^2/2)}{(2\pi)^{(N+m_j)/2}|D_j|^{1/2}|NM_j(\xi) + D_j^{-1}|^{1/2}} \; , \end{aligned} \tag{1.14}$$

where

$$\{\Sigma_N\}_{ii} = \sigma^2/r_i, \;\; r_i = N\xi(x_i),$$

$$s_{Nj}^2 = \min_{\theta_j} \sigma^{-2} \sum_{i=1}^{N} [y_i - \theta_j^T f_j(x_i)]^2 \; , \quad F_j = \{f_j(x_i)\}_1^N \; .$$

Again, $\Upsilon(y)$ depends on the observations y and we need to introduce $\bar{\Upsilon} = E\{\Upsilon(y)\}$ in order to be able to select a design prior to an experiment. If Υ_0 is the prior information, then from (1.12) and the convexity of the function $t \ln t$ it follows that

$$\bar{\Upsilon}(\xi) - \Upsilon_0 \geq \frac{1}{2} \sum_{j=1}^{K} \pi_j \ln |NM_j(\xi) + D_j^{-1}| = -\frac{1}{2} \sum_{j=1}^{K} \pi_j \ln |D_j^{-1}| = \Delta(\xi) \; . \tag{1.15}$$

Once more we are faced with a criterion which is a particular case of compound D-optimality, with weights coinciding with the prior probabilities of the competing models. As in the Kiefer equivalence theorem it can be verified that the two optimization problems

$$\xi^* = \arg \min_{\xi} \max_{x \in X} \sum_{j}^{K} \pi_j f_j^T(x) [NM_j(\xi) + D_j^{-1}]^{-1} f_j(x)$$

and

$$\xi^* = \arg \max_{\xi} \Delta(\xi)$$

are equivalent; see (1.11).

5. OPEN PROBLEMS

• We have not managed to derive any reasonable upper bounds for the criterion $E\{\max_{x\in X^*} \mathrm{Var}\{\eta(x)|y\}\}$ which depend in a constructive way on the variance between models. Optimization of the lower bound may sometimes give a design which is good for the original criterion. Despite this caveat, the optimization leads to sound practical results when compared with those of Box and Hill (1967) and of Meyer et al. (1996). Only designs minimizing the upper bound can guarantee that the value of the criterion is not greater than a known level.

• If $\sigma^2/N \to 0$, the average increment of information Υ will be approximately defined by the function

$$\varphi(x,\xi_N) = \sum_{j=1}^{K} \pi_{Nj}[d_j(x,N) + (\bar{\eta}_{Nj}(x) - \bar{\eta}(x))^2] \,, \tag{1.16}$$

$$d_j(x,N) = f_j^T(x)[NM_j(\xi) + D_j^{-1}]^{-1} f_j(x) \,.$$

See Fedorov and Pazman (1968) and Sebastiani and Wynn (2000). Fedorov and Pazman proposed an adaptive design based on (1.16) in which

$$x_{N+1} = \arg\min_{x\in X} \varphi(x,\xi_N). \tag{1.17}$$

While simulations show that the adaptive procedure works well, not only at each step but also asymptotically, there are no rigourous results on its asymptotic behaviour. The similarity (both formal and in interpretation) of the function $\varphi(x,\xi)$ to the celebrated function $d(x,\xi)$ from Kiefer's equivalence theorem makes us expect the equivalence the following optimization problems:

$$\max_{\xi} \bar{\Upsilon}(\xi) \quad \text{and} \quad \min_{\xi} \max_{x\in X} E\{\mathrm{Var}\{\eta(x)|y\}\} \,,$$

the latter expression occurring also in (1.9).

References

Atkinson, A.C. and Bailey, R.A. (2001). One hundred years of designed experiments on and off the pages of "Biometrika". *Biometrika* **88**. (In press).

Atkinson, A.C. and Cox, D.R. (1974). Planning experiments for discriminating between models (with discussion). *Journal of the Royal Statistical Society* B **36**, 321–348.

Atkinson, A.C. and Donev, A.N. (1992). *Optimum Experimental Designs*. Oxford: Oxford University Press.

Atkinson, A.C. and Fedorov, V.V. (1975a). The design of experiments for discriminating between two rival models. *Biometrika* **62**, 57–70.

Atkinson, A.C. and Fedorov, V.V. (1975b). Optimal design: experiments for discriminating between several models. *Biometrika* **62**, 289–303.

Box, G. E. P. and Hill, W. J. (1967). Discrimination among mechanisitc models. *Technometrics* **9**, 57–71.

Box, G.E.P. and Hunter, W.G. (1965). Sequential design of experiments for nonlinear models. In *Proceedings IBM Scientific Computing Symposium: Statistics* Ed. J. J. Korth, pp. 113–137. White Plains: IBM.

Box, G.E.P. and Lucas, H.L. (1959). Design of experiments in nonlinear situations. *Biometrika* **46**, 77–90.

Chaloner, K. and Larntz, K. (1989). Optimal Bayesian design applied to logistic regression experiments. *Journal of Statistical Planning and Inference* **21**, 191–208.

Chaloner, K. and Verdinelli, I. (1995). Bayesian experimental design: a review. *Statistical Science* **10**, 273–304.

Chernoff, H. (1953). Locally optimal designs for estimating parameters. *Annals of Mathematical Statistics* **24**, 586–602.

Cover, T.M. and Thomas, M. (1991). *Elements of Information Theory*. New York: Wiley.

de la Garza, A. (1954). Spacing of information in polynomial regression. *Annals of Mathematical Statistics* **25**, 123–130.

Elfving, G. (1952). Optimum allocation in linear regression theory. *Annals of Mathematical Statistics* **23**, 255–262.

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

Fedorov, V.V. and Hackl, P. (1997). *Model-Oriented Design of Experiments*. Lecture Notes in Statistics 125. New York: Springer.

Fedorov, V. and Khabarov, V. (1986). Duality of optimal designs for model discrimination and parameter estimation. *Biometrika* **73**, 183–190.

Fedorov, V.V. and Malyutov, M.B. (1972). Optimal designs in regression problems. *Math. Operationsforsch. u. Statistik* **3**, 281–308.

Fedorov, V.V. and Pázman, A. (1968). Design of physical experiments (statistical methods). *Fortschritte der Physik* **24**, 325–345.

Guest, P.G. (1958). The spacing of observations in polynomial regression. *Annals of Mathematical Statistics* **29**, 294–299.

Hoel, P.G. (1958). Efficiency problems in polynomial estimation. *Annals of Mathematical Statistics* **29**, 1134–1145.

Hoeting, J.A., Madigan, D., Raftery A.E. and Volinsky C.T. (1999). Bayesian model averaging: a tutorial. *Statistical Science* **14**, 382–417.

Kiefer, J. (1958). On the nonrandomized optimality and randomized nonoptimality of symmetrical designs. *Annals of Mathematical Statistics* **29**, 675–699.

Kiefer, J. (1959). Optimum experimental designs (with discussion). *J. Royal Statistical Soc.* B **21**, 272–319.

Kiefer, J. and Wolfowitz, J. (1960). The equivalence of two extremum problems. *Canadian Journal of Mathematics* **12**, 363–366.

Läuter, E. (1974). Experimental design in a class of models. *Mathematische Operationsforschung und Statistik, Serie Statistik* **5**, 379–398.

Meyer, R.D., Steinberg, D.M. and Box, G.E.P. (1996). Follow-up designs to resolve confounding in multifactor experiments. *Technometrics* **38**, 303–332.

O'Hagan, A. (1994). *Kendall's Advanced Theory of Statistics, Volume 2B: Bayesian Inference*. London: Edward Arnold.

Pukelsheim, F. (1993). *Optimal Design of Experiments*. New York: Wiley.

Sebastiani, P. and Wynn, H.P. (2000). Maximum entropy sampling and optimal Bayesian experimental design. *J. Royal Statistical Soc.* B **62**, 145–157.

Silvey, S.D. and Titterington, D.M. (1973). A geometric approach to optimal design theory. *Biometrika* **60**, 15–19.

Smith, K. (1918). On the standard deviations of adjusted and interpolated values of an observed polynomial function and its constants and the guidance they give towards a proper choice of the distribution of observations. *Biometrika* **12**, 1–85.

Wald, A. (1943). On the efficient design of statistical investigations. *Annals of Mathematical Statistics* **14**, 134–140.

Wynn, H.P. (1970). The sequential generation of D-optimal experimental designs. *Annals of Mathematical Statistics* **41**, 1055–1064.

Chapter 2

OPTIMAL DESIGNS FOR THE EVALUATION OF AN EXTREMUM POINT

R.C.H. Cheng
Department of Mathematics, University of Southampton,
Highfield, Southampton, SO17 1BJ, UK
rchc@maths.soton.ac.uk

V.B.Melas and A.N. Pepelyshev
Faculty of Mathematics and Mechanics,
St.Petersburg State University,
Bibliotechnaya sq.2, St.Petersburg, 198904, Russia.
viatcheslav.melas@pobox.spbu.ru

Abstract This paper studies the optimal experimental design for the evaluation of an extremum point of a quadratic regression function of one or several variables. Experimental designs which are locally optimal for arbitrary dimension k among all approximate designs are constructed (although for $k > 1$ an explicit form proves to be available only under a restriction on the location of the extremum point). The result obtained can be considered as an improvement of the last step of the well-known Box-Wilson procedure.

Introduction

Let us consider the problem of optimal design for estimating the point of extremum — $\beta_2/2\beta_3$ in the quadratic regression model

$$y_i = \beta_1 + \beta_2 x_i + \beta_3 x_i^2 + \varepsilon_i, \ \ i = 1, \ldots, n,$$

where y_i is the result of an observation at point $x_i \in [-1, 1]$ and $\{\varepsilon_i\}$ are i.i.d. random values such that $E\varepsilon_i = 0$ and $E\varepsilon_i^2 = 1$.

This problem is a nonlinear one. There exist a few standard approaches to such problems: sequential, minimax, Bayesian and locally optimal. All these approaches have been implemented for the above model. Ford and Silvey (1980) and Müller and Pötscher (1992) considered sequential procedures,

Mandal (1978), Chaloner (1989), Mandel and Heiligers (1992), Müller (1995) and Müller and Pazman (1998) studied the problem from Bayesian or minimax points of view.

Locally optimal designs were considered by Fedorov and Müller (1997). The authors of that paper investigated also the multivariate regression model. They suggested to use a known reparametrization of the problem that allows to obtain a convenient representation of the information matrix (the reparametrization was also used by Chatterjee and Mandal (1981), Buonaccorsi and Iyer (1986) and others).

The present paper can be considered as a further development of this approach. We consider the multivariate quadratic regression model on the hypercube $[-1,1]^k$. We find, analytically, locally optimal designs for this model under a restriction on the location of the extremum point. More precisely, our analytical solution is appropriate if the extremum point belongs to the hypercube $[-1/2,1/2]^k$.

This problem can be considered as part of a more general problem, that of extremum point evaluation for a function of general form. To be specific, we will consider estimation of the minimum point. Investigation of this problem was initiated in the well-known paper by Box and Wilson (1951). The approach to the problem, elaborated in that paper, is based on the combination of the steepest descent technique with methods of planning factorial experiments. Box and Wilson suggested that experiments be performed in the vicinity of a certain basic point, according to a complete or fractional factorial design from which a linear model can be built using the resulting observations. If this model proves to be adequate, it gives an unbiased estimate of the function gradient. A few test steps are then performed in the direction of the antigradient until a decrease in the measured values of the function is observed. The last successful point is taken as a new basic point and so on. In the vicinity of an extremum point the linear model will be inadequate and, when this occurs, a quadratic model can be built for the final estimation of location of the extremum point.

A review of some other approaches to this general problem can be found in Pronzato and Walter (1993).

The analytical solution elaborated in the present paper can be used to optimize the last stage of the Box–Wilson procedure. The formal outline of the problem is given in Sections 1 and 2. Further, in Section 3, we formulate our basic results. A short discussion is given in Section 4. Proofs of the basic results are concentrated in the Appendix.

1. PRELIMINARY OUTLINE OF THE PROBLEM

Consider a quadratic function of several variables:

$$\eta(x) = \eta(x, A, \beta, \gamma) = x^T A x + \beta^T x + \gamma, \tag{2.1}$$

where A is a positive definite $k \times k$ matrix, β is a k dimensional vector, γ is a real number. This function attains, as is well known, its minimal value equal to $c = \gamma - \beta^T A^{-1}\beta/4$ at the point $x = x^* = b = -\frac{1}{2}A^{-1}\beta$.

Suppose the function can only be measured with a random error at design points x belonging to the hypercube $\mathcal{X} = [-1, 1]^k$. More precisely, let the experimental results at the design points $x_{(i)}$, $i = 1, 2, \ldots, n$, $x_{(i)} \in \mathcal{X}$ be described by the equation

$$y_i = \eta(x_{(i)}, A, \beta, \gamma) + \varepsilon_i, \quad i = 1, 2, \ldots, n, \tag{2.2}$$

where $\{\varepsilon_i\}$ are random errors such that $E\varepsilon_i = 0$, $E\varepsilon_i\varepsilon_j = 0$ $i \neq j$, $E\varepsilon_i^2 = \sigma^2$, $i, j = 1, 2, \ldots, n$.

Elements of the matrix A and of the vector β as well as γ are unknown. It is required to construct an estimate of the vector $x^* = b$ and to find an optimal (in a sense to be yet defined) experimental design

$$\xi = \{x_{(1)}, \ldots, x_{(n)}; m_1, \ldots, m_n\},$$

where $x_{(i)} \in \mathcal{X}$ and $m_i > 0$ are respectively the design point and the proportion of the total number of available experiments to be performed at that design point, for $i = 1, 2, \ldots, n$, with $\sum_{i=1}^n m_i = 1$.

2. THE APPROACH TO THE PROBLEM

Rewrite the regression function in the form

$$\bar{\eta}(x, \theta) = (x - b)^T A(x - b) + c, \tag{2.3}$$

where

$$\begin{aligned} \theta &= (b_1, \ldots, b_k, a_{11}, \ldots, a_{kk}, a_{12}, \ldots, a_{1k}, a_{23}, \ldots, a_{k-1k}, c)^T, \\ b &= -\frac{1}{2}A^{-1}\beta, \ c = \gamma - \beta^T A^{-1}\beta/4. \end{aligned}$$

This function depends on parameters $b_1, \ldots, b_k$ in a nonlinear way. We apply to our case well-known results for nonlinear models.

Let $\xi = \{x_{(1)}, \ldots, x_{(n)}; m_1, \ldots, m_n\}$ be an experimental design. Consider the matrix

$$M(\xi) = \left(\sum_{l=1}^{n} \frac{\partial \bar{\eta}(x_{(l)}, \theta)}{\partial \theta_i} \frac{\partial \bar{\eta}(x_{(l)}, \theta)}{\partial \theta_j} m_{(l)} \right)_{i,j=1}^{r},$$

$r = k + k(k + 1)/2 + 1$.

An immediate calculation gives

$$M(\xi) = \begin{pmatrix} 2A & 0 \\ 0 & I \end{pmatrix} \bar{M}(\xi) \begin{pmatrix} 2A & 0 \\ 0 & I \end{pmatrix},$$

where I is the identity matrix,

$$\begin{aligned}\bar{M}(\xi) &= \bar{M}(\xi,b)=\sum_{l=1}^{n} f(x_{(l)})f^T(x_{(l)})m_l,\\ f(x) &= f(x,b)=\big((b_1-x_1),\dots,(b_k-x_k),(x_1-b_1)^2,\dots,(x_k-b_k)^2,\\ &\qquad (x_1-b_1)(x_2-b_2),\dots,(x_{k-1}-b_{k-1})(x_k-b_k),1\big)^T.\end{aligned}$$

Note that for $n = r$ we have $\det \bar{M}(\xi,b) = \det \bar{M}(\xi,0)$. This can be verified through the equality $\bar{M}(\xi,b) = F^T F$, where $F = \big(\sqrt{m_j} f_i(x_{(j)})\big)_{i,j=1}^{r}$ and linear transformations of columns of the matrix F. Then, with $n \geq r$, the Binet-Cauchy formula states that $\det \bar{M}(\xi,b) \neq 0$ for an arbitrary b if $\det \bar{M}(\xi,0) \neq 0$.

An arbitrary design ξ will be called a *nonsingular* design if $\det \bar{M}(\xi,0) \neq 0$. From the above remark we have $\det \bar{M}(\xi,b) \neq 0$ for an arbitrary vector b if the design ξ is a nonsingular design.

Consider the (nonlinear) least squares estimate of the vector θ for the regression function (2.3):

$$\hat{\theta} = \arg\min_{\theta} \sum_{l=1}^{N} \big(\bar{\eta}(x_{(l)},\theta) - y_l\big)^2. \tag{2.4}$$

Since $\bar{\eta}(x,\theta) = \eta(x,A,\beta,\gamma)$ we have $\hat{b} = -\frac{1}{2}\bar{A}^{-1}\bar{\beta}$, where $\bar{A}$ is the matrix consisting of least squares estimates of elements of the matrix A in the linear regression function $\eta(x,A,\beta,\gamma)$ under the equation (2.2) and $\bar{\beta}$ is the similar estimate for β. Thus the construction of the estimate $\hat{\theta}$ is easy.

At the same time we have the following proposition:

Proposition 2.1. *Let ξ be an arbitrary nonsingular design and $\hat{\theta}$ be determined by (2.4), where $y_1,\dots,y_N$ are results obtained from Nm_j experiments at point $x_{(j)}$, $j = 1,\dots,n$. Then $\hat{\theta}$ is a strongly consistent estimate for θ and with $N \to \infty$ the vector $\sqrt{N}(\hat{\theta}-\theta)$ has asymptotically the normal distribution with zero expectation and the covariance matrix $\mathcal{D}_{\hat{\theta}} = \sigma^2 M^{-1}(\xi)$.*

This proposition is a particular case of results obtained in Jennrich (1969).

Rewrite the matrix $\bar{M}(\xi)$ in the block form

$$\bar{M}(\xi) = \begin{pmatrix} M_1 & M_2^T \\ M_2 & M_3 \end{pmatrix},$$

where M_1 is a $k \times k$ matrix. Let

$$M_s = M_s(\xi) = M_1 - X^T M_3 X,$$

where $X = M_3^{-1} M_2$ if the matrix M_3 is nonsingular and X is an arbitrary solution of the equation $M_3 X = M_2$ otherwise. It is known (Karlin & Studden, 1966, §10.8) that M_s does not depend on the choice of the solution.

Let $\xi = \alpha\xi_1 + (1-\alpha)\xi_2$, where ξ_1 is a nonsingular design and ξ_2 is an arbitrary design, $0 < \alpha < 1$. Then the covariance matrix of $\sqrt{N}(\hat{b} - b)$, where $\hat{b}$ is the nonlinear least squares estimate of the vector b, takes the form

$$\mathcal{D}_{\hat{b}}(\xi) = \frac{\sigma^2}{4} A^{-1} M_s^{-1}(\xi) A^{-1}.$$

Definition 2.1. *Any design ξ_2 minimizing the magnitude*

$$\lim_{\alpha \to 0} \det \mathcal{D}_{\hat{b}}(\xi) = \left(\frac{\sigma^2}{4}\right)^k (\det A)^{-2} \det M_s^{-1}(\xi_2, b)$$

for a fixed value $b = b^{(0)}$ will be called a locally optimal design for estimating an extremum point of the regression function (2.1). Note that the locally optimal design depends on b and does not depend on the true values of A and c.

It is evident that for large N the majority of experiments should be performed in accordance with a locally optimal design where $b^{(0)}$ is the current estimate of the vector b.

The problem of finding locally optimal designs is equivalent to the problem $\det M_s(\xi, b) \to \max_\xi$, where the maximum is taken over all (approximate) designs and $b = b^{(0)}$. The remainder of the paper considers the solution of this problem.

3. BASIC RESULTS

Let k be an arbitrary natural number and $b = b^{(0)} \in \mathrm{Int}[-1, 1]^k$. Consider all hyperparallelepipeds with centre at the point b inscribed in the hypercube $[-1, 1]^k$ and take the maximal one. Let ξ^* be the experimental design that consists of all vertices of this hyperparallelepiped with equal weights, $m_l = 1/n$, $l = 1, \ldots, n$, $n = 2^k$.

Our basic result is the following theorem:

Theorem 1. *For an arbitrary k, the design ξ^* described above is a locally optimal design for estimation of an extremum point of the regression function (2.1) if and only if $|b_i| \le \frac{1}{2}$, $i = 1, 2, \ldots, k$.*

The result of Theorem 1 for the case $k = 2$ is illustrated in Fig. 2.1.

The exact analytical solution given by Theorem 1 allows us to study the efficiency of the locally optimal design. This design is more accurate than the usual D-optimal design for the estimation of all parameters of the regression function (2.1). For example, suppose $b = (\frac{1}{2}, \ldots, \frac{1}{2})^T$. In this case we need s_k times less observations to receive the same accurancy under dimension k, where $s_1 = 1.5$, $s_2 \approx 1.78$, $s_3 \approx 2.08$, $s_4 \approx 2.38$, $s_5 \approx 2.68$. We see that the efficiency of the locally optimal design increases with k.

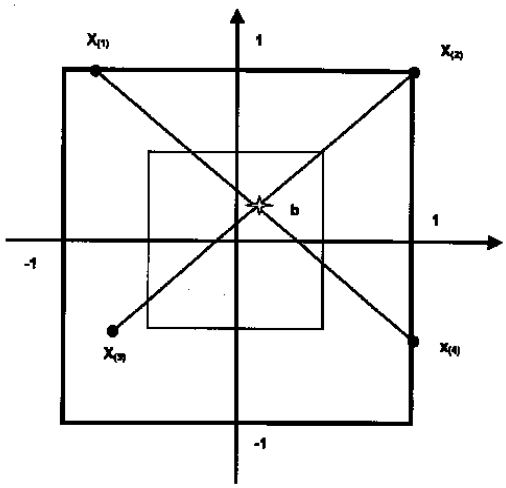


Figure 2.1 Points of the design ξ^* for $k = 2$.

Note that the design ξ^* corresponds to the full factorial design. For $k \geq 4$ it is possible to construct a locally optimal design with the number of points less than 2^k. Assume that $k \geq 4$. Let ν be a natural number such that

$$2^{\nu-1} \geq k, \ \nu \leq k. \tag{2.5}$$

Let $z_{(1)}, \ldots, z_{(n)}$, $n = 2^\nu$ be distinct vertices of the hypercube $[-1, 1]^\nu$. Assign to the number $j = \nu + 1, \ldots, k$ a vector $(j_1, \ldots, j_\alpha)$ whose components are distinct numbers from the set $\{1, \ldots, \nu\}$, with α an odd number, $\alpha \geq 3$. The assignment is such that distinct vectors correspond to distinct numbers, which is possible since the number of the distinct vectors equals

$$\sum_{i=2}^{[\nu/2]} C_\nu^{2i-1} = 2^{\nu-1} - \nu \geq k - \nu.$$

Introduce the vectors $\tilde{z}_{(i)} = (\tilde{z}_{(i)1}, \ldots, \tilde{z}_{(i)k})^T$, $i = 1, 2, \ldots, n$,

$$\begin{aligned} \tilde{z}_{(i)j} &= z_{(i)j}, \ j = 1, \ldots, \nu, \\ \tilde{z}_{(i)j} &= z_{(i)j_1} \ldots z_{(i)j_\alpha}, \ j = \nu + 1, \ldots, k. \end{aligned}$$

Now consider the vectors $x_{(i)} = (x_{(i)1}, \ldots, x_{(i)k})^T$, $i = 1, \ldots, n$, which are vertices of the hyperparallelepiped described above and correspond to the vectors $\tilde{z}_{(i)}$ in such a way that $\operatorname{sign}(x_{(i)j} - b_j) = \operatorname{sign}(\tilde{z}_{(i)j})$, $i = 1, \ldots, n = 2^\nu$, $j = 1, \ldots, k$. Let

$$\xi^+ = \{x_{(1)}, \ldots, x_{(n)}; \frac{1}{n}, \ldots, \frac{1}{n}\}.$$

Theorem 2. *If $b \in [-1/2, 1/2]^k$, and $k \geq 4$, then ξ^+ is a locally optimal design for the estimation of an extremum point of the function (2.1).*

Note that with k satisfying the inequality (2.5) we can construct a locally optimal design ξ^+ with $n = 2^\nu$ distinct points. Let ν^* be the minimal number satisfying (2.5) for a given k, and set $n^*(k) = 2^\nu$. It seems important that when k increases the corresponding value $n^*(k)$ increases very slowly. For example, with $5 \leq k \leq 10$ we have $n^*(k) = 16$.

4. DISCUSSION

Consider the results obtained in the context of the Box-Wilson procedure. It is evident that at the last stage of this procedure we have already a rough estimate of the extremum point and it is only necessary to make it more exact.

If the extremum point lies rather far from the boundary of the design region, it is reasonable to use the design $\alpha\xi_1 + (1-\alpha)\xi_2$, where ξ_1 is the usual D-optimal design and ξ_2 is the locally optimal design given by Theorem 1. It can be shown that the asymptotically best value of α is $\alpha = \alpha^*/\sqrt{N}$, where α^* does not depend on N.

However the calculation of α^* is reasonably combined in practice with the elaboration of a sequential procedure . This problem exceeds the scope of the present article and will be considered elsewhere. If the extremum point is close to the boundary, or outside the design region, our approach can also be applied. However, in this case, Theorem 1 does not give a locally optimal design. This also will be the subject of further study.

5. APPENDIX

Proof of Theorem 1

Introduce the notation $f^{(1)}(x) = f^{(1)}(x, b)$, $f^{(2)}(x) = f^{(2)}(x, b)$,

$$\begin{aligned} f^{(1)}(x,b) &= (b_1 - x_1, \ldots, b_k - x_k)^T, \\ f^{(2)}(x,b) &= \left((x_1 - b_1)^2, \ldots, (x_k - b_k)^2\right., \\ &\quad \left.(x_1 - b_1)(x_2 - b_2), \ldots, (x_{k-1} - b_{k-1})(x_k - b_k), 1\right)^T, \\ d_s(x, \xi, X) &= \left(f^{(1)}(x) - X^T f^{(2)}(x)\right)^T M_s^{-1}(\xi) \left(f^{(1)}(x) - X^T f^{(2)}(x)\right). \end{aligned}$$

Note that for a fixed $b = b^{(0)}$, the matrix $\bar{M}(\xi)$ coincides with the information matrix for the linear regression function $\bar{\theta}^T f(x, b)$. Therefore a locally optimal design for estimation of an extremum point of the regression function (2.1) is a truncated D-optimal design for estimation of the first k parameters of the regression function $\bar{\theta}^T f(x, b)$ and *vice versa*. We can apply the corre-

sponding equivalence theorem from Karlin & Studden, 1966, §10.8, which we reformulate in a form convenient for our purpose:

Lemma 5.1. *The following definitions are equivalent:*
1) a design $\tilde{\xi}$ maximizes $\det M_s(\xi)$,
2) there exists a matrix X such that

$$M_3(\tilde{\xi})X = M_2(\tilde{\xi}), \quad \max_x d_s(x, \tilde{\xi}, X) = k.$$

Besides, if one of these conditions is fulfilled then $d_s(\tilde{x}_{(i)}, \tilde{\xi}, X) = k$, $(i = 1, \ldots, n)$ where $\tilde{x}_{(i)}$, $i = 1, \ldots, n$ are the points of the design $\tilde{\xi}$.

Consider now the design ξ^*. Due to the symmetry properties of the hypercube, we can restrict consideration to the case $b \in [0,1]^k$. It is easy to verify that

$$\begin{aligned} M_1(\xi^*) &= \operatorname{diag}\{(1-b_1)^2, \ldots, (1-b_k)^2\},\ M_2 = 0,\ M_s = M_1, \\ M_3 &= M_3(\xi^*) = \begin{pmatrix} G_1^T G_1 & 0 & G_1^T \\ 0 & G_2 & 0 \\ G_1 & 0 & 1 \end{pmatrix}, \end{aligned}$$

where

$$\begin{aligned} G_1 &= ((1-b_1)^2, \ldots, (1-b_k)^2) \text{ is a row vector, and} \\ G_2 &= 4\operatorname{diag}\{(1-b_1)^2(1-b_2)^2, \ldots, (1-b_{k-1})^2(1-b_k)^2\} \end{aligned}$$

is a diagonal matrix of size $k(k-1)/2 \times k(k-1)/2$.

Let

$$X^T = \left(\operatorname{diag}\left\{ \frac{1}{2(1-b_1)}, \ldots, \frac{1}{2(1-b_k)} \right\}, O, \left(-\frac{1-b_1}{2}, \ldots - \frac{1-b_k}{2} \right)^T \right),$$

where O is the zero matrix of size $k \times k(k-1)/2$. Then

$$M_3 X = M_2 = 0 \tag{2.6}$$

and

$$d_s(x, \xi^*, X) = \sum_{i=1}^{k} s_i^2(x_i),$$

where

$$s_i(x_i) = \frac{(x_i + 1 - 2b_i)^2 - 2(1-b_i)^2}{2(1-b_i)^2}.$$

The function $s_i(x_i)$ is a quadratic polynomial and $s_i'(x_i) = 0$ with $x_i = 2b_i - 1$. Since for $b_i \in [0, 1/2]$ we have

$$s_i^2(1) = 1,\ s_i^2(2b_i - 1) = 1,\ s_i^2(-1) = \left(1 - \frac{2b_i}{(1-b_i)^2}\right)^2 \le 1,$$

then $\max_{x_i\in[-1,1]} s_i^2(x_i) = 1$ and $\max_{x\in[-1,1]^k} d(x,\xi^*,X) = k$. Therefore, from Lemma 5.1, the design ξ^* is a locally optimal design under the condition $b \in [0,1/2]^k$.

Now let $1 \geq b_i > 1/2$ for some $i = i^*$. Suppose that ξ^* is a locally optimal design. An arbitrary solution of equation (2.6) has the form

$$X^T = (W^T \vdots O \vdots V),$$

where $W = (w_{ij})_{i,j=1}^k$ is an arbitrary matrix, O is the zero matrix of size $k \times k(k-1)/2$ and

$$V = \left(-\sum_{i=1}^k w_{i1}(1-b_i)^2, \ldots, -\sum_{i=1}^k w_{ik}(1-b_i)^2 \right)^T.$$

Therefore the function $d_s(x,\xi^*,X)$ is of the form

$$\sum_{i=1}^k \left\{ \left(b_i - x_i - \sum_{j=1}^k w_{ji} \left[(x_j - b_j)^2 - (1-b_j)^2 \right] \right)^2 / (1-b_i)^2 \right\}.$$

For an arbitrary $j = 1,\ldots,k$, consider the points $x_{(1)} = (2b_1 - 1, \ldots, 2b_k - 1)$ and $x_{(2)} = (2b_1 - 1, \ldots, 2b_{j-1} - 1, 1, 2b_{j+1} - 1, \ldots, 2b_k - 1)$. From Lemma 5.1 $d_s(x,\xi^*,X)'_{x_i} = 0$ for $x = x_{(1)}$ and arbitrary $i = 1,\ldots,k$ and for $x = x_{(2)}$ and $i \neq j$. From this it follows that $w_{ij} = 0$ for $i \neq j$ and $w_{ii} = 1/(2(1-b_i))$. Therefore, $d_s(x,\xi^*) > k$ at the point $\bar{x} = (\bar{x}_1,\ldots,\bar{x}_k)^T$, where $\bar{x}_i = -|b_i|/b_i$, $i = i^*$ and $\bar{x}_i = 2b_i - 1$ otherwise. This contradicts our supposition. Hence, in this case, ξ^* is not a locally optimal design. ■

Proof of Theorem 2

It was shown in the proof of Theorem 1 that the conditions

$$\begin{array}{rcl} M_1(\xi) & = & \mathrm{diag}\{(1-|b_1|)^2,\ldots,(1-|b_k|)^2\}, \\ M_2(\xi) & = & O \end{array} \tag{2.7}$$

are sufficient for local optimality of the design ξ if $b \in [-1/2,1/2]^k$. Now let $\alpha = 1,\ldots,\nu, j_1,\ldots,j_\nu \in \{1,\ldots,\nu\}$, $n = 2^\nu$. Introduce vectors $w_{(j_1,\ldots,j_\alpha)} \in \mathcal{R}^n$ by the formula

$$w_{(j_1,\ldots,j_\alpha)i} = z_{(i)j_1} \ldots z_{(i)j_\alpha}, \;\; i = 1,\ldots,n.$$

It is easy to check that all these vectors are orthogonal to the vector $(1,\ldots,1)$. Therefore,

$$\sum_{i=1}^n z_{(i)j_1} \ldots z_{(i)j_\alpha} = 0, \;\; \alpha = 1,\ldots,\nu,$$

if $\alpha = 1$ or $\alpha \geq 2$ and at least two of the numbers $j_1,\ldots,j_\alpha$ are distinct.

From these relationships we can verify by an immediate calculation that conditions (2.7) are satisfied for the design $\xi = \xi^+$. ■

References

Box, G.E.P. and Wilson, K.B. (1951). On the experimental attainment of optimum conditions. *J. Royal Statistical Soc.* B **13**, 1–38.

Buonaccorsi, J.P. and Iyer, Y.K. (1986). Optimal designs for ratios of linear combinations in the general linear model. *JSPI* **13**, 345–356.

Chaloner, K. (1989). Optimal Bayesian experimental design for estimation of the turning point of a quadratic regression. *Communications in Statistics, Theory and Methods* **18**, 1385–1400.

Chatterjee, S.K. and Mandal, N.K. (1981). Response surface designs for estimating the optimal point. *Bull. Calcutta Statist. Ass.* **30**, 145–169.

Fedorov, V.V. and Müller, W.G. (1997). Another view on optimal design for estimating the point of extremum in quadratic regression. *Metrika* **46**, 147-157.

Jennrich, R. J. (1969). Asymptotic properties of non-linear least squares estimators. *Ann. Math. Stat.* **40**, 633–643.

Karlin, S. and Studden, W. (1966). *Tchebysheff Systems: with Application in Analysis and Statistics*. New York: Wiley.

Mandal, N.K. and Heiligers, B. (1992). Minimax designs for estimating the optimum point in a quadratic response surface. *JSPI* **31**, 235–244.

Mandal, N.K. (1978). On estimation of the maximal point of single factor quadratic response function. *Bull. Calcutta Statist. Assoc.* **27**, 119–125.

Müller, W.G. and Pötscher, B.M. (1992). Batch sequential design for a nonlinear estimation problem. In *Model-Oriented Data Analysis 2* Eds V.V. Fedorov, W.G. Müller and I. Vuchkov, pp. 77–87. Heidelberg: Physica-Verlag.

Müller, Ch.H. (1995). Maximin efficient designs for estimating nonlinear aspects in linear models. *JSPI* **44**, 117–132.

Müller, Ch.H. and Pazman, A. (1998). Applications of necessary and sufficient conditions for maximin efficient designs. *Metrika* **48**, 1–19.

Pronzato, L. and Walter, E. (1993). Experimental design for estimating the optimum point in a response surface. *Acta Applic. Mathemat.*, **33**, 45–68.

Chapter 3

ON REGRESSION EXPERIMENT DESIGN IN THE PRESENCE OF SYSTEMATIC ERROR

S.M. Ermakov
Department of Statistical Modelling,
St Petersburg State University,
Bibliotechnaya Sq.2, St Petersburg, 198904, Russia
Sergej.Ermakov@pobox.spbu.ru

Abstract Different approaches to experimental design in the presence of systematic error are considered. Randomisation of designs allows us to study the problems from a unified viewpoint. Some new results concerning random replication in the linear regression model are elucidated.

Keywords: regression, linear model, randomisation, systematic error

Introduction

Consider the classical formulation of the regression problem. Suppose that a random function $\zeta(x)$, $x \in \mathcal{X} \subset \mathcal{R}^s$, can be measured (observed) at some given points x_i from $\mathcal{X}$, $i = 1, 2, \ldots, N$, $\zeta_i = \zeta(x_i)$, and the expectation $E(\zeta/x)$ can be represented, exactly or approximately, in the form

$$E(\zeta/x) = f(x) \cong \sum_{j=1}^{m} \theta_j \phi_j(x). \tag{3.1}$$

Here $m \leq N$, θ_j are parameters, with the $\phi_j(x)$ determined and linearly independent on $\mathcal{X}$. We discuss the simplest case $\zeta(x) = f(x) + \varepsilon,\ E\varepsilon(x) \equiv 0$,

$$E\varepsilon_i\varepsilon_k = \begin{cases} \sigma^2 & i = k \qquad \varepsilon_i = \varepsilon(x_i) \\ 0 & i \neq k \qquad i, k = 1, \ldots, N. \end{cases}$$

The experimental design problem in this case has been studied in detail supposing that the equality (3.1) is exact and that the parameters θ_j are estimated by least squares. As is well known, a set of points $x_1, \ldots, x_N$ or a measure

on $\mathcal{X}$ are called an experimental design (exact design or continuous design, respectively).

The problem of optimal experimental design is in essence multicriterial. Usually some compromise criterion is chosen. If the equality (3.1) is not exact then the resulting systematic error can also be reflected in a corresponding compromise criterion.

Suppose estimators $\hat{\theta}_j$ of parameters θ_j can be represented by the Cramer formula

$$\hat{\theta}_j = \frac{\det \|[\phi_1, \phi_l], \ldots, [\phi_{j-1}, \phi_l], [\zeta, \phi_l], [\phi_{j+1}, \phi_l], \ldots, [\phi_m, \phi_l]\|_{l=1}^m}{det\|[\phi_k, \phi_l]\|_{k,l=1}^m}. \quad (3.2)$$

Here, for any given points x_i, $[f, g]$ is

$$[f, g] = \frac{1}{N} \sum_{i=1}^{N} f(x_i) g(x_i).$$

Suppose also that the determinant $\det \|[\phi_k, \phi_l]\|_{k,l=1}^m$ is nonzero. After some transformations we can represent $\hat{\theta}_j$ in the form $\hat{\theta}_j = \sum\limits_{i=1}^{N} \zeta_i \, t_{i,j}(x_1, \ldots, x_N)$ where $t_{i,j}(x_1, \ldots, x_N) =$

$$\frac{\det \|[\phi_1, \phi_l], \ldots, [\phi_{j-1}, \phi_l], \frac{1}{N}\phi_l(x_i), [\phi_{j+1}, \phi_l], \ldots, [\phi_m, \phi_l]\|_{l=1}^m}{\det \|[\phi_i, \phi_j]_{i,j=1}^m}. \quad (3.3)$$

Thus the estimator $\hat{f}(x)$ of the regression function $f(x)$ on $\mathcal{X}$ can be represented as

$$\hat{f}(x) = \sum_{j=1}^{m} \hat{\theta}_j \phi_j(x) \text{ or } \hat{f}(x) = \sum_{i=1}^{N} l_i(x) \zeta_i, \quad (3.4)$$

where

$$l_i(x) = \sum_{j=1}^{m} \phi_j(x) \, t_{i,j}(x_1, \ldots, x_N). \quad (3.5)$$

Note that from (3.4) it follows that

$$\hat{f}(x) = \sum_{i=1}^{N} l_i(x) \, f(x_i) \, + \sum_{l=1}^{N} l_i(x) \varepsilon_i. \quad (3.6)$$

If we let $Q = (x_1, \ldots, x_N)$ we have

$$E(\hat{f}(x)/Q) = \sum_{i=1}^{N} l_i(x) f(x_i) \text{ and } \mathcal{D}(\hat{f}(x)/Q) = \sigma^2 \sum_{i=1}^{N} l_i^2(x).$$

The value

$$R(x, Q) = f(x) - E\hat{f}(x) \tag{3.7}$$

is the systematic error.

In the classical statement of the problem of experimental design one supposes that

$$E\,(\hat{f}(x)/Q) = f(x) \quad \text{i.e. } R(x, Q) \equiv 0. \tag{3.8}$$

An example of an optimality criterion is

$$E\left(\int\limits_{\mathcal{X}} \left(f(x) - \hat{f}(x)\right)^2 dx/Q\right), \tag{3.9}$$

known as I-optimality. The transition from an exact design $(x_1, \ldots, x_N)$ to the corresponding continuous design is performed by $\nu(dx) = \frac{1}{N}\sum\limits_{i=1}^{N} \delta_{x_i}(x)$, where $\delta_{x_i}(x)$, is a δ-measure concentrated at the point x_i. In the general case the continuous design is $\nu(dx) = \sum\limits_{i=1}^{N} p_i \delta_{x_i}(x),\ p_i \geq 0$ or an arbitrary probability measure.

1. SYSTEMATIC ERRORS IN REGRESSION ANALYSIS

As has been noted in many papers, the supposition (3.8) is not realistic for a number of problems in regression analysis. Usually $E(\hat{f}(x)/Q) \neq f(x)$ and

$$E\left(\int\limits_{\mathcal{X}} \left(f(x) - \hat{f}(x)\right)^2 dx/Q\right) = \int\limits_{\mathcal{X}} \left[f(x) - \sum_{i=1}^{N} l_i(x) f(x_i)\right]^2 dx$$

$$+\sigma^2 \int\limits_{\mathcal{X}} \sum_{i=1}^{N} l_i^2(x)\, dx. \tag{3.10}$$

If the right side of (3.10) is chosen as a criterion of optimality for the experimental design, two important problems appear:

1 To estimate each term on the right side of (3.10) by experimentation, i.e. to separate the systematic and random components of error;

2 To construct experimental designs for the minimisation of both systematic and random components on the right side of (3.10). If the relation between the components is unknown this problem involves two criteria.

We consider two approaches to solving these problems.

The first approach is connected with the method of criterion construction. Suppose that f belongs to some class $\mathcal{F}$ of differentiable functions. Since the metric in (3.10) is square then $\mathcal{F}$ can be an example of a Sobolev space. In this case there exists $K(Q)$ such that

$$E\Big(\int_{\mathcal{X}} \left(f(x) - \hat{f}(x)\right)^2 / Q\Big) \leq \|f\|^2 \cdot K^2(x_1, \ldots, x_N) + \sigma^2 \int_{\mathcal{X}} \sum_{i=1}^{N} l_i^2(x), \tag{3.11}$$

with the equality holding for some function (the worst) in this space. The results of numerical calculation of (exact) optimal designs, i.e. designs minimising

$$K^2(x_1, \ldots, x_N) + \left(\frac{\sigma}{\|f\|}\right)^2 \int_{\mathcal{X}} \sum_{i-1}^{N} l_i^2(x) dx \tag{3.12}$$

are published in Ermakov and Makhmudov (1977), for the simplest case of $\mathcal{X} = [0, 1]$ and the supposition that $\int_0^1 \left(f'(x)\right)^2 dx$ has some prescribed value. The calculations were performed for some specified values of the parameters $\sigma/\|f\|$. Unfortunately, the calculation of $K(Q)$ in more general situations is too complicated and does not allow a solution to this problem in more general cases.

The second approach is based on randomisation of the experimental designs. Determine a measure on the set of the exact designs - a common distribution of random points $x_1, \ldots, x_N$. Then the two components of error are random and a procedure similar to the analysis of variance can be constructed for dividing these error components.

Since unbiasedness is the usual requirement for estimators we demand for the distributions of Q that

$$E\hat{f}(x) = f(x) \tag{3.13}$$

for any f from the given class of functions $\mathcal{F}$.

This equality determines, but not completely, the randomisation of the experimental design. The additional condition can be $E\, H(\zeta, Q) = \min$ where H is a given function and minimisation is performed by choosing the distribution of Q (H-optimality).

Designs of such a kind, with a common distribution of $(x_1, \ldots, x_N)$, are called unbiased H-optimal designs. The problem with their construction is that, in the general case, one has to solve an infinite dimensional linear programming problem. The precise statement of the problem, along with a proof of the existence of the solution, can be found in Ermakov (1970).

Since the continuous design is a probability measure, we can give another interpretation of (3.13). Namely, we can replace the expectation $E_Q(E(\hat{f}/Q))$

on the right side of (3.13) by the expectation $E_\nu(E\hat{f}/\nu)$. Then

$$E_\nu(E\hat{f}/\nu) = f(x). \tag{3.14}$$

If a design ν is such that the equality (3.14) holds for any function f from a set $\mathcal{F}$ of functions, then ν is also called an unbiased design (in the set $\mathcal{F}$). (This name is probably inadequate). The problems of construction of such unbiased designs are close to those formulated by Box and Hunter (1959). On the one hand, the theory of unbiased designs in the sense of (3.14) is close to cubature formula theory (Ermakov and Sedunov, 1974). On the other hand, this theory has many common features with the classical theory of continuous designs (Ermakov and Melas, 1995). There are many other results about this topic but our article is constrained in length. Now we discuss instead some new results devoted to the randomisation of designs.

2. EXAMPLES OF UNBIASED EXPERIMENTAL DESIGNS

Let $g(Q)$ be the density of $\mu^n(Q)$, where μ^n is a product measure and μ is a probability measure on $\mathcal{X}$. It is natural to suppose that g is symmetrical with respect to the permutation of arguments $x_1, \dots, x_N$. Then if the equality (3.13) is satisfied for all $f \in L^2$,

$$\int \mu(dx_2) \dots \int \mu(dx_N) \frac{g(x_1, \dots, x_N)}{\det \|[\phi_j, \phi_l]\|_{j,l=1}^m} \times$$

$$\times \det \|[\phi_1, \phi_l] \dots, [\phi_{k-1}, \phi_l], \phi_l(x_1), [\phi_{k+1}, \phi_l], \dots, [\phi_m, \phi_l]\|_{l=1}^m$$

$$= N\phi_k(x_1) \pmod{\mu}, \quad k = 1, \dots, m. \tag{3.15}$$

It can also be shown that there exists a function g, satisfying (3.15) for any μ-integrable ϕ_j, and that such a function is not unique. These statements are confirmed by the next examples.

Example 1. The next result is given by Ermakov and Zolotukhin (1960). The function $g(Q) = C \det \|[\phi_j, \phi_l]\|_{j,l=1}^m$, complies with (3.15). Here C is a normalisation constant. The variance $\hat{f}(x)$ has also been studied in the case $N = m$. The results are used in the problem of variance reduction in the Monte-Carlo method. There are many publications on this problem. For example, a simpler explanation is based on the Binet-Cauchy formula (Ermakov and Schwabe, 1999).

Now consider the following procedure for estimation of $\hat{f}(x)$. Let $\mathcal{X}$ be a given set of N points. At random choose M points $x_{i_1}, \dots, x_{i_M}$, $(M < N)$. The set of points $x_{i_1}, \dots, x_{i_M}$, or a set of numbers $\gamma = (i_1, \dots, i_M)$, will be called a random replicate. We suppose that the probabilities of replicates $p(\gamma)$ are preassigned.

Two different cases are distinguished.

1. Replicates are chosen at random and values ζ_{i_j} are independent both within and between replicates.

2. Randomisation is carried out after receiving all N values of a function ζ (a resampling procedure or a passive experiment).

Let $\phi_i(x)$ be orthonormalised on the set $x_1, \ldots, x_N$ (μ-orthonormalised). Write

$$[\phi_i, \phi_j]_\gamma = \frac{1}{M} \sum_{r=1}^{M} \phi_i(x_{i_r}) \phi_j(x_{i_r})$$

and let $\hat{f}_\gamma(x)$ be the least squares estimator for f replicate γ. Choose

$$p(\gamma) = \frac{\det \|[\phi_i, \phi_j]_\gamma\|}{\sum\limits_{\gamma'} \det \|[\phi_j, \phi_j]_{\gamma'}\|}. \tag{3.16}$$

The next theorem holds.

Theorem 1. *If $p(\gamma)$ is chosen according to (3.16) then*

$$E(\hat{f}_\gamma(x)/\zeta) = \hat{f}(x) \tag{3.17}$$

for both active and passive experiments.

The proof can be performed by checking the unbiasedness condition (3.15) or, when $M = m$, by the Binet-Cauchy formula.

It is important for us to calculate the second moment of $\hat{f}(x)$. For a wide class of functions ϕ_i (D regular functions) the calculation can be performed in the case of $M = m$. The results are reflected in the next theorem (Ermakov and Schwabe, 1999).

Theorem 2 . Let the number of independent replicates γ_k in correspondence with distribution (3.16) be N_1. Then the following equalities hold:

$$Var(\frac{1}{N_1} \sum_{k=1}^{N_1} \theta_l(\gamma_k)) = \frac{1}{N_1}(S^2 + \sigma^2), \tag{3.18}$$

in the case of an active experiment, and

$$Var(\frac{1}{N_1} \sum_{k-1}^{N_1} \hat{\theta}_l(\gamma_k)) = \frac{1}{N_1} S^2 + \frac{N + N_1 - 1}{N\, N_1}\, \sigma^2 \tag{3.19}$$

in the case of resampling. Here $l = 1, 2, \ldots, m$, $\hat{\theta}_l(\gamma_k)$ is the least squares estimator of the l th regression coefficient for replicate γ_k and

$$S^2 = \frac{1}{N} \sum_{i=1}^{N} \left(f(x_i) - \sum_{l=1}^{m} [f, \phi_l] \phi_l(x_i)\right)^2 \tag{3.20}$$

is the systematic component of the variance.

Thus an example of the procedure of division for the systematic and random components of variance under least squares estimators in the linear regression model is shown. The procedure demands a preliminary evaluation of the probabilities $p(\gamma)$. Similar results are given by Ermakov (1975) for active experiments when $\mathcal{X}$ is some domain in R^s.

As was pointed out, see (3.16) there exist various distributions with the property (3.15) of unbiasedness. We can construct an infinite number of such distributions (Ermakov, 1975), but usually it is very difficult to compare them by a simple analysis.

Now we show one more example where the analysis of variance for the estimator is comparatively simple.

Example 2. Distribution for the reduction of systematic error. We have an active experiment, $\mathcal{X} \subset R^s$, with the random replicates $x_1^r, \ldots, x_m^r$ subject to some additional deterministic conditions. Also let $M > m$ and, under these conditions, let

$$[\phi_i, \phi_j]_\gamma = \begin{cases} 1, & i = j \\ 0, & i \neq j, \end{cases} \qquad i, j = 1, \ldots, m. \tag{3.21}$$

Then it is easy to see that for each replicate

$$\hat{\theta}_l^r = \frac{1}{M} \sum_{k=1}^{M} \zeta_k^r \phi_l(x_k^r). \tag{3.22}$$

If the replicates have a uniform distribution on a hypersurface defined by (3.21), supposing that such a hypersurface exists, the estimators (3.22) are evidently unbiased estimators for $\hat{\theta}_l$.

Two particular cases are discussed in what follows.

A. *(Degenerate)*. If we choose, in a deterministic way, a set of points $x_1, \ldots, x_M$ so that the equalities

$$\frac{1}{M} \sum_{i=1}^{M} \phi_i(x_k)\phi_j(x_k) = \int_{\mathcal{X}} \phi_i(x)\phi_j(x)\, dx \;\; i, j = 1, 2, \ldots, m \tag{3.23}$$

hold, then the conditions of unbiasedness are formally satisfied.

In addition the equalities (3.23) mean that $x_k, \ldots, x_M$ are nodes of a cubature formula with equal coefficients which is precise for all functions $\phi_i \cdot \phi_j (i, j = 1, \ldots, m)$. The connection between the cubature formulae and the theory of design for minimising systematic error has been thoroughly studied (Ermakov and Sedunov, 1974); Ermakov and Melas, 1995).

B. There exist a few situations in which we can simulate uniform distributions on the hypersurface (3.21). These situations are connected with the theory of

the random cubature formulae. A corresponding result for the s-dimensional hypercube is known. So for $M = 2^s \cdot 3$, random replicates consist of points equidistributed on the unit sphere and symmetric with respect to the co-ordinate planes (Ermakov, 1970). The case $M = s + 1$ can easily be investigated on the basis of the results of Haber (1968).

If we note that our estimators are invariant with respect to some group of transformations, then we can see that the component of systematic error is decreased when compared with the case $g(Q) = C \det \|[\phi_i, \phi_j]\|_{i,j=1}^m$. This result follows from the general theory of random cubature formulae. If we repeatedly get the values of $\zeta(x)$ at the same points, we can construct the procedure for division of the systematic and random components in the usual way.

References

Box, G.E.P. and Hunter, W.G. (1959). A basis for the selection of a response surface design. *J. American Statist. Assoc.* **54**, 622–654.

Ermakov, S.M. (1970). On the optimal unbiased design of regression experiments. *Trudy Math. Inst., Academy of Sciences* **3**, 252–258 (Russian).

Ermakov, S.M. (1975) *Monte-Carlo Techniques and Related Problems*. Moscow: Nauka (in Russian).

Ermakov, S.M. (1989). Random interpolation in the theory of experimental design. *Comp. Stat.& Data Anal.* **8**, 75–80.

Ermakov, S.M. and Makhmudov, A.A. (1977). On regression designs minimising a systematic error. *Zavodskaya Laboratoria* **7**, 854–858 (in Russian).

Ermakov, S.M. and Melas, V.B. (1995). *Design and Analysis of Simulation Experiments*. Dordrecht: Kluwer Academic Publishers.

Ermakov, S.M. and Schwabe, R. (1999). On randomizing estimators in linear regression models. Preprint Nr.A/5-99 Fachbereich Math. und Informatik, Frei Universitat Berlin, Serie A-Mathematik, 1999.

Ermakov, S.M. and Sedunov, E.V. (1974). On unbiased regression designs in finite dimensional spaces of functions. *Vestnik LGU* **1**, 12–20.

Ermakov, S.M. and Zolotukhin, V.G. (1960). Polynomial approximations and Monte-Carlo method. *Probability Theory and its Applic.* **5**, 473–476.

Haber, S. (1968). A combination of Monte-Carlo and classical methods for evaluating multiple integrals. *Bull. Amer. Math. Soc.* **74**, 683–686.

Chapter 4

GRÖBNER BASIS METHODS IN MIXTURE EXPERIMENTS AND GENERALISATIONS

Beatrice Giglio and Henry P. Wynn
Department of Statistics, University of Warwick, Coventry, UK
B.Giglio@warwick.ac.uk and H.P.Wynn@warwick.ac.uk

Eva Riccomagno
EURANDOM, Eindhoven, NL
E.Riccomagno@tue.nl

Abstract The theory of mixture designs has a considerable history. We address here the important issue of the analysis of an experiment having in mind the algebraic interpretation of the structural restriction $\sum x_i = 1$. We present an approach for rewriting models for mixture experiments, based on constructing homogeneous orthogonal polynomials using Gröbner bases. Examples are given utilising the approach.

Keywords: mixture experiments, Gröbner Bases, orthogonal polynomials, homogeneous polynomials.

Introduction

The use of Gröbner bases (G-bases) in experimental design, initiated by Pistone and Wynn (1996), has now been further investigated in a series of papers: Giglio, Riccomagno and Wynn (2000); Riccomagno, Pistone and Wynn (1998); Riccomagno, Pistone and Wynn (1999) and Riccomagno and Wynn (2000). It will also be the topic of a forthcoming monograph (Pistone, Riccomagno and Wynn, 2000).

In this paper, after first briefly outlining the methodology, we specialise to an important subclass of problems, namely mixture experiments, in which factors

$x_1, \ldots, x_d$ are constrained to the proportions:

$$\sum_{i=1}^{d} x_i = 1,\ x_i \geq 0\ \ (i = 1, \ldots, d).$$

As we shall see, algebraic restrictions of the form $\sum x_i = 1$ fall naturally into the G-basis theory which makes use of computational algebraic geometry, that is, symbolic computations with polynomials. The approach to mixtures generalises naturally to cases where more complex algebraic restrictions are placed in the design. These are here called structural restrictions. A most important issue, which we shall discuss, is the effect on allowable models of such restrictions.

The G-basis theory has at its heart the study of *monomials*. For a list $\alpha = (\alpha_1, \ldots, \alpha_d)$ of non-negative integer exponents we define a monomial as

$$x^\alpha = x_1^{\alpha_1} \ldots x_d^{\alpha_d}.$$

A class (ring) of polynomial models is defined when a list of monomials is given:

$$\eta = \sum_{\alpha \in L} \theta_\alpha x^\alpha. \tag{4.1}$$

Thus if $d = 2$ and the list is $L = \{1, x_1, x_2, x_1 x_2\}$ then

$$\eta = \theta_{00} + \theta_{10} x_1 + \theta_{01} x_2 + \theta_{11} x_1 x_2.$$

An experimental design is a set of values (points) $x = (x_1, \ldots, x_d)$ in R_d

$$D = \{x\}_{x \in D}, \ \text{ or sometimes } \ D = \{x^{(1)}, \ldots, x^{(n)}\}.$$

The G-basis method depends on the initial specification of a monomial ordering τ, described below.

1. THE G-BASIS METHOD

The basic methodology is given in some detail in Pistone and Wynn (1996), Giglio, Riccomagno and Wynn (2000) and Riccomagno (1997). We only sketch it here for completeness.

Consider an experimental design for d factors as a zero-dimensional polynomial *variety*. That is to say, the design points are considered as the zeros of a set of polynomial equations. In the language of algebraic geometry we also discuss the *ideal* of all polynomials with zeros containing the design points. We call this the *design ideal* and indicate it by $I(D)$. Such an ideal is described by a finite basis (see Hilbert's basis theorem in Cox, Little and O'Shea, 1997) that is, a special finite system of polynomial equations whose solution set is

given by the design points. Different systems of equations can have the same design as solution set, but there are different bases of the same ideal. G-bases are particular bases of an ideal depending on a term-ordering.

A term-ordering is a total ordering on the set of all monomials in d indeterminates. Among the properties of a term-ordering we recall the so-called *divisibility condition*: for monomials $x^\alpha, x^\beta, x^\gamma$ if $x^\alpha < x^\beta$ then $x^\alpha x^\gamma < x^\beta x^\gamma$.

In this paper we mainly use the `tdeg` term-ordering. In the `tdeg` ordering $\alpha \prec_{\texttt{tdeg}} \beta$ if and only if $\sum_{i=1}^d \alpha_i < \sum_{i=1}^d \beta_i$ or $\sum_{i=1}^d \alpha_i = \sum_{i=1}^d \beta_i$ and the right-most non-zero entry of $\alpha - \beta$ is negative. Thus the `tdeg` ordering depends on the ordering in which the indeterminates are arranged. We call this ordering *initial ordering*. For example, if the initial ordering of the three factors l, s, c is $l \prec s \prec c$ then the ordered list of monomial terms will be:

$$1 \prec l \prec s \prec c \prec l^2 \prec ls \prec lc \prec s^2 \prec sc \prec c^2 \prec l^3 \prec l^2 s \prec l^2 c \prec \ldots.$$

If, instead, the variables are initially ordered so that $c \prec l \prec s$ then we obtain:

$$1 \prec c \prec l \prec s \prec c^2 \prec cl \prec cs \prec l^2 \prec ls \prec s^2 \prec c^3 \prec c^2 l \prec c^2 s \prec \ldots.$$

Notice that the `tdeg` term-ordering first takes into account the total degree of monomials and then orders according to the initial ordering, giving priority to the monomials with highest power of the leftmost variables. For example, for the initial ordering $l \prec s \prec c$, it is

- $l^2 s^3 \prec l^2 s^3 c$: the total degree of $l^2 s^3$ is less than the total degree of $l^2 s^3 c$
- $l^2 s^3 c \prec l^2 s^2 c^2$: the two monomials have the same total degree, the variable l has the same degree in the two monomials, but the next variable, s, has a higher degree in the first monomial.

For $c \prec l \prec s$ it is instead $c^2 l^2 s^2 \prec c l^2 s^3$, because c comes first in the initial ordering.

Given a term-ordering τ a Gröbner basis for the design ideal with respect to τ, $\mathcal{G} = \langle g_1(x), \ldots, g_m(x) \rangle$, is a special basis, that is the solutions of $g_1(x) = \ldots = g_m(x) = 0$ are the design points in D. The design ideal consists of all polynomials of the form

$$s_1(x) g_1(x) + \ldots + s_m(x) g_m(x)$$

where $s_j(x)$ are polynomials. It follows that any polynomial $p(x)$ can be written in the form

$$p(x) = \sum_{j=1}^{m} s_j(x) g_j(x) + r(x),$$

where the remainder $r(x)$ has a special form and is unique, given the monomial ordering.

x_1	x_2	x_3
1	0	0
0	1	0
0	0	1
1/2	1/2	0
0	1/2	1/2
1/2	0	1/2
1/3	1/3	1/3

Table 4.1 A simple example

We say that the set of all polynomials $r(x)$ is the quotient $K(x)/I(D)$ of the ring $K[x] = K[x_1, \ldots, x_d]$ of all real polynomials in d variables, with respect to the design ideal $I(D)$. The monomials $r(x)$ form a subset Est_τ, which consists precisely of all monomials not divisible by the *leading terms* of $g_1(x), \ldots, g_m(x)$ with respect to the monomial ordering. Moreover Est_τ has exactly n terms where n is the sample size and is unique given D and τ.

In summary, given an ordering τ and a design $D = \{x^{(1)}, \ldots, x^{(n)}\}$, the method allows us to obtain the following:

(i) A list $Est_\tau(D) = \{x^\alpha : \alpha \in L\}$ with

$$\#(L) = \#(D) = n \quad \text{(saturated model)};$$

(ii) A model (4.1) based on L and fully estimable, that is the $n \times n$ X-matrix

$$X = \{x^\alpha\}_{x \in D, \alpha \in L}$$

has full rank n;

(iii) $Est_\tau(D)$ has the order ideal property: $x^\alpha \in Est_\tau(D)$ implies $x^\beta \in Est_\tau(D)$ for all x^β dividing x^α.

We add the following property especially for this paper:

(iv) The method allows the incorporation of structured conditions such as $\sum x_i = 1$.

2. MIXTURES AND HOMOGENISATION

It is informative to begin with a brief example. Consider the simple mixture experiment in 3 factors x_1, x_2, x_3 given in Table 4.1.

For the `tdeg` ordering with initial order $x_1 \prec x_2 \prec x_3$ we obtain the G-basis:

$$\begin{cases} g_1(x) = x_3 + x_2 + x_1 - 1 \\ g_2(x) = x_2x_1^2 + 1/2x_1^3 - 1/2x_2x_1 - 3/4x_1^2 + 1/4x_1 \\ g_3(x) = x_2^2x_1 + 1/2x_1^3 - 1/2x_2x_1 - 3/4x_1^2 + 1/4x_1 \\ g_4(x) = x_2^3 - x_1^3 - 3/2x_2^2 + 3/2x_1^2 + 1/2x_2 - 1/2x_1 \\ g_5(x) = x_1^4 - 11/6x_1^3 + x_1^2 - 1/6x_1. \end{cases}$$

To repeat: setting $g_i(x) = 0$, $i = 1, \ldots, 5$ exactly gives the design. Also, the leading terms are

$$Lt = \{x_3, x_2x_1^2, x_2^2x_1, x_2^3, x_1^4\} \text{ and } Est_{\texttt{tdeg}} = \{1, x_1, x_2, x_1^2, x_2x_1, x_2^2, x_1^3\},$$

that is, Est is given by all the terms that are not divided by the leading terms in Lt. We see immediately that x_3 does not appear in $Est_{\texttt{tdeg}}(D)$. This is essentially because the condition $\sum x_i - 1 = 0$ holds. We consolidate this remark into a Lemma.

Lemma 1. *For a design D satisfying $\sum x_i = 1$, a monomial ordering τ and an initial order $x_1 \prec \ldots \prec x_d$, $Est_\tau(D)$ excludes at least one of $x_1, \ldots, x_d$.*

Proof. Suppose that all of $x_1, \ldots, x_d$ were contained in $Est_\tau(D)$. Then since 1 always appears in $Est_\tau(D)$ we must have $\{1, x_1, \ldots, x_d\} \subseteq Est_\tau(D)$. But any polynomial based on $Est_\tau(D)$ which is zero on the design should be identically zero. However $\sum_{i=1}^d x_i - 1$ is such a polynomial, leading to a contradiction. Thus at least one x_k does not appear in $Est_\tau(D)$ and by the order ideal property (iii) no monomial containing x_k appears. □

The theory of mixture designs has a considerable history. See for example Cornell (1990). We show how the G-basis theory clarifies some points. In the literature two main issues have arisen: (i) how to construct good designs and (ii) how to analyse the experiment having in mind the structural restriction $\sum x_i = 1$. Here we partially address issue (ii).

To aid the discussion we return again to the example. First note that the G-basis theory always yields a non-homogeneous model. In this case the model is:

$$\eta = \theta_0 + \theta_{10}x_1 + \theta_{01}x_2 + \theta_{20}x_1^2 + \theta_{11}x_2x_1 + \theta_{02}x_2^2 + \theta_{03}x_2^3.$$

This model can be homogenised by simply using the condition $\sum x_i = 1$ to bring the order up to the maximal order. This has three effects:

(i) The terms of the model are no longer simple monomials;

(ii) The excluded factor(s) according to Lemma 1 are typically reintroduced;

(iii) All terms are of the same degree.

For the above model we obtain:

$$\begin{aligned}\eta = \; & (\theta_0 + \theta_{10} + \theta_{20})\, x_1^3 + (\theta_0 + \theta_{01} + \theta_{02} + \theta_{03})\, x_2^3 + (\theta_0)\, x_3^3 \\ & + (3\theta_0 + 2\theta_{10} + \theta_{11} + \theta_{20} + \theta_{01})\, x_1^2 x_2 + (3\theta_0 + 2\theta_{10} + \theta_{20})\, x_1^2 x_3 \\ & + (6\theta_0 + 2\theta_{10} + 2\theta_{01} + \theta_{11})\, x_1 x_2 x_3 \\ & + (3\theta_0 + 2\theta_{01} + \theta_{02} + \theta_{11} + \theta_{10})\, x_1 x_2^2 + (3\theta_0 + \theta_{10})\, x_1 x_3^2 \\ & + (3\theta_0 + 2\theta_{01} + \theta_{02})\, x_2^2 x_3 + (3\theta_0 + \theta_{01})\, x_2 x_3^2.\end{aligned}$$

We present an approach for rewriting models for mixture experiments, which are a hybrid between the full homogenisation described above and the simple exclusion of factors according to Lemma 1. We express this as a simple principle: *homogenise within terms of the same degree*. However, we consider this solution not quite sufficient and combine it with the construction of orthogonal polynomials. In summary the method will be based on constructing homogeneous orthogonal polynomials out of the G-basis method.

The monomial ordering τ provides an order for the construction of orthogonal polynomials with respect to the design D. Thus, define X_τ to be the full $n \times n$ X-matrix for $Est_\tau(D)$, being careful to order the columns of X_τ in exactly the τ-order. Construct the Cholesky factorization $X_\tau^T X_\tau = U_\tau^T U_\tau$, where U_τ is upper triangular. Rewrite the model $Y = X_\tau \theta + \epsilon$ as $Y = Z_\tau \phi + \epsilon$, where $Z_\tau = X_\tau U_\tau^{-1}$ and $\phi = U_\tau \theta$. Then $Z_\tau^T Z_\tau = I_{n \times n}$ (identity matrix) and

$$[g_\alpha(x)]_{\alpha \in L} = (U_\tau^T)^{-1} [x_\alpha]_{\alpha \in L}$$

are a list of orthogonal polynomials indexed by α. Namely, for certain coefficients $\{c_{\alpha,\beta}\}$

$$g_\alpha(x) = \sum_{\beta \prec_\tau \alpha} c_{\alpha,\beta} x^\beta$$

and

$$\sum_{x \in D} g_\alpha(x) g_\gamma(x) = \begin{cases} 1 & \alpha = \gamma \\ 0 & \text{otherwise.} \end{cases}$$

Rewriting the model as

$$Y = \sum_{\alpha \in L} \phi_\alpha g_\alpha(x)$$

leads to a full orthogonal analysis:

(i) under the usual regression assumptions the estimates $[\hat{\phi}_\alpha] = U\, [\hat{\theta}_\alpha]$ are uncorrelated;

(ii) the contribution to the regression sum of squares (RegSS) for each $g_\alpha(x)$ is $(\hat{\phi}_\alpha)^2, \alpha \in L$.

The attractive feature of this analysis is that if we choose a so-called "graded" ordering τ, such as `tdeg`, which respects degree, we may decompose the RegSS into contributions from each degree. The underlying algebraic theory is our guarantee of consistency in the construction of the $g_\alpha(x)$.

For the example the orthogonal polynomials are:

$$\left\{\begin{array}{rcl}
g_{(0,0)} & = & 1/2 \\
g_{(1,0)} & = & 1/2 + 1/2\sqrt{11}\, x_1 \\
g_{(0,1)} & = & 1/2 + 3/22\sqrt{11}\, x_1 + 1/11\sqrt{66}\, x_2 \\
g_{(2,0)} & = & 15/16 + 41/176\sqrt{11}\, x_1 + 1/33\sqrt{66}\, x_2 + 1/6\sqrt{6}\, x_1^2 \\
g_{(1,1)} & = & 15/16 + 17/176\sqrt{11}\, x_1 + 17/264\sqrt{66}\, x_2 + 7/96\sqrt{6}\, x_1^2 \\
 & & +1/32\sqrt{42}\, x_1\, x_2 \\
g_{(0,2)} & = & 1/2 + 3/11\sqrt{11}\, x_1 + 5/88\sqrt{66}\, x_2 + 1/32\sqrt{6}\, x_1^2 \\
 & & -1/96\sqrt{42}\, x_1 x_2 + 1/12\sqrt{3}\, x_2^2 \\
g_{(3,0)} & = & 20/27 + 58/297\sqrt{11}\, x_1 + 50/891\sqrt{66}\, x_2 + 37/324\sqrt{6}\, x_1^2 \\
 & & +5/756\sqrt{42}\, x_1 x_2 + 1/27\sqrt{3}\, x_2^2 + 1/189\sqrt{14}\, x_1^3.
\end{array}\right.$$

Each orthogonal polynomial g_α can now be turned into a homogeneous polynomial h_α. For the example the result is:

$$\left\{\begin{array}{rcl}
h_{(0,0)} & = & 1/2 \\
h_{(1,0)} & = & (1/2 + 1/2\sqrt{11})\, x_1 + 1/2 x_2 + 1/2 x_3 \\
h_{(0,1)} & = & (1/2 + 3/22\sqrt{11})\, x_1 + (1/2 + 1/11\sqrt{66})\, x_2 + 1/2 x_3 \\
h_{(2,0)} & = & (15/16 + 41/176\sqrt{11} + 1/6\sqrt{6})\, x_1^2 + (15/16 + 1/33\sqrt{66})\, x_2^2 \\
 & & +15/8\, x_3^2 + (15/8 + 41/176\sqrt{11} + 1/33\sqrt{66})\, x_1 x_2 \\
 & & +(15/8 + 41/176\sqrt{11})\, x_1 x_3 + (15/8 + 1/33\sqrt{66})\, x_2 x_3 \\
 & \vdots & \\
h_{(3,0)} & = & \ldots .
\end{array}\right.$$

Each $h_\alpha(x)$ has the same value on D as the $g_\alpha(x)$ from which it was constructed and the h_α are therefore also orthogonal. Moreover the h_α fall into groups according to degree. The following lemma is useful to control what to expect.

Lemma 2. *For a fixed design D and for every graded ordering the number of terms of given degree k is the same.*

Proof The proof derives from the simple property of the so-called Hilbert function and of the design ideal and is omitted. □

Let us verify the case where we take `tdeg` but with all six permutations of the initial ordering $x_1 \prec x_2 \prec x_3$, $x_1 \prec x_3 \prec x_2$, etc. Since the design D is symmetric we would obtain the same Est_τ for any permutation of the initial

x_1	x_2	x_3
1	0	0
0	1	0
0	0	1
1/2	1/2	0
1/2	0	1/2
1/3	1/3	1/3

Table 4.2 Design D'

ordering. Thus, to make this more interesting, let us destroy the symmetry of the problem and consider the design D' where the point $(0, 1/2, 1/2)$ has been removed (see Table 4.2).

Table 4.3 gives all the $Est_\tau(D')$ and it is seen that the number of terms of degree 1, 2 and 3 is the same as is expected from Lemma 2, namely the constant term plus two second degree terms and three third degree terms. Each initial ordering potentially leads to different sets of h_α.

Now, using the method, write the new homogenised orthogonal polynomial model as

$$\eta = \sum_{\alpha \in L} \phi_\alpha h_\alpha(x) = \phi_0 + \sum_{\sum \alpha_i = 1} \phi_\alpha h_\alpha + \sum_{\sum \alpha_i = 2} \phi_\alpha h_\alpha + \dots .$$

The claim is that interpretation of the analysis for mixture experiments is considerably clarified. Here is a summary of features:

(i) An orthogonal analysis of $\hat{\phi}_\alpha$ together with appropriate plots;

(ii) An analysis of variance based on degree;

(iii) A clear definition of an "effect".

We carry out the analysis for a real case study in the next section. Here, we wish to elaborate briefly on point (iii) above.

A conundrum of the idea of effect in a mixture experiment arises from the possible incompatibility between the idea of an effect as simply a parameter such as ϕ_α (or its estimate) and the more physically motivated definition as a change in the model η (or its estimate) as a factor (or factors) are varied in the design region. For mixture experiments the second definition should take into account the fact that the design region is $\sum x_i = 1,\ x_i > 0,\ i = 1, \dots, d$.

Consider, then, a direction vector s on the plane $\sum x_i = 1$:

$$s = (s_1, \dots, s_d),\ \sum s_i = 0,\ \sum s_i^2 = 1,$$

so that, at some base point $x^{(0)}$, the perturbation $x^{(0)} + \epsilon s\ (\epsilon > 0)$ remains on $\sum x_i = 1$. Then the natural interpretation of the local effects of a factor x_j are

Initial ordering	$Est_{\texttt{tdeg}}(D')$
$x_1 \prec x_2 \prec x_3$	$1, x_1, x_2, x_1^2, x_2x_1, x_2^2$
$x_1 \prec x_3 \prec x_2$	$1, x_1, x_3, x_1^2, x_3x_1, x_3^2$
$x_2 \prec x_3 \prec x_1$	$1, x_2, x_3, x_2^2, x_3x_2, x_3^2$
$x_2 \prec x_1 \prec x_3$	$1, x_2, x_1, x_2^2, x_1x_2, x_1^2$
$x_3 \prec x_2 \prec x_1$	$1, x_3, x_2, x_3^2, x_2x_3, x_2^2$
$x_3 \prec x_1 \prec x_2$	$1, x_3, x_1, x_3^2, x_3x_1, x_1^2$

Table 4.3 Example: Est for D' for the the six initial orderings

given by selection of s, e.g.

$$\|s\| = \left(\frac{1}{n-1}, \ldots, 1, \ldots, \frac{1}{n-1}\right)$$

(1 is in entry j), the local effect at $x^{(0)}$ for deviation s

$$c_s = \sum_j \frac{\partial \eta}{\partial x_i} s_j \Bigg|_{x=x^{(0)}} = \sum_j \left(\sum_{\sum \alpha_i = 1} \phi_\alpha \frac{\partial h_\alpha}{\partial x_i} + \sum_{\sum \alpha_i = 2} \phi_\alpha \frac{\partial h_\alpha}{\partial x_i} + \ldots \right) s_j \Bigg|_{x=x^{(0)}}$$

and, under standard conditions,

$$c_s = \sum_j \left(\sum_{\sum \alpha_i = 1} \hat{\phi}_\alpha \frac{\partial h_\alpha}{\partial x_i} + \sum_{\sum \alpha_i = 2} \hat{\phi}_\alpha \frac{\partial h_\alpha}{\partial x_i} + \ldots \right) s_j \Bigg|_{x=x^{(0)}}.$$

Tests for $c_\alpha = 0$ can easily be constructed.

3. A CASE STUDY

Table 4.4 represents the experimental design table for a mixture experiment with factors a, b, c, d, e, f, g, h and single output Y. It is not the role of this paper to discuss the nature or quality of the design. The background is that a, b, c, d, e, f, g, h are proportions of materials in an experiment on the design of composite materials and Y is one performance characteristic when the composite is tested in a laboratory as a candidate for inclusion in the final product.

Several strategies to select the initial ordering are suggested in Giglio, Riccomagno and Wynn (2000) and Bates, Giglio, Riccomagno and Wynn (1998). In this case study a combination of an initial linear screening of the variables and the judgement of an expert suggested the following initial ordering:

$$h \prec g \prec f \prec e \prec d \prec c \prec b \prec a.$$

a	b	c	d	e	f	g	h	Y
0.04	0.45	0.02	0.04	0.00	0.10	0.12	0.23	0.448
0.04	0.19	0.02	0.12	0.12	0.10	0.12	0.29	0.198
0.12	0.25	0.02	0.12	0.12	0.02	0.12	0.23	0.318
0.12	0.27	0.02	0.04	0.12	0.10	0.04	0.29	0.298
0.12	0.16	0.10	0.12	0.00	0.10	0.12	0.28	0.162
0.12	0.39	0.02	0.12	0.00	0.02	0.04	0.29	0.240
0.04	0.31	0.10	0.04	0.00	0.10	0.12	0.29	0.250
0.12	0.17	0.10	0.12	0.12	0.10	0.04	0.23	0.313
0.10	0.29	0.02	0.12	0.00	0.10	0.12	0.25	0.340
0.10	0.37	0.10	0.04	0.00	0.02	0.12	0.25	0.330
0.04	0.47	0.02	0.12	0.00	0.02	0.04	0.29	0.330
0.04	0.19	0.10	0.12	0.12	0.02	0.12	0.29	0.238
0.04	0.45	0.10	0.12	0.00	0.02	0.04	0.23	0.315
0.12	0.27	0.10	0.04	0.12	0.02	0.04	0.29	0.300
0.04	0.48	0.02	0.05	0.12	0.02	0.04	0.23	0.375
0.04	0.25	0.10	0.12	0.12	0.10	0.04	0.23	0.313
0.04	0.27	0.10	0.04	0.12	0.02	0.12	0.29	0.255
0.08	0.32	0.06	0.08	0.06	0.06	0.08	0.26	0.295

Table 4.4 A case study: design and response Y

Constant			h	e	c	g	d	f	b
0.361			0.134	0.016	0.010	0.008	0.003	0.002	0.001
eg	ch	g^2	h^2	fg	gh	fh	bh	eh	dh
0.007	0.006	0.001	0.001	0.001	0.000	0.000	0.000	0.000	0.000

Constant	Total Linear	Total Quadratic	Total Regression SS
0.361	0.174	0.016	0.551

Table 4.5 Regression sums of squares: constant, linear and quadratic term contributions

This gives the 18 terms:

$$Est_{\texttt{tdeg}} = \{1, h, g, f, e, d, c, b, h^2, gh, fh, eh, dh, ch, bh, g^2, fg, eg\}.$$

Table 4.5 shows the contribution of each term to the regression sum of squares after an orthogonal analysis has been performed and the total contribution of the constant, the linear and the quadratic terms.

4. GENERALISATIONS

As mentioned, we consider generalisation of the mixture case in which it is known that the design points lie on some algebraic variety defined by equations $u_j(x) = 0$, $j = 1, \ldots, m$. This implies that the corresponding ideal is contained in the design ideal $I_u = \langle u_j(x),\ j = 1, \ldots, m \rangle \subseteq I(D)$. Using a proof very similar to that of Lemma 2 we can show the following result, which helps to predict how the structural restrictions, sometimes called "counfounding relations", affect $Est_\tau(D)$.

Lemma 3. *Given a design D, a monomial ordering τ and $Est_\tau(D) = \{x^\alpha, \alpha \in L\}$, it is not possible for any polynomial $r(x) \in I(D)$ to take the form $r(x) = \sum_{\alpha \in L} \theta_\alpha x^\alpha$ with at least one non-zero θ_α. Equivalently, for at least one θ_α for which $\theta_\alpha \neq 0$, the monomial x^α cannot be in $Est_\tau(D)$.*

Corollary 1. *Given an experimental design D satisfying a structural condition of the form $x^\beta = c_\beta$, it is not possible for $x^\beta \in Est_\tau(D)$ for any nominal ordering τ or any x^γ with $\gamma \geq \beta$.*

This corollary confirms the intuition from classical experimental design that side conditions used to construct the design may not appear as model terms. The constraint $x^\beta = c_\beta$ is sometimes described by saying x^β is "confounded" or "aliased with the constant term". Here is a less standard example:
Example For a two-level full factorial design on $\{-1, 1\}$ consider $d = 4$ and the side condition $x_1x_2 + x_3x_4 = 0$, leading to the design

$$\begin{array}{rrrr} 1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 \\ -1 & 1 & 1 & 1 \\ -1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{array}$$

Having chosen the `tdeg` ordering, the G-basis is

$$x_4^2 - 1,\ x_3^2 - 1,\ x_2^2 - 1,\ x_1^2 - 1, x_2x_3 + x_1x_4,\ x_1x_3 + x_2x_4,\ x_1x_2 + x_3x_4.$$

$Est_{\texttt{tdeg}}$ is $Est_{\texttt{tdeg}} = \{1, x_4, x_3, x_2, x_1, x_3x_4, x_2x_4, x_1x_4\}$. Lemma 3 is verified by noting that $Est_{\texttt{tdeg}}$ does not contain x_1x_2.

References

Bates, R., Giglio, B., Riccomagno, E. and Wynn, H.P. (1998). Gröbner basis methods in polynomial modelling. In *COMPSTAT98: Proceedings in Computational Statistics* Eds. R.L. Payne and P.J. Green, pp. 179–184. Heidelberg: Physica-Verlag.

Caboara, M., Pistone, G., Riccomagno, E. and Wynn, H.P. (1997). The fan of an experimental design. *SCU Report Number 10*, University of Warwick.

Cornell, J. (1990). *Experiments with Mixtures*. New York: Wiley.

Cox, D., Little, J. and O'Shea, J. (1997). *Ideals, Varieties and Algorithms, 2nd Edn*. New York: Springer–Verlag.

Giglio, B., Riccomagno, E. and Wynn, H.P. (2000). Gröbner basis strategies in regression. *Applied Statistics* (to appear).

Draper, N. R. and Smith, H. (1981). *Applied Regression Analysis, 2nd Edn*. New York: Wiley.

Holliday, T., Pistone, G., Riccomagno, E. and Wynn, H.P. (1999). The application of computational geometry to the design and analysis of experiments: a case study. *Computational Statistics* (in press).

Pistone, G., Riccomagno, E. and Wynn, H.P. (2000). *Algebraic Statistics*. London: Chapman and Hall, CRC (in preparation).

Pistone, G. and Wynn, H.P. (1996). Generalised confounding with Gröbner Bases. *Biometrika* **83**, 653–666.

Riccomagno, E. (1997). *Algebraic Identifiability in Experimental Design and Related Ttopics*. PhD thesis, University of Warwick.

Riccomagno, E., Pistone, G. and Wynn, H.P. (1998). Gröbner bases and factorisation in discrete probability and Bayes. *Statistics and Computing* (in press). (Special issue for the workshop "Statistics and Computing", Montreal, 1998).

Riccomagno, E., Pistone, G. and Wynn, H.P. (1999). Polynomial encoding of discrete probability. *ISI Helsinki, 1999* Abstract. Submitted to *Int. Stat. Review*.

Riccomagno, E. and Wynn, H.P. (2000). Gröbner bases, ideals and a generalised divided difference formula. *Num. Math.* (submitted).

Robbiano, L. (1985). Term orderings on the polynomial rings. *EUROCAL'85* LNC Series **204**, 513–517.

L. Robbiano (1998). Gröbner bases and Statistics In *Proceedings of the International Conference "33 years of Gröbner Bases"*. Cambridge University Press, London Mathematical Society Lecture Notes (in press).

Chapter 5

EFFICIENT DESIGNS FOR PAIRED COMPARISONS WITH A POLYNOMIAL FACTOR

H. Großmann and H. Holling
Psychologisches Institut IV, Westfälische Wilhelms-Universität,
Fliednerstr. 21, D-48149 Münster, Germany
grossman@psy.uni-muenster.de and holling@psy.uni-muenster.de

U. Graßhoff
Institut für Mathematik I, Freie Universität Berlin,
Arnimallee 2-6, D-14195 Berlin, Germany
grasshoff@math.fu-berlin.de

R. Schwabe
Institut für Medizinische Biometrie, Eberhard-Karls-Universität,
Westbahnhofstr. 55, D-72070 Tübingen, Germany
rainer.schwabe@uni-tuebingen.de

Abstract In psychological research paired comparisons, which demand judges to evaluate the trade-off between two alternatives, have been shown to yield valid estimates of the judges' preferences. For this situation we present optimal and efficient designs in a response surface setting where the alternatives are modelled by a polynomial.

Keywords: paired comparisons, optimal design, equivalence theorem, symmetrization, polynomial regression, additive model, marginal design

Introduction

In psychological research paired comparisons, which demand judges to evaluate the trade-off between two alternatives, have been shown to yield valid estimates of the judges' preferences (Bateson, Reibstein and Boulding, 1987). In addition, over the past decades a sound measurement theoretical framework for

paired comparisons has been established by mathematical psychologists (David, 1963). As a result of these developments, paired comparisons have been widely adopted as a method for data collection in the social sciences and especially in marketing research. Recently, psychometric methods and, in particular, ideas of experimental design have attracted growing interest in psychological research (McClelland (1997)).

In this note, we present optimal and efficient designs for paired comparisons in a response surface setting where the alternatives are modeled by a linear, quadratic or cubic polynomial. These components can be readily used as bricks for constructing efficient designs for more than one active effect.

1. GENERAL SETTING

As in the monograph by van Berkum (1987) we confine ourselves to the solution of the design problem for the approximating linear model. To be more specific we consider the general linear model

$$\mu(t) = \beta_0 + f(t)^\top \beta$$

with an explicit constant term β_0 for modeling the value of an alternative $t \in \mathcal{T}$. Here f is a vector of known regression functions and β is the vector of parameters of interest. In contrast to standard design problems the values for the alternatives are not directly observable. Only observations $Y(s,t)$ are available for comparing pairs of alternatives s and t. Nevertheless, the observations are still properly described by a linear model

$$Y(s,t) = (f(s) - f(t))^\top \beta + Z(s,t)$$

with settings $x = (s,t)$ for the paired comparisons chosen from the design region $\mathcal{X} = \mathcal{T} \times \mathcal{T}$. Note that for paired comparisons the regression function is given by $x = (s,t) \to F(s,t) := f(s) - f(t)$. Note also that it is fashionable in the literature to model the value $\mu(t)$ without a constant term β_0 (cf. van Berkum (1987)) which results in the same model for the paired comparisons but which might be misleading for finding optimal designs. Additionally, different experiments are assumed to result in uncorrelated and homoscedastic errors Z.

Due to the general complexity of the problem we mainly consider generalized designs ξ which are defined as finitely supported probability measures on the design region $\mathcal{X}$. For the definition and properties of generalized (approximate) designs we refer to the seminal paper by Kiefer (1959) and the monograph by Fedorov (1972). However, we will also point out how certain exact designs can be constructed which are highly efficient.

The quality of a design ξ is reflected by its information matrix $M(\xi)$ which is essentially proportional to the inverse of the covariance matrix for the best linear unbiased estimator of the parameter vector β. As a performance measure

we use the most popular criterion of D–optimality which aims at maximizing the determinant of the information matrix $M(\xi)$. In the present setting the information matrix equals $\sum_{(s,t)}(f(s)-f(t))(f(s)-f(t))^\top \xi(s,t)$ where $(f(s)-f(t))(f(s)-f(t))^\top$ is the information contained in the paired comparison of s and t and $\xi(s,t)$ is the weight associated with it.

Let us note the obvious fact that there is zero information in the pairs (t,t), i. e. in the comparison of t with itself. Moreover, it is also clear that the information for the pair (s,t) is the same as for (t,s), the pair with the internal order reversed. However, it is worthwhile to distinguish between the comparison of s with t and the comparison of t with s in view of the last section on models with more than one factor of influence - although those comparisons seem to be indistinguishable. As a side–effect we can always replace a design ξ by its symmetrized version $\bar{\xi}$ with respect to exchanging s and t and get the same information matrix, $M(\bar{\xi}) = M(\xi)$, with the additional benefit that $\sum_{(s,t)}(f(s)-f(t))\bar{\xi}(s,t) = 0$. We will take advantage of this in the final section when constructing optimal designs for models where the response depends on more than one factor.

Our next remark is that invariance and equivariance properties are inherited by passing from the model μ for the value to the linear model for observing the paired comparison as long as the constant term is preserved. If the augmented regression function $(1, f^\top)^\top$ is linearly equivariant with respect to a transformation g, i. e. there is a matrix Q_g such that $(1, f(g(t))^\top)^\top = Q_g \cdot (1, f(t)^\top)^\top$ for all $t \in \mathcal{T}$, then the regression function $F : (s,t) \to f(s) - f(t)$ of the paired comparison is also linearly equivariant. The corresponding transformation matrix for F is obtained by deleting the first row and column of Q_g.

In the following we will deal with polynomial response of degree p, i. e. $\mu(t) = \beta_0 + \sum_{i=1}^p t^i\beta_i$ to model paired comparisons

$$Y(s,t) = \sum_{i=1}^{p}(s^i - t^i)\beta_i + Z(s,t)$$

for a possibly continuously varying factor. According to the preceding considerations the following result is straightforward from the equivariance of polynomial regression:

Lemma 1. *If the design ξ^* is D–optimal for paired comparisons with polynomial response of degree p on the design region of the standard unit interval, $\mathcal{T} = [-1,1]$, then the properly transformed design $g_{a,b} \circ \xi^*$ is D–optimal for paired comparisons with polynomial response of degree p on the general interval, $[a,b]$, where $g_{a,b}$ is the linear transformation which maps $[-1,1]^2$ onto $[a,b]^2$.*

Hence, it is sufficient for D–optimality to work with the unit interval $\mathcal{T} = [-1, 1]$. Due to the equivariance with respect to sign changes and due to the quasi–concavity of the D–criterion the following property is also directly inherited:

Lemma 2. *If the design ξ^* is D–optimal for paired comparisons with polynomial response on the unit interval, $\mathcal{T} = [-1, 1]$, then there exists a symmetric D–optimal design $\bar{\xi}^*$ with $\bar{\xi}^*(s, t) = \bar{\xi}^*(-s, -t)$. Moreover, the information matrix $M(\xi^*) = M(\bar{\xi}^*)$ has a chess–board structure, i. e. all entries with odd index sum vanish.*

2. LINEAR REGRESSION

For linear response, $\mu(t) = \beta_0 + t\beta$, i. e. $Y(s, t) = (s - t)\beta + Z(s, t)$, the information matrix reduces to a nonnegative number. It is well–known that this number will be maximized when ξ^* assigns all mass to the largest modulus of the regression function $F(s, t) = s - t$.

Theorem 1. *The design ξ^* is D–optimal for paired comparisons with linear response on the unit interval, $\mathcal{T} = [-1, 1]$, if and only if ξ^* is concentrated on $(1, -1)$ and $(-1, 1)$.*

Hence, the policy for linear response is to compare the extreme settings for the factor. Note that, in particular, the one–point designs concentrated on either the comparison $(1, -1)$ or on $(-1, 1)$ are optimal. The corresponding value for the variance of β equals $\frac{1}{4}$ which coincides with the minimal attainable variance for the slope in the directly observable model with response μ.

3. QUADRATIC REGRESSION

Next we consider the situation where each alternative in the paired comparisons is modelled by a quadratic polynomial, i. e. the response is given by

$$Y(s, t) = (s - t)\beta_1 + (s^2 - t^2)\beta_2 + Z(s, t) \qquad s, t \in [-1, 1].$$

As has been pointed out by Silvey (1972), the possible support points of an optimal design are those points where the boundary of a circumscribing ellipsoid touches the image $\{(s - t, s^2 - t^2); s, t \in [-1, 1]\}$ of the regression function $F(s, t)$. This geometric approach suggests that an optimal design will be of the form

$$\xi_{\tau,w} = \tfrac{1}{4}w(\epsilon_{(1,\tau)} + \epsilon_{(\tau,1)} + \epsilon_{(-1,-\tau)} + \epsilon_{(-\tau,-1)}) + \tfrac{1}{2}(1 - w)(\epsilon_{(1,-1)} + \epsilon_{(-1,1)}),$$

where $\epsilon_{(s,t)}$ denotes the Dirac measure concentrated on (s, t) and the optimal setting τ and weight w have to be determined (see Figure 5.1).

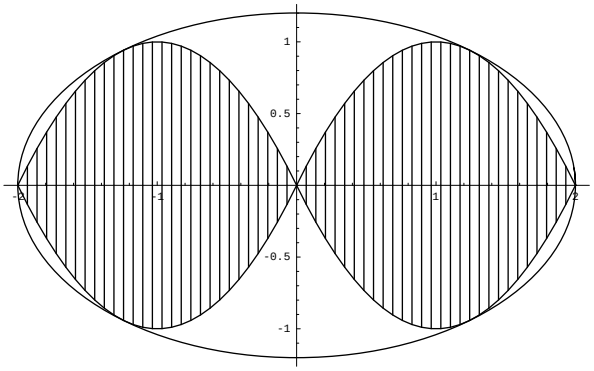


Figure 5.1 The image $\{F(s,t)\}$ with its circumscribing ellipsoid for quadratic response

The equal weights $\frac{1}{4}w$ associated with the support points $(1,\tau)$, $(\tau,1)$, $(-1,-\tau)$, $(-\tau,-1)$ and $\frac{1}{2}(1-w)$ associated with the points $(1,-1)$, $(-1,1)$ are motivated by the natural requirement that the design should be invariant under the reflections at both diagonals of $\mathcal{X}=[-1,1]^2$. The information matrix $M(\xi_{\tau,w})$ for the design $\xi_{\tau,w}$ is diagonal and equals

$$M(\xi_{\tau,w}) = \begin{pmatrix} w(1-\tau)^2+4(1-w) & 0 \\ 0 & w(1-\tau^2)^2 \end{pmatrix}.$$

The determinant of $M(\xi_{\tau,w})$ is given by

$$|M(\xi_{\tau,w})| = w(1-\tau^2)^2(w(1-\tau)^2+4(1-w))$$

which is maximized by the weight

$$w(\tau) = \frac{2}{(3-\tau)(\tau+1)}$$

where the setting τ is fixed. Hence,

$$|M(\xi_{\tau,w(\tau)})| = \frac{4(1-\tau)^2(1+\tau)}{3-\tau}$$

is maximized for the optimal setting $\tau^* = 2-\sqrt{5} \approx -0.236068$. The corresponding optimal weight $w^* = w(\tau^*) = (\sqrt{5}+1)/4 \approx 0.8090$ is closely related to the golden section ratio.

The information matrix of the design ξ_{τ^*,w^*} is given by

$$M(\xi_{\tau^*,w^*}) = \begin{pmatrix} 2 & 0 \\ 0 & 4(5\sqrt{5}-11) \end{pmatrix}$$

resulting in the variance function

$$d(s,t) = (s-t)^2 \left(\frac{1}{2} + \frac{(s+t)^2}{20\sqrt{5} - 44} \right)$$

(see Figure 5.2).

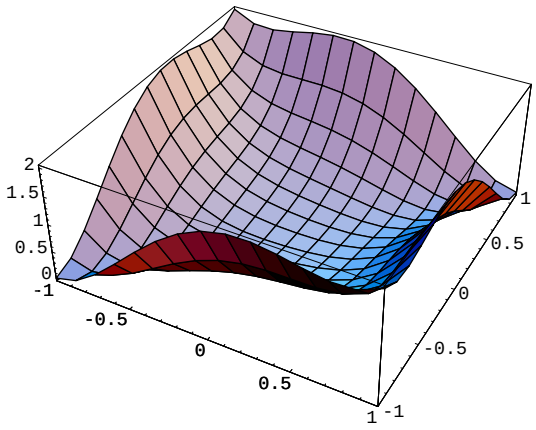


Figure 5.2 Variance function for the D–optimal quadratic design

Now, we want to show that the variance function $d(s,t)$ does not exceed 2 which proves the D–optimality of the design ξ_{τ^*,w^*} by virtue of the celebrated Kiefer–Wolfowitz equivalence theorem (Kiefer and Wolfowitz, 1960). Due to the symmetry of the variance function $d(s,t)$ we can restrict ourselves to the case $s \geq t$. Note that $d(t,t) = 0$.

First we consider the variance function on line segments parallel to the main diagonal, $s = t + c$, $c > 0$. Then $d(t+c,t)$ is a polynomial in t of degree two which attains its maximum on the boundary, $s = 1$ or $t = -1$. Due to the symmetry, again, we have $\sup_{s,t} d(s,t) = \sup_t d(1,t)$. Now, $d(1,t)$ is a polynomial in t of degree four with positive leading coefficient. Moreover, $d(1,1) = 0$, $d(1,-1) = d(1,\tau^*) = 2$ and, additionally, the derivative $\frac{\partial}{\partial t} d(1,t)$ equals 0 for $t = \tau^*$. Thus, it follows that $d(1,t)$ and, hence, $d(s,t)$ is bounded by 2 which proves the D–optimality of the design ξ_{τ^*,w^*}. Moreover, the shape of the variance function $d(s,t)$ establishes that the support of every D–optimal design has to be included in that of ξ_{τ^*,w^*}.

Theorem 2. *The design ξ^* is D–optimal for paired comparisons with quadratic response on the unit interval, $\mathcal{T} = [-1,1]$, if and only if $\xi^*(1,\tau^*) + \xi^*(\tau^*,1) = \xi^*(-\tau^*,-1) + \xi^*(-1,-\tau^*) = \frac{1}{2}w^*$ and $\xi^*(1,-1) + \xi^*(-1,1) = 1 - w^*$ with $\tau^* = 2 - \sqrt{5}$ and $w^* = (\sqrt{5}+1)/4$.*

In particular, the D–optimal three-point design assigns equal weights $\frac{1}{2}w^* \approx 0.4045$ to each of the pairs $(1, \tau^*)$ and $(-\tau^*, -1)$, respectively, where $\tau^* \approx -0.2361$ and the remaining weight $1 - w^* \approx 0.1910$ to the extreme comparison $(1, -1)$.

Remark 1. van Berkum (1987) has obtained numerical solutions for optimal designs in the paired comparison setup with quadratic response when more than one influential factor is present. These solutions are in accordance with our theoretical findings that the extreme settings have to be compared with an intermediate setting exceeding the center point and, additionally, with each other. Note that the optimal settings for the objects under comparison differ from those suitable for the situation of directly observable values. In the latter case the D–optimal design is known to allocate equal weights to the three settings $t = -1$, 0 and 1, respectively.

Remark 2. Atkins and Cheng (1999) considered a related situation where responses are observed directly and coming in pairs with a specified within pair correlation ϱ. For that situation the optimal designs show the same structure as in the present setup of paired comparisons. In fact, it can be seen that by formally letting the correlation ϱ tend to 1, the support and the weights of the optimal design obtained by Atkins and Cheng (1999) tend to the optimal values for paired comparisons given in Theorem 2. This could be explained as follows: in the model with correlated observations coming in pairs, the D–optimal design for the whole parameter vector $(\beta_0, \beta^\top)^\top$ is also D–optimal for β, i. e. for all parameters besides the constant term β_0 (cf. Schwabe (1996), p.17, for the uncorrelated case). If $\widetilde{M}_{\varrho,\beta}(\xi)$ is the information matrix for β in the situation of paired correlated observations then $(1-\varrho^2)\widetilde{M}_{\varrho,\beta}(\xi)$ tends to the information matrix $M(\xi)$ for paired comparisons as ϱ approaches 1.

The design of Theorem 2 can be used as a surrogate for the optimal design in the situation of correlated observations coming in pairs when the correlation is high. Finally, it is worthwhile noting that the within pair correlation is eliminated by passing to paired comparisons.

Example 1. As an alternative to the D–optimal design ξ^* we propose a design ξ' with minimal support of the form $\{(1, \tau), (-\tau, -1)\}$. This is obtained by omitting the extreme comparison from ξ^* which bears the smallest weight. Since the number of support points coincides with the number of unknown parameters the D–optimal weights are uniform, i. e. $\frac{1}{2}$ on each pair. It remains to maximize the determinant $|M(\xi')| = (1-\tau)^4(1+\tau)^2$ of the information matrix. This is achieved for the setting $\tau' = -1/3$ with a resulting determinant $|M(\xi')| = 1024/729 \approx 1.4046$. As for the D–optimal design $|M(\xi^*)| \approx 1.4427$ we can calculate the D–efficiency of the design ξ' according to $\mathrm{eff}_D(\xi') = (|M(\xi')|/|M(\xi^*)|)^{1/2} \geq 98.6\%$ showing that ξ' is highly

efficient. Due to the simple weight structure the design ξ' can be identified as an efficient exact design when the number n of paired comparisons is even. Then $\frac{1}{2}n$ observations are to be taken at each of the comparisons of $s = 1$ with $t = -\frac{1}{3}$ and $s = \frac{1}{3}$ with $t = -1$, respectively.

Example 2. The design ξ'' which assigns equal weight $\frac{1}{3}$ to the comparisons $(1, -1)$, $(1, 0)$ and $(0, -1)$ of the standard settings for directly observable alternatives results in a determinant $|M(\xi'')| = 4/3$ of the information matrix $M(\xi'')$. Hence, ξ'' has a D–efficiency $\text{eff}_D(\xi'') \geq 96\%$. The design ξ'' can be identified as an efficient exact design if the number of comparisons is a multiple of three.

4. CUBIC REGRESSION

For the cubic response

$$Y(s,t) = (s-t)\beta_1 + (s^2 - t^2)\beta_2 + (s^3 - t^3)\beta_3 + Z(s,t)$$

the optimization problem becomes even more complicated. Therefore we start with a characterization of the shape of the variance function $d(s,t) = (f(s) - f(t))^\top M(\xi)^{-1}(f(s) - f(t))$:

Lemma 3. *If the design ξ is symmetric with respect to sign change, $\xi(-s,-t) = \xi(s,t)$, then the variance function $d(s,t)$ attains its maximum at, at most, twelve points in $\mathcal{X} = [-1,1]^2$. These points can be characterized by three settings τ_1, τ_2 and $\tau_3 \in [-1,1]$, $\tau_2 < \tau_3$, according to $(s,t) = (1,-1)$, $(\tau_1, -\tau_1)$, $(1,\tau_2)$, and $(1,\tau_3)$ as well as their corresponding images under reflection with respect to the diagonals.*

Proof. According to Lemma 2 the information matrix $M(\xi)$ has a chess–board structure which is shared by its inverse $M(\xi)^{-1}$. Hence, the variance function can be written as

$$d(s,t) = (s-t)^2(c_1 + c_2(s+t)^2 + c_3(s^2+st+t^2)^2 + c_4(s^2+st+t^2))$$

for some constants $c_1, \ldots, c_4$, where, in particular, $c_3 > 0$. Note that the variance function itself is symmetric with respect to reflections at both diagonals. For notational convenience we only consider those pairs on one side of the main diagonal, $s \geq t$, keeping in mind that the variance function attains its minimum at the main diagonal, $d(t,t) = 0$.

On line segments parallel to the main diagonal, $s = t + c$, $c > 0$, the variance function $d(t+c,t)$ is a polynomial in t of degree four with positive leading term. Hence, due to symmetry, $d(t+c,t)$ may attain its maximum only at the boundary, $s = 1$ or $t = -1$, or at the center of the line segment, $s = -t$. This

consideration reduces the possible locations for the maxima of the variance function to the boundary of the design region and to the secondary diagonal, $s = -t$.

On the secondary diagonal the variance function $d(t, -t)$ is a polynomial in t of degree six with positive leading term. As $d(0, 0) = 0$ and the variance function $d(t, -t)$ is symmetric to 0 it may attain its maximum only at the boundary, $t = 1$, or at, at most, one interior point $t = \tau_1 > 0$.

For the boundary, due to symmetry, we have to investigate only one segment of the boundary, say $s = 1$. On that segment $d(1, t)$ is a polynomial in t of degree six with positive leading term. As $d(1, 1) = 0$ the variance function $d(1, t)$ may attain the maximum only at $t = -1$ and at, at most, two different interior points, $t = \tau_2$ or $t = \tau_3$, $-1 < \tau_2 < \tau_3 < 1$.

Combining these arguments we see that the variance function $d(s, t)$ may attain its maximum, at most, at the points $(1, -1)$, $(\tau_1, -\tau_1)$, $(1, \tau_2)$ and $(1, \tau_3)$ as well as their corresponding images under reflections. □

As an immediate consequence we obtain the shape of the D–optimal design by virtue of the equivalence theorem of Kiefer and Wolfowitz (1960):

Theorem 3. *For paired comparisons with cubic response the D–optimal design ξ^* is supported by, at most, twelve pairs which can be characterized by three settings τ_1, τ_2 and $\tau_3 \in [-1, 1]$, $\tau_2 < \tau_3$, according to $(s, t) = (1, -1)$, $(\tau_1, -\tau_1)$, $(1, \tau_2)$, and $(1, \tau_3)$ as well as their corresponding images under reflection with respect to the diagonals.*

For finding the symmetric D–optimal design it remains to specify six parameters, the three optimal location parameters $\tau_1, \ldots, \tau_3$ and the corresponding optimal weights $w_1, \ldots, w_3$ which are to be uniformly distributed on the support sets $\{(\tau_1, -\tau_1), (-\tau_1, \tau_1)\}$, $\{(1, \tau_2), (-\tau_2, -1), (\tau_2, 1), (-1, -\tau_2)\}$ and $\{(1, \tau_3), (-\tau_3, -1), (\tau_3, 1), (-1, -\tau_3)\}$, respectively, where the remaining weight $1 - w_1 - w_2 - w_3$ is meant to be assigned to $\{(1, -1), (-1, 1)\}$. Again, the number of comparisons may be restricted to those pairs with $s > t$.

Example 3. Using a computer program by E. Rafajłowicz (personal communication) for optimizing the weights on an equidistant grid with step size 0.1 we could identify the pairs $(1, -1)$, $(0.6, -0.6)$, $(1, 0.2)$ and $(-0.2, -1)$ as suitable candidate points for a highly efficient design with four different paired comparisons. In view of the symmetry, this design has the shape required in Theorem 3 with $w_3 = 0$. By a further numerical optimization step using the internal structure of the variance function we obtained the optimal settings $\tau_1 = 0.5984$ and $\tau_2 = 0.1802$. The optimal weights were found to be $w_1 = 0.3189$ for the pair $(\tau_1, -\tau_1)$, $\frac{1}{2}w_2 = 0.2403$ for each of the pairs $(1, \tau_2)$ and $(-\tau_2, -1)$ and the remaining weight $1 - w_1 - w_2 = 0.2005$ on the pair of

extreme comparisons $(1,-1)$. The determinant of the information matrix can be calculated as $|M(\xi)| = 0.121658$. The corresponding variance function is plotted in Figure 5.3 and shows the good performance of the design.

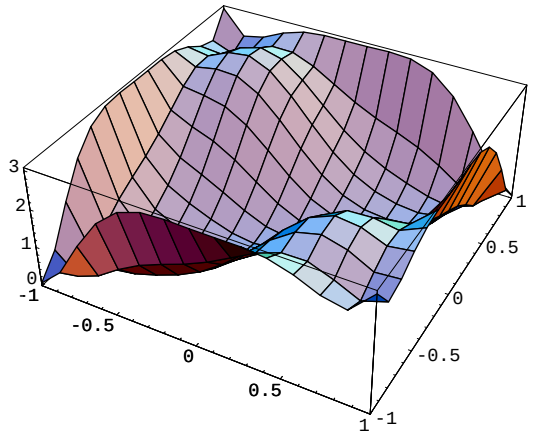


Figure 5.3 Variance function for the four point cubic design of Example 3

A closer inspection of the variance function $d(s,t)$ shows that its numerical maximum is attained at the extreme comparison $(1,-1)$ and exceeds $p = 3$ by less than 0.01%. According to Atwood (1969) a lower bound for the D–efficiency is, hence, given by $\text{eff}_D(\xi) \geq p/\sup_{s,t} d(s,t) \geq 99.99\%$. Thus the D–optimal design seems to be concentrated on four points.

Example 4. As a simple alternative we propose a design ξ_0 with minimal support of the form $\{(\tau_1,-\tau_1),(1,\tau_2),(-\tau_2,-1)\}$. This is, similarly to Example 1, obtained by omitting the extreme comparison which bears the smallest weight from the design of Example 3. Again, due to the support being minimal, the D–optimal weights are uniform, i. e. $\frac{1}{3}$ on each pair. It remains to optimize the determinant $|M(\xi_0)| = \left(\frac{2}{3}\right)^3 (1-\tau_2^2)^2((2\tau_1^2+(1-\tau_2)^2)(2\tau_1^6+(1-\tau_2^3)^2)-(2\tau_1^4+(1-\tau_2)(1-\tau_2^3))^2)$ of the information matrix with respect to the settings τ_1 and τ_2. A numerical solution is given by $\tau_1 = 0.645828$ and $\tau_2 = 0.208013$ with a resulting determinant $|M(\xi_0)| = 0.0987493$. Hence, the minimal support design ξ_0 has 93.2% D–efficiency compared to the four point design of the previous example. Note that the settings τ_1 and τ_2 are very close to the corresponding solutions in Example 3. The design ξ_0 can be used as an efficient exact design when the number of observations is a multiple of 3.

5. GENERATING OPTIMAL DESIGNS FOR ADDITIVE MODELS

Finally, we consider additive two–factor models $\mu(t_1, t_2) = \beta_0 + f_1(t_1)^\top \beta_1 + f_2(t_2)^\top \beta_2$ for the value of an alternative depending on two factors t_1 and t_2 which may be adjusted independently, $t = (t_1, t_2) \in \mathcal{T}_1 \times \mathcal{T}_2$. Thus the observations for the paired comparisons are described by

$$Y(s,t) = (f_1(s_1) - f_1(t_1))^\top \beta_1 + (f_2(s_2) - f_2(t_2))^\top \beta_2 + Z(s,t)$$

and no interactions occur among the components. A result of Rafajłowicz and Myszka (1992) states that D–optimal designs can be constructed as the cross–product of the optimal designs in the corresponding marginal models (see also Schwabe (1996), section 5.2):

Theorem 4. *If the designs ξ_i^* are D–optimal for the marginal models given by the response functions $E(Y^{(i)}(s_i, t_i)) = (f_i(s_i) - f_i(t_i))^\top \beta_i$, $s_i, t_i \in \mathcal{T}_i$, $i = 1, 2$, and if, additionally, $\xi_2^* = \bar{\xi}_2^*$ is symmetric with respect to exchanging s_2 and t_2, then the product design $\xi_1^* \otimes \xi_2^*$ is D–optimal for paired comparisons in the additive model.*

Proof. Due to the orthogonality condition $\sum (f_2(s_2) - f_2(t_2)) \bar{\xi}_2^*(s_2, t_2) = 0$ the result follows immediately from Theorem 5.23 in Schwabe (1996). □

Moreover, if ξ_1 and ξ_2 are competing designs in the marginal models which show high marginal D–efficiencies, $\mathrm{eff}_D(\xi_i) \geq 1 - \varepsilon$, and for which ξ_2 shares the symmetry property of Theorem 4, then the corresponding product design $\xi_1 \otimes \xi_2$ has also D–efficiency of, at least, $1 - \varepsilon$. Hence, the optimal and efficient designs obtained in the previous sections can be used as bricks to build optimal or efficient designs for paired comparisons in additive models with more than one active effect.

Remark 3. Note that the additive structure does not cover the common quadratic response surface model for which D–optimal designs have been obtained by van Berkum (1987). If we use the product of the D–optimal marginals of Theorem 2 in that situation the D–efficiencies of the corresponding product designs equal 88.5%, 76.6%, and 70.2% compared to the D–optimal designs of van Berkum (1987), when the number K of factors is respectively equal to 2, 3 and 4. However, optimal product designs can be obtained numerically which depend on the number of factors and which have substantially higher efficiencies. These findings have to be compared with the corresponding properties in case of directly observable experiments (see Dette and Studden (1997), pp.165-166).

Acknowlegdement. This work was partly supported by the research grant Ho 1286/2-1 of the Deutsche Forschungsgemeinschaft. The authors wish to express their thanks to Professor Ewaryst Rafajłowicz, Wrocław, for supporting them with his computer program for optimizing designs and to an anonymous referee for helpful comments.

References

Atkins, J.E. and Cheng, C.-S. (1999). Optimal regression designs in the presence of random block effects. *J. Statist. Plann. Inference* **77**, 321–335.

Atwood, C.L. (1969). Optimal and efficient designs of experiments. *Ann. Math. Statist.* **40**, 1570–1602.

Bateson, J.E.G., Reibstein, D.J. and Boulding, W. (1987). Conjoint analysis reliability and validity: A framework for future research. In *Review of Marketing* Ed. M.J. Houston, pp. 451–481. Chicago: American Marketing Association.

David, H.A. (1963). *The Method of Paired Comparisons*. London: Griffin.

Dette, H. and Studden, W.J. (1997). *The Theory of Canonical Moments with Applications in Statistics, Probability, and Analysis.* New York: Wiley.

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

Kiefer, J. (1959). Optimal experimental designs. *J. R. Statist. Soc.* B **21**, 272–304.

Kiefer, J. and Wolfowitz, J. (1960). The equivalence of two extremum problems. *Can. J. Math.* **12**, 363–366.

McClelland, G.H. (1997). Optimal design in psychological research. *Psychological Methods* **2**, 3–19.

Rafajłowicz, E. and Myszka, W. (1992). When product type experimental design is optimal? - Brief survey and new results. *Metrika* **39**, 321–333.

Schwabe, R. (1996). *Optimum Designs for Multi–Factor Models. Lecture Notes in Statistics* **113**. New York: Springer-Verlag.

Silvey, S.D. (1972). Discussion of the paper by H.P. Wynn. *J. R. Statist. Soc.* B **34**, 174–175.

van Berkum, E.E.M. (1987). *Optimal Paired Comparison Designs for Factorial Experiments, CWI Tract* **31**.

Chapter 6

ON GENERATING AND CLASSIFYING ALL Q^{N-M} REGULAR DESIGNS FOR SQUARE-FREE Q

Patrick J. Laycock and Peter J. Rowley
Department of Mathematics
UMIST, Sackville St, Manchester M60 1QD
pjlaycock@umist.ac.uk and p.j.rowley@umist.ac.uk

Abstract By extending results developed for q a prime or a prime power we develop methods for generating all q^{n-m} regular designs for q a product of distinct primes, along with their confounded interactions or defining contrasts. The method of generation produces a unique and decodable number for each such design and explicit formulae are supplied for each step of the coding and decoding algorithm. The case where q is a product of two separate primes is studied in detail, with indications given for extensions to more primes and different values of q for each factor, this latter case covering in particular the situation where each q is a single, possibly distinct, prime.

Keywords: Fractional factorials; block designs; complete sub-spaces; algorithms; design numbers; confounded interactions; resolution numbers; aberration

Introduction

This paper is concerned with methods for systematically generating, labelling and classifying all regular versions of q^{-m} th fractions (or corresponding blocking arrangements) of q^n factorial designs, for q a product of distinct primes. That is to say, where the fraction (or principal block) corresponds to a subgroup with a suitable arithmetic. Extensions to arbitrary q and mixed q will be described. The methods are adapted and extended versions of those used in Laycock & Rowley (1995, 1999), which dealt with cases where q is a prime or power of a prime; that is, $q = 2, 3, 4, 5, 7, 8, 9, 11, 13, ...,$ whilst this paper covers $q = 6, 10, 14, 15, ...$ This paper does not however deal with the blocking of fractions.

The justification, analysis and construction of such fractions or blocks utilizing the formal structures of group theory dates back to the origins of the subject

with Finney (1945) and Kempthorne (1947). More recently there are the books by Raktoe *et al* (1981), Street & Street (1987) and Constantine (1987). The number of possible designs or subspaces increases rapidly with n, and the popular methods of generating specific designs are usually 'target specific', such as the 'design keys' of Patterson & Bailey (1978), the generalized cyclic methods of John & Dean (1975) and John & Williams (1995), or the direct search methods of Wynn *et al* (1996).

Our algorithmic procedure supplies a unique numbering scheme for all such designs, and these numbers are easily decoded so as to recover their originating design. This means, for instance, that tables of 'optimal' or 'good' designs, relative to one particular criterion, can be presented as a short list of design numbers.

A regular q^{-m}th fraction of a q^n factorial divides the q^n estimable effects or interactions into q^m alias sets each with q^{n-m} members. It is the sums of the members of these sets which are estimated by the corresponding contrasts of data means in the usual linear model formulation. Alternatively, a regular q^{-m}th blocking of a q^n factorial divides the complete design into q^m sets each with q^{n-m} members. When q is a product of distinct primes, rather than a single prime or power of a single prime, some of the major statistical properties, such as balance, of these regular fractions or blocks are lost or weakened.

An important number used to describe and discriminate amongst regular designs is the *resolution number*, r, which is the minimum number of factors involved in any of the aliases of the grand mean. A regularly blocked design of resolution r will not confound any $(r-2)$th, or lower order, interactions; a regular fractional design of resolution r contains complete factorials in any $r-1$ (or fewer) factors and no f-factor interactions will be aliased whenever r exceeds $2f$.

In this paper we will show how to generate systematically and efficiently all possible blockings or fractional designs in the class q^{n-m}, for given n, $k = n-m$, and q a product of distinct primes; along with their coset representatives, confounded interactions, aliases, resolution numbers, aberration numbers and design generation numbers. In Section 2, the underlying group theoretic context of our methods and their extension to the cases covered by this paper are set out, whilst in Section 3 explicit formulae are supplied for every aspect of the construction methods described. These formulae are provided here for the first time for both q a product of distinct primes - as in this paper, and for q a prime or power of a prime - as in the earlier paper, Laycock & Rowley (1995). In Section 4, examples of designs produced by these methods are presented.

1. OUTLINE OF THEORY AND ALGORITHM

We consider the case $q = p_1p_2$ where p_1 and p_2 are distinct primes ordered so that $p_1 < p_2$. We use $\mathbf{A}$ to denote the direct product of n copies of $\mathbb{Z}_q$, the cyclic group of order q. The theory and algorithm presented here apply equally to the case when $q = p_1p_2...p_h$, where $p_1, p_2, ..., p_h$ are distinct primes and also when we have different levels $q_1, q_2, ..., q_n$ for each factor, provided only that each q_i is the product of distinct primes. For ease of exposition we just describe the case $r = 2$. For $i = 1, 2$, let $\mathbf{A}_i$ be the subgroup of $\mathbf{A}$ generated by the elements of order p_i. Since $\mathbb{Z}_q \cong \mathbb{Z}_{p_1} \times \mathbb{Z}_{p_2}$, we have $\mathbf{A} = \mathbf{A}_1\mathbf{A}_2$ with $\mathbf{A}_i$ isomorphic to a direct product of n copies of $\mathbb{Z}_{p_i}$. Hence we may regard $\mathbf{A}_i$ as an n-dimensional $GF(p_i)-$ vector space. Let $\mathbf{B}$ be a subgroup of $\mathbf{A}$ of order q^k. This will represent the whole fraction (or principal block) of our design. Then $\mathbf{B} = (\mathbf{B} \cap \mathbf{A}_1)(\mathbf{B} \cap \mathbf{A}_2)$ with $|\mathbf{B} \cap \mathbf{A}_i| = p_i^k$. So regarding $\mathbf{A}_i$ as a $GF(p_i)-$ vector space, $\mathbf{B} \cap \mathbf{A}_i$ is a k-dimensional subspace of $\mathbf{A}_i$. Using Laycock & Rowley (1995), we may systematically enumerate all the k-subspaces of the n-dimensional $GF(p_i)$ – vector space $\mathbf{V}_i$, $i = 1, 2$. Then after suitably identifying $\mathbf{V}_i$ with $\mathbf{A}_i$, $i = 1, 2$, we may obtain a complete list of all subgroups of $\mathbf{A}$ of order q^k. Moving onto specific details we define, for $i = 1, 2$, $\phi_i : V_i \longrightarrow \mathbf{A}$ by

$$\phi_i : (a_{i1}, a_{i2}, ..., a_{in}) \longmapsto (p_{3-i}.a_{i1}, p_{3-i}.a_{i2}, ..., p_{3-i}.a_{in}).$$

Here we regard $GF(p_i)$ as $\{0, ..., p_i - 1\}$ $(\subseteq \mathbb{Z})$ with addition and multiplication modulo p_i. It is easily verified that ϕ_i is a group isomorphism, and we shall use ϕ_i to identify $\mathbf{V}_i$ and $\mathbf{A}_i$, $i = 1, 2$. We define an inner product on A with values in $\mathbb{Z}_q$ by

$$(\mathbf{a}, \mathbf{b}) = \sum_{i=1}^{n} a_i b_i \tag{2.1}$$

where $\mathbf{a}, \mathbf{b} \in \mathbf{A}$ with $\mathbf{a} = (a_1, a_2, ..., a_n)$, $\mathbf{b} = (b_1, b_2, ..., b_n)$. (If different levels were present then the inner product would be defined differently - see Wynn *et al* (1996)). For $\mathbf{B}$ a subgroup of $\mathbf{A}$ we define $\mathbf{B}^\perp$ by

$$\mathbf{B}^\perp = \{\mathbf{a} \mid (\mathbf{a}, \mathbf{b}) = 0 \text{ for all } \mathbf{b} \in \mathbf{B}\}. \tag{2.2}$$

Now $\mathbf{B}^\perp$ is a subgroup of $\mathbf{A}$,and for a regular design can be identified with the defining contrasts or confounded interactions of the design, as in Bailey *et al* (1977). Also see the computer generation methods in John & Williams (1995). Following Rowley & Laycock (1995) we define

$$\#(\mathbf{b}) = |\{i | b_i \neq 0\}| \tag{2.3}$$

and for $d \in \mathbb{N}$

$$\omega_t(\mathbf{B}) = |\{\mathbf{b} \in \mathbf{B} | \#(\mathbf{b}) = t\}|\,; \tag{2.4}$$

$$\rho(\mathbf{B}) = \min\{d \in \mathbb{N} | \omega_t(\mathbf{B}) \neq 0\}\,; \text{ and} \tag{2.5}$$

$$\sigma(\mathbf{B}) = |\{\mathbf{x} \in \mathbf{B} | \#(\mathbf{b}) = \rho(\mathbf{B})\}|\,. \tag{2.6}$$

The resolution of $\mathbf{B}$, $r(\mathbf{B})$, is then $\rho(\mathbf{B}^{\perp})$ and the aberration count of $\mathbf{B}$, $s(\mathbf{B})$, is $\sigma(\mathbf{B}^{\perp})$.

Suppose $|\mathbf{B}| = q^k$ with $\mathbf{B} = (\mathbf{B} \cap \mathbf{A}_1)(\mathbf{B} \cap \mathbf{A}_2)$ and let $\mathbf{U}_i$ be the k-subspace of $\mathbf{V}_i$ for which $\phi_i(\mathbf{U}_i) = \mathbf{B} \cap \mathbf{A}_i$ $(i = 1, 2)$. Then it may be readily checked that $\mathbf{B}^{\perp} = \phi_1(\mathbf{U}_1^{\perp})\phi_2(\mathbf{U}_2^{\perp})$. Here $\mathbf{U}_i^{\perp}$ is the subspace of $\mathbf{V}_i$ orthogonal to $\mathbf{U}_i$. The scheme in Laycock & Rowley (1995) allows us quickly to calculate $\mathbf{U}_i^{\perp}$ and hence we may obtain $\mathbf{B}^{\perp}$.

2. FORMULAE FOR THE NUMBERING SCHEME AND ITS DECODING

It is a consequence of the scheme described in Laycock & Rowley (1995), that the unique number, $L = L(q, k)$, for each regular k-dimensional subspace, $U_L = U_L(n, q, k)$ of an n-dimensional $GF(q) -$ vector space (with q a prime or prime power) can be uniquely decomposed into a sum of powers of q plus a remainder. It should be noted that in this notation, the designs $\mathbf{U}_i$ in Section 2 above would become U_{L_i}. We can write

$$L = \sum_{t<f} q^{j_t} + g \tag{3.1}$$

where f determines the count of powers of q, and $1 \leq g \leq q^{j_f}$, where $j_t = \#\Lambda(\omega_t)$ is the number of 'starred' (i, j) pairs defined by the permutation, $\omega_t = (i_1, i_2, ..., i_k, i_{k+1}, ..., i_n)$, $t = 1, 2, ...\binom{n}{k}$. The k leading places in each permutation are ordered integers chosen from $\{1, ..., n\}$, the balance of $(n - k)$ places being occupied by the ordered remainder. These permutations are arranged in lexicographical order. The set of starred pairs, $\Lambda = \Lambda(\omega_t)$ for each permutation, is defined by

$$\left\{(i, j) \mid 1 \leq i < j \leq n,\ \omega_t^{-1} i \in \{k + 1, ..., n\} \text{ and } \omega_t^{-1} j \in \{1, ..., k\}\right\}, \tag{3.2}$$

whilst the permutation numbered f defines a set of square matrices

$$\left\{\beta_i,\ i = 1, .., q^{j_f}\right\}$$

the gth of which contains in its columns a set of generators for the design. For an example, consider the pair of designs $D_1 = U_7(3, 2, 2)$ and $D_2 = U_9(3, 3, 2)$, each having 9 design points for three factors, the first on 2 levels, the second on 3. We will show how to construct this pair of designs and then how to merge them so as to create the design $U_{87}(3, 6, 2)$. We have

$$\omega_1 = (1, 2, 3),\ \text{so } \omega_1^{-1} = (1, 2, 3) \text{ and } \Lambda(\omega_1) = \phi,\ \ \therefore\ j_1 = 0,$$

$$\omega_2 = (1,3,2),\ \text{so}\ \omega_2^{-1} = (1,3,2)\ \text{and}\ \Lambda(\omega_2) = \{(2,3)\}\,,\ \ \therefore j_2 = 1,$$

$$\omega_3 = (2,3,1),\ \text{so}\ \omega_3^{-1} = (3,1,2)\ \text{and}\ \Lambda(\omega_3) = \{(1,2),(1,3)\}\,,\ \ \therefore j_3 = 2.$$

So that

$$f = 3,\ \text{and}\ L = 7 = 2^0 + 2^1 + 4\ \text{for}\ D_1,\ \text{and}\ L = 9 = 3^0 + 3^1 + 5\ \text{for}\ D_2.$$

Each β_g is an identity matrix plus j_f extra 'starred' entries, $\{b_i\}$, above the diagonal in positions given by $\Lambda(\omega_f)$ in its natural order, and these entries run over all possible q^{j_f} ordered selections from $\{0,...,q-1\}$. So that for D_1

$$\{\beta_g\} = \left\{ \begin{pmatrix} 1 & * & * \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\}$$

which can generate a set of $2^2 = 4$ matrices for D_1 (which uses β_4). Alternatively, this can generate a set of $3^2 = 9$ matrices for D_2 (which uses β_5). Here and elsewhere, we have used an adapted form of the algorithm by O'Flaherty & Mackenzie (1982) to form these ordered 'starred' selections. This algorithm can be summarised by two formulae for the general situation with q^N selections. For the selection numbered g, the recursive relation

$$T_{i+1} = T_i \bmod (q^{N-i}),\ \text{with}\ T_1 = g - 1 \tag{3.3}$$

is coupled to

$$b_i = (T_i - T_{i+1})\,/q^{N-i},\ \ i = 1,...,(N-1),\ \text{with}\ b_N = T_N \tag{3.4}$$

and this scheme is easily adapted to the mixed levels situation, as in the original algorithm by O'Flaherty & Mackenzie (1982). For our sample designs, these equations produce:

$$\{(b_1,b_2)\} = \{(0,0),(0,1),(1,0),(1,1)\}$$

for 2^2 selections, and hence $(b_1,b_2) = (1,1)$ when $g = 4$ as for D_1. And

$$\{(b_1,b_2)\} = \{(0,0),(0,1),(0,2),(1,0),(1,1),(1,2),(2,0),(2,1),(2,2)\}$$

hence $(b_1,b_2) = (1,1)$ when $g = 5$ as for D_2. A set containing k generators for U_L can then be found in the sequentially numbered columns of β_g as selected by the first k places in ω_f, with the complementary *rows* of β_g^{-1} (over $GF(q)$) containing generators for $U_L^{\perp}$. Because β_g arrives in upper triangular form with units down the diagonal, inversion can be performed with just one pass around

a single arithmetic statement taken from the standard Gram-Schmidt inversion algorithm. For our pair of example designs, we have

$$\beta_4(D_1) = \beta_5(D_2) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and since $\omega_3 = (2, 3, 1)$, this gives generators $< (110), (101) >$ for both the designs D_1 and D_2. Then

$$\beta_4^{-1}(D_1) = \beta_4(D_1),\ \beta_5^{-1}(D_2) = \begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

giving generator $< (111) >$ for the defining contrast of design D_1 and $< (122) >$ for design D_2. So both constituent designs for our merged design $U_{87}(3, 6, 2)$ will have (maximum) resolution, 3 and they also have minimum aberration within their own class.

When decoding L so as to find f and g, (and hence calculate β_g) we use

$$f = Arg\left\{ \underset{f}{Max}\left[\sum_{t<f} q^{j_t} < L \right] \right\},\ \ g = L - \sum_{t<f} q^{j_t}. \tag{3.5}$$

For example,

$$2^{j_1} = 2^0 < 7,\ 2^{j_1} + 2^{j_2} = 2^0 + 2^1 < 7,\ 2^{j_1} + 2^{j_2} + 2^{j_3} = 2^0 + 2^1 + 2^2 \geq 7$$
$$\therefore f = 3 \text{ and } g = 7 - 2^{j_1} + 2^{j_2} = 4 \text{ for } D_1.$$

For $q = p_1 p_2$ we will obtain $L_1 = L(p_1)$ and $L_2 = L(p_2)$, using the decomposition scheme described in Section 2 followed by the numbering scheme described above, for each subspace. We then define a unique number for the combined design by

$$L = (L_1 - 1) * S_{n,k;p_2} + L_2 \tag{3.6}$$

where $S_{n,k;q}$ is the standard formula for the number of k-subspaces in an n-dimensional $GF(q) -$ vector space. See Carmichael (1937) and Laycock & Rowley (1995). For example, we have

$$S_{3,2;3} = 13 \text{ and } L = (7 - 1) * S_{3,2;3} + 9 = 87$$

for $D_1 = U_7(3, 2, 2)$ and $D_2 = U_9(3, 3, 2)$. Conversely, it is clearly easy to recover L_1 and L_2 from L, when given n, k, and $q = p_1 p_2$ where $p_1 < p_2$.

The two sets of generators $\{u_{1i}\}$ for U_{L_1} and $\{u_{2i}\}$ for U_{L_2} must then be separately mapped over into $\mathbb{Z}_q$ by setting

$$u^*_{1i} = u_{1i}(q/p_1) \text{ and } u^*_{2i} = u_{2i}(q/p_2).$$

They can then be combined into the whole k-space $\{v_i\}$, without duplication of terms, via

$$\{v_i\} = \left\{\sum[\lambda_{1i}u^*_{1i} + \lambda_{2i}u^*_{2i}] \mid \lambda_{ji} \in \mathbb{Z}_{p_j}, j = 1, 2, \forall i\right\} \tag{3.7}$$

with arithmetic modulo q, where $q = p_1p_2$. For our example designs D_1 and D_2 this becomes

$$\{[\lambda_{11}(3,3,0) + \lambda_{12}(3,0,3) + \lambda_{21}(2,2,0) + \lambda_{22}(2,0,2)]\}$$

where $\lambda_{1i} \in (0,1)$, $\lambda_{2i} \in (0,1,2)$, $i = 1,2$, thus generating all $2 \times 2 \times 3 \times 3 = 36$ design points in $U_{87}(3,6,2)$.

A similar method applies for the defining contrasts $U^{\perp}_{L_j}$, $j = 1,2$, as derived from the generators in β_g^{-1}, except that when different levels are present $U^{\perp}_{L}$ would have to be determined by a direct search for orthogonal vectors over the whole space, as in Wynn *et al* (1996). For our example design $U_{87}(3,6,2)$ the defining contrasts are $< (155) >$, implying a (minimum aberration) design with resolution 3 and aberration count 5.

For generating and numbering the first k places of the permutation, ω_t, we have adapted an algorithm suggested in a private communication by David Bedford (1995a). For this,

$$t = \binom{n}{k} - \sum_{j=1}^{k} \binom{n - i_j}{k - j + 1} \tag{3.8}$$

whilst for decoding t when $t > 1$ ($t = 1$ is the identity permutation), we set $i_0 = 0$, $m_0 = t$, and use the recursive relation

$$i_j = i_{j-1} + Arg\left\{\underset{u}{Max}\left[m_j(u) = m_{j-1} - \sum_{s=1}^{u}\binom{n - i_{j-1} - s}{k - j} > 0\right]\right\} \tag{3.9}$$

which generates these permutations in lexicographical order.

3. SOME SPECIMEN DESIGNS

In this section we present a small selection of interesting q^{n-m} fractional designs which illustrate the scope of the algorithm given in Sections 2 & 3. We also find orbits in some cases for the minimum aberration designs, that is, those sets of such designs which are invariant under re-labelling of the factor names and levels.

3.1. $N = 3, K = 2, Q = 6$

There are $7 \times 13 = 91$ designs in this class, of which the 4 minimum aberration designs have resolution 3. There are two orbits amongst these 4 minimum aberration designs, of size 3 and 1 respectively. A representative of the first orbit is

$$U_{87} = \ < (330), (303), (220), (202) >$$

which is not a latin square, although it contains a complete factorial design in any pair of its factors. The notation $< \cdots \cdots >$ used here and elsewhere indicates that the selected space can be generated from the listed vectors. This can be done in the usual way using arithmetic modulo q, although for efficient generation to avoid duplication of entries when q is a multiple of distinct primes, the modified method set out in (3.7) above should be used.

The second orbit has one member and this is given by the combination of designs D_1 and D_2 examined in section 3 above:

$$U_{91} =< (330), (303), (420), (402) >$$

and this is a cyclic Latin Square.

3.2. $N = 4, K = 3, Q = 6$

There are $15 \times 40 = 600$ designs in this class, of which the 8 minimum aberration designs have resolution 4. There are three orbits, of sizes 4, 3 and 1 respectively. A representative of the first orbit is

$$U_{587} =< (3300), (3030), (3003), (2200), (2020), (2002) >$$

A representative of the second orbit is U_{588}, whilst the last orbit has representative U_{600}.

3.3. $N = 4, K = 2, Q = 6$

There are $35 \times 130 = 4550$ designs in this class, of which the 48 minimum aberration designs have resolution 2. There are two orbits, each of size 24. A representative of the first orbit is

$$U_{2431} =< (3300), (3033), (2220), (2402) >$$

Unlike the situation for minimum aberration designs when q is a prime power and k=2, this is not a pair of mutually orthogonal Latin Squares (MOLS). Although all 36 pairs are represented in the body of the square, only one of the individual squares is Latin. This is a neccessary consequence of Euler's conjecture, which certainly holds (but possibly only) for n=6. For a survey of this mathematically interesting area, see Bedford (1995b) and Mullen (1995). A representative for the second orbit is U_{2433}.

3.4. $N = 5, K = 2, Q = 6$

There are $155 \times 1210 = 187550$ designs in this class. This complete set of designs was generated and inspected for resolution and aberration in 435 seconds on a 233MHZ Pentium PC. The 1920 minimum aberration designs have resolution 2. One of these minimum aberration designs is

$$U_{99824} =<< (33300), (03033) >, < (02220), (22402) >>$$

Naturally, this is not a set of MOLS. Although all levels are equally represented and it can be laid out as a three factor 6×6 square, not all of the $\binom{6}{2}$ factor pairings are fully represented and only one of the individual squares is Latin.

3.5. $N = 5, K = 3, Q = 6$

There are $155 \times 1210 = 187550$ designs in this class. The 1920 minimum aberration designs have resolution 3. There are 17 orbits, each of size 60 or 120. One of these minimum aberration designs is

$$U_{99824} =< (33000), (00330), (30303), (02200), (22020), (24002) >$$

whilst the design numbers of representatives for the other sixteen orbits are L = 100132, 100342, 105068, 105259, 109780, 110007, 145966, 148558, 148787, 163214, 166157, 167157, 167941, 172676, 177749, 181978, and 104883.

3.6. $N = 4, K = 2, Q = 10$

There are $35 \times 806 = 28210$ designs in this class of which 1152 have minimum aberration with resolution 2. One of these minimum aberration designs is

$$U_{14847} =< (5500), (5055), (2220), (2402) >$$

This comprises two superimposed Latin Squares, but they are not mutually orthogonal. However pairs of 10×10 MOLS are known to exist: Parker (1960).

3.7. $N = 5, K = 2, Q = 10$

There are $155 \times 20306 = 3147430$ designs in this class of which 23040 have minimum aberration with resolution 2. One of these minimum aberration designs is

$$U_{1673707} =< (55500), (05055), (22220), (24602) >$$

This can be laid out as a three factor 10×10 square, but only one of the individual squares is then Latin.

Acknowledgement We are grateful to an anonymous referee for comments and suggestions which significantly improved the presentation of these results.

References

Bailey, R.A., Gilchrist, F.H.L. and Patterson, H.D. (1977). Identification of effects and confounding patterns in factorial designs. *Biometrika* **64**, 347–354.

Bedford, D. (1995a). *Private communication*. University of Keele.

Bedford, D. (1995a). Orthomorphisms and near orthomorphisms of groups and orthogonal Latin squares: a survey. Preprint, University of Keele.

Constantine, G.M. (1987). *Combinatorial Theory and Statistical Design*. New York: Wiley.

Finney, D.J. (1945). The fractional replication of factorial arrangements. *Ann. Eugen., Lond.* **12**, 291–301.

John, J.A. and Dean, A.M. (1975). Single replicate factorial experiments in generalized cyclic designs. I. Symmetrical arrangements. *J. R. Statist. Soc.* B **37**, 72–76.

John, J.A. and Williams, E. R. (1995). *Cyclic and Computer Generated Designs*. London: Chapman & Hall.

Kempthorne, O. (1947). A simple approach to confounding and fractional replication in factorial experiments. *Biometrika* **34**, 255–272.

Laycock, P.J. and Rowley, P.J. (1995). A Method for generating and labelling all regular fractions or blocks for qˆ(n-m) designs. *J. Roy. Statist. Soc.* B **57**, 191–204.

Laycock, P.J. and Rowley, P.J. (1999). On generating and classifying all qˆ(n-m-l) regularly blocked fractional designs. In *Stochastic Simulation Methods* Eds. S.M. Ermakov, N. Balakrishnan and V.B. Melas, pp. 159–170. Boston: Birkhauser.

Mullen, G.L. (1995). A candidate for the "Next Fermat Problem". *The Mathematical Intelligencer* **17**, No.3. New York: Springer-Verlag.

Parker, E.T. (1960). Orthogonal Latin squares. *Proc. Nat. Acad. Sci. U.S.A.* **21**, 859–862.

Patterson, H.D. and Bailey, R.A. (1978). Design keys for factorial experiments. *Appl. Statist.* **27**, 335–343.

Raktoe, B.L., Hedayat, A. and Federer, W.T. (1981). *Factorial Designs*. New York: Wiley.

Street, A.P. and Street, D.J. (1987). *Combinatorics of Experimental Design*. Oxford: Clarendon Press.

Wynn, H.P., Sivaloganathan, S., Buck, R.J. and Lewis, S.M. (1996). GENERATE: An algorithm for the computer generation of orthogonal arrays with specified alias structure. Preprint, City University, London.

Chapter 7

SECOND-ORDER OPTIMAL SEQUENTIAL TESTS

M.B. Malyutov
Department of Mathematics, Northeastern University
Boston MA 02115 USA
MLTV@neu.edu

I.I.Tsitovich
Institute for Information Transmission Problems, B. Karetny 19
Moscow, 101447, Russia
cito@iitp.ru

Abstract An asymptotic lower bound is derived involving a second additive term of order $\sqrt{|\ln \alpha|}$ as $\alpha \to 0$ for the mean length of a controlled sequential strategy s for discrimination between two statistical models in a very general nonparametric setting. The parameter α is the maximal error probability of s.

A sequential strategy is constructed attaining (or almost attaining) this asymptotic bound uniformly over the distributions of models including those from the indifference zone. These results are extended for a general loss function $g(N)$ with the power growth of the strategy length N.

Applications of these results to change-point detection and testing homogeneity are outlined.

Keywords: change-point detection, controlled experiments, general risk, second-order optimality, sequential test

Introduction

There are numerous references on sequential hypothesis testing and the quick detection of parameter changes in a stochastic system. The aim of the present paper is three-fold:
1. To construct second-order optimal sequential strategies strengthening the traditional ones;

2. To do this for a non-parametric setting with control, indifference zone and general risk;
3. To show the applicability of the above ideas and constructions for change-point detection and testing homogeneity.

We begin with sequential discrimination between a **finite** set of distributions in Section 1.2 which illuminates the further general exposition.

1. BACKGROUND

1.1. SETTING OF THE PROBLEM

To start we outline the non-parametric setting of the problem following Malyutov and Tsitovich (1997a) and Malyutov and Tsitovich (2000) which dealt with first-order optimal tests.

No control case. Let $(X, \mathcal{B}, \mu)$, $X \subset \mathbf{R}$, be a probability space and $(\mathcal{P}, d(\cdot))$ be a metric space, where $\mathcal{P}$ is a subset of the set $\mathcal{A}$ of mutually absolutely continuous (AC) probability measures which are also AC with respect to a sigma-finite measure μ. Their densities are denoted by the corresponding lower case letters.

Denote by $\mathbf{E}_P f(X)$ the expectation of $f(X)$, where X is a random variable (RV) with distribution P. Let $I(P,Q) = \mathbf{E}_P \log \frac{p(X)}{q(X)}$ be the relative entropy (Kullback-Leibler divergence) with the usual conventions (logarithms are to the base e, $0 \log 0 = 0$ etc.) and the metric d be *I-uniformly continuous* on $\mathcal{P}$, i.e. for every $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that $d(P,Q) < \varepsilon$ for every P, Q from $\mathcal{P}$ such that $I(P,Q) < \delta$.

The set $\mathcal{P}$ is partitioned into sets $\mathcal{P}_0$, $\mathcal{P}_1$ and the indifference zone $\mathcal{P}_+ = \mathcal{P} \setminus (\mathcal{P}_1 \cup \mathcal{P}_0)$. We test $H_0 : P \in \mathcal{P}_0$ versus $H_1 : P \in \mathcal{P}_1$, every decision is good for $P \in \mathcal{P}_+$.

Suppose that the distance between the hypotheses is positive, i.e.

$$\inf_{P \in \mathcal{P}_0, Q \in \mathcal{P}_1} d(P,Q) \geq \delta_0 > 0. \tag{7.1}$$

A *strategy* s consists of a stopping time N and a measurable binary decision $\delta : \delta = r$ means that H_r , $r = 0, 1$, is accepted. Observations $X_i, i = 1, \ldots, n$ are assumed independent and P-identically distributed (i.i.d.(P)), $P \in \mathcal{P}$; whether $N > n$ or not depends only on $X_i, i = 1, \ldots, n$ for every n.

For an $\alpha > 0$ introduce *α- strategies* s satisfying

Condition $G(\alpha) : \max_{r=0,1} \sup_{P \in \mathcal{P}_r} \mathbf{P}_P(\delta = 1 - r) \leq \alpha$.

Remark 1. For simplicity of notation we have confined ourselves here to testing two hypotheses with the same maximal error probabilities. The more general condition $G(\alpha, \mathbf{c}, \mathbf{d}) : \sup_{P \in \mathcal{P}_r} \mathbf{P}_P(\delta = 1 - r) \leq c_r \alpha^{d_r}, c_r > 0, d_r > 0$ can be tackled similarly (Tsitovich, 1990). For an extension to testing several hypotheses see e.g. Lai (1995).

Let $\mathbf{E}_P^s N$ be the *mean length* (MEL) of a strategy s. Our first aim is to find an expansion for the lower bound for MEL as $\alpha \to 0$.

Define $I(P, \mathcal{R}) = \inf_{Q \in \mathcal{R}} I(P, Q)$ for $\mathcal{R} \subset \mathcal{P}$; $A(P) = \mathcal{P}_{1-r}$ for $P \in \mathcal{P}_r$ as the alternative set in $\mathcal{P}$ for P. For $P \in \mathcal{P}_+$, if $I(P, \mathcal{P}_0) \le I(P, \mathcal{P}_1)$, then $A(P) = \mathcal{P}_1$, otherwise, $A(P) = \mathcal{P}_0$. Finally $k(P) = I(P, A(P))$.

Theorem 1 states: for every α-strategy s under mild regularity conditions

$$\mathbf{E}_P^s N \ge \frac{|\log \alpha|}{k(P)} + O(\sqrt{|\log \alpha|}). \tag{7.2}$$

In Theorem 3 under stricter regularity conditions we construct the α-strategy s^* attaining equality in (7.2).

Controlled experiments. Here metric spaces $\mathcal{P}^u$ of some distributions $P^u \in \mathcal{A}$ labelled by controls $u \in U = \{1, \dots, m\}$ are given. Introduce a **mixed control** $\mathbf{u} = (\kappa_1, \dots, \kappa_m)$, where $\kappa_u \ge 0$ and $\sum_{u=1}^m \kappa_u = 1$; U^* is the set of mixed controls, $P^{\mathbf{u}} = \sum \kappa_u P^u, \mathcal{P}^{\mathbf{u}} = \{P^{\mathbf{u}}\}$. Introduce an m-**tuple of measures** $P = \{P^u, u \in U, P^u \in \mathcal{P}^u\}, \mathcal{P}$ is the set of such m-tuples.

We assume that metrics d_u are given which are I-uniformly continuous on $\mathcal{P}^u = \{P^u : P \in \mathcal{P}\}$ for each u with respect to the relative entropy and (7.1) holds for the metric $d(P, Q) = \max_{u \in U} d_u(P^u, Q^u)$.

After obtaining the n-th observation the experimenter either decides to stop or to choose a mixed control for the $(n+1)$st experiment. Let $\mathcal{F}_n$ be the σ-algebra generated by the observations and controls up to time n. We suppose that the $(n+1)$st experiment is predictable, i.e. the corresponding distribution on U is $\mathcal{F}_n$-measurable, the strategy length N is a stopping time under the flow $\mathcal{F}_n$ and a decision δ is $\mathcal{F}_N$-measurable. A strategy s now consists of a rule $\mathbf{u}(\cdot)$ for choosing the mixed controls, a stopping time N and a decision δ. For more details about constructing a probability space and a control strategy see Malyutov (1983).

Define

$$I_{\mathbf{u}}(P, Q) = \sum_{i=1}^m \kappa_i I(P^i, Q^i)$$

for a mixed control $\mathbf{u}$ and introduce

$$k^*(P) = \max_{\mathbf{u} \in U^*} \inf_{Q \in A_{\mathbf{u}}(P)} I_{\mathbf{u}}(P, Q) > 0, \tag{7.3}$$

((7.1) implies the inequality in (7.3)); $\mathbf{u}^* = \mathbf{u}^*(P)$ is a control such that

$$k^*(P) = \inf_{Q \in A_{\mathbf{u}^*}(P)} I_{\mathbf{u}^*}(P, Q),$$

where, for the mixed control $\mathbf{u}$ $A_{\mathbf{u}}(P)$ is the alternative set in $\mathcal{P}^{\mathbf{u}}$ for $P^{\mathbf{u}}$.

Our main result for controlled sequential testing consists in proving a modified lower bound (7.10) and constructing a second-order optimal α-admissible strategy s satisfying (7.10).

1.2. FINITE SET OF DISTRIBUTIONS

Sequential controlled discrimination between distributions from a finite set $\mathcal{P}$ was pioneered in Chernoff (1959). Indifference zones were first incorporated in Schwarz (1962). Chernoff (1997) surveys the subsequent development of asymptotically optimal sequential testing of the first order (AOSP1) with the MEL satisfying $\mathbf{E}_P^s N = |\log(\alpha)| k(P)^{-1}\big(1+o(1)\big)$ in a Bayesian framework. The wideness of the class of AOSP1 allows excessive values of the risk function for finite samples. Constructing procedures of the higher (second-) order of optimality (AOSP2) therefore became desirable. The results of Keener (1984) on Bayesian sequential discrimination between a finite number of distributions imply that (7.4) holds for the optimal strategy with some nonnegative K. Under additional conditions a strategy is constructed in Tsitovich (1984) with MEL exceeding the optimal value by $O(\log|\log(\alpha)|)$.

No control case.

Introduce $z(P,Q,x)=\log\frac{p(x)}{q(x)}$, $L_n(P,Q)=\sum_{i=1}^n z(P,Q,x_i)$ and the set $\mathcal{Q}(P)$=$\{Q: k(P)=I(P,Q), Q\in A(P)\}$, $l:=\#\mathcal{Q}(P)$.

Proposition 1. *Let RV $z(P,Q,x)$ possess fourth moments for all P and Q. For every α-strategy s and $l=1$*

$$\mathbf{E}_P^s N \geq \frac{|\log(\alpha)|}{k(P)} + O(1)$$

and, for $l>1$,

$$\mathbf{E}_P^s N \geq \frac{|\log(\alpha)|}{k(P)} + K\sqrt{\frac{|\log(\alpha)|}{k(P)}}\big(1+o(1)\big), \tag{7.4}$$

where $K=\mathbf{E}(g(\zeta))>0$, $\zeta=(\zeta_1,\dots,\zeta_l)$ is a normally distributed RV with mean **0** *and covariance matrix $\Sigma=(\Sigma_{ij})$, $\Sigma_{ij}=\mathbf{E}_P(z(P,Q_i,x)z(P,Q_j,x))$, $g(\zeta)=\max_{i=1,\dots l}\zeta_i$.*

Sketch of proof (the full proof is in Tsitovich, 1993). Define $\mathbf{L}_n(P)=(L_n(P,Q_1),\dots,L_n(P,Q_l))$, $\mathbf{x}_0=\log(\alpha)\mathbf{1}$, $\mathbf{1}=(1,\dots,1)$.

Introduce τ, the time when the likelihood ratio of the true distribution P with respect to every alternative first exceeds the prescribed level $|\log(\alpha)|$. It is not a stopping time (for unknown P) but it is shown in Keener (1984), Lemma 5.1, that every strategy has an MEL exceeding $\mathbf{E}_P\tau+const$ for all α.

Hence, the lower bound can be obtained by minimizing the mean time until the first entry into $R^+=\{\mathbf{x}\in R^l: x_i\geq 0, i=1,\dots,l\}$ for the process $\mathbf{x}_0+\mathbf{L}_n(P)$.

Applying Theorem 4.1 (Keener, 1984) for our minimization problem we get the results of Proposition 1.

Controlled experiments. For simplicity let the optimal control u^* be unique and $u_1,\dots,u_t$ be controls in U such that $u^*(P)=(\kappa_1,\dots,\kappa_t,0,\dots,0)$, $\kappa_i>$

0, and $\sum_{i=1}^{t} \kappa_i = 1$. Introduce vectors $\mu_i = (I(P^i, Q_1^i), \ldots, I(P^i, Q_l^i)), i = 1, \ldots, t$, where $Q_j \in A(P)$ are such that $I_{u^*(P)}(P, Q_j) = k(P), j = 1, \ldots, l$. Usually $l \geq 2$ for controlled discrimination.

The **regular case** is defined such that $\mu_i, i = 1, \ldots, t$, span the subspace L and $\dim(L) = l$. In the **degenerate case** $\dim(L) < l$.

An example of the degenerate case is the following: $\mathcal{P} = \{P_0, P_1, P_{-1}\}$, $\mathcal{P}_0 = \{P_0\}, \mathcal{P}_1 = \{P_1, P_{-1}\}, U = \{1, 2\}$, and $I(P_0^u, P_r^u) = I, r = 1, -1, u = 1, 2$.

Let $F_{st}(P, \mathbf{x})$ be the minimal value of the function

$$f = \sum_{i=1}^{t} \kappa_i \tag{7.5}$$

under the constraints

$$\kappa_i \geq 0, i = 1, \ldots, t, \mathbf{x} + \sum_{i=1}^{t} \kappa_i \mu_i \geq \mathbf{0} \tag{7.6}$$

and

$$g(\zeta) = F_{st}(P, \mathbf{x}_0 + \zeta) - |\log \alpha| k(P)^{-1}.$$

Proposition 2. *Let the RV $z(P^u, Q^u, x)$ possess fourth moments for all P, Q, and u. For every α-strategy s in the regular case*

$$\mathbf{E}_P^s N \geq \frac{|\log(\alpha)|}{k(P)} + O(1)$$

and, in the degenerate case,

$$\mathbf{E}_P^s N \geq \frac{|\log(\alpha)|}{k(P)} + K\sqrt{\frac{|\log(\alpha)|}{k(P)}}\big(1 + o(1)\big), \tag{7.7}$$

where $K = \mathbf{E}(g(\zeta)) > 0$, $\zeta = (\zeta_1, \ldots, \zeta_l)$ is a normally distributed RV with mean $\mathbf{0}$ and covariance matrix

$$\Sigma = (\Sigma_{j_1 j_2}), \Sigma_{j_1 j_2} = \sum_{i=1}^{t} \kappa_i \mathbf{E}_P(z(P^i, Q_{j_1}^i, x) z(P^i, Q_{j_2}^i, x)).$$

Sketch of proof (the full proof is in Tsitovich, 1993). For a finite set U of controls the asymptotically optimal control rule should provide the fastest approach to the positive octant R^+ by the vector composed of likelihood ratios. This problem was studied in Lalley and Lorden (1986). A simpler non-stochastic control problem, analogous to (7.5) and (7.6), was studied in Keener (1984) for sequential discrimination in a Bayesian setting.

Under the conditions of Proposition 2, strategies s^* with MEL $\frac{|\log(\alpha)|}{k(P)}+O(1)$ for the regular case and with

$$\mathbf{E}_P^{s^*} N = \frac{|\log(\alpha)|}{k(P)} + K\sqrt{\frac{|\log(\alpha)|}{k(P)}} + O(1), \tag{7.8}$$

where K is the same constant as in (7.7), for the degenerate case, are constructed in Tsitovich (1993). In contrast to the regular case, the optimal control from Lalley and Lorden (1986) does not apply for the last phase of the strategy with condition (7.8). A full description of the procedure under the condition (7.8) is outlined in Tsitovich (1993).

1.3. OUTLINE OF CHANGE-POINT DETECTION

Our strategy and bounds appear readily applicable to the non-parametric detection of abrupt change in the distribution of an i.i.d. sequence *without* an indifference zone. We use the methodology outlined in Lorden (1971).

Let the observations $X_1, \dots, X_n, \dots$ be independent and, for $n < \nu$, all having a distribution $P_0 \in \mathcal{P}_0$, while all the X_n with $n \geq \nu$ have an unknown distribution $P_1 \in \mathcal{P}_1$, where $\nu \geq 1$ is an unknown integer, and (7.1) holds.

Let N be a change-point estimate, and $a^+ = a$, if $a \geq 0$, and $a^+ = 0$ otherwise. Following Lorden (1971) we use the functional

$$\bar{\mathbf{E}}^s(N) = \sup_{\nu \geq 1} \operatorname{ess\,sup} \mathbf{E}_{P_1}^s \left((N - \nu + 1)^+ \, | X_1, \dots, X_\nu \right),$$

with index $P_1 \in \mathcal{P}_1$ suppressed, as the optimality criterion of the strategy s under the restriction that for a given $\alpha > 0$

$$\mathbf{E}_P^s (N) \geq \alpha^{-1}. \tag{7.9}$$

Let s be a sequential strategy for testing the hypothesis $H_1 : P \in \mathcal{P}_1$ versus $H_0 : P \in \mathcal{P}_0$ with a one-sided decision δ such that $\sup_{P \in \mathcal{P}_0} \mathbf{P}_P(\delta = 1) \leq \alpha$ and let T be its stopping time. For every t we denote by T_t the stopping time of s based on the observations $\mathbf{X}_t = (x_t, x_{t+1}, \dots)$. Define $N = \inf_t (T_t + t)$ as a stopping time of change-point detection. By Theorem 2 in Lorden (1971) $\mathbf{E}_P^s (N) \geq \alpha^{-1}$ for $P \in \mathcal{P}_0$. Hence asymptotically optimal one-sided strategies for hypothesis testing are also asymptotically optimal for change-point detection under the condition (7.9).

Theorem 3 in Lorden (1971) states that, under the condition (7.9), $\bar{\mathbf{E}}^s(N) \geq |\log \alpha| I_1^{-1}(1 + o(1))$, where $I_1 = I_1(P_1) = I(P_1, \mathcal{P}_0)$, if the condition C1 in section 2 is valid.

Lai (1995) gives a detailed survey of connections between sequential testing and change-point detection for finite or parametric $\mathcal{P}$.

We generalize these results for a non-parametric setting and study AOSP2 of change-point detection in subsection 4.2.

2. NON-PARAMETRIC SECOND ORDER BOUNDS

C1. There is some $c > 0$ such that, for all $P \in \mathcal{P}, Q \in \mathcal{P}$ and $u \in U$, $\mathbf{E}_P^u \left(z(P^u, Q^u, X)\right)^2 < c$.

We prove the following lower bound extending that of Malyutov and Tsitovich (2000).

Theorem 1. i. *Under the condition C1, every α-strategy s for the no-control problem satisfies* (7.2) *for every $P \in \mathcal{P}$.*
ii. *For controlled experiments and every $P \in \mathcal{P}$ the following inequality holds*

$$\mathbf{E}_P^s N \geq \frac{|\log \alpha|}{k^*(P)} + O\left(\sqrt{|\log \alpha|}\right). \tag{7.10}$$

The proof is a rather straightforward generalization of a lower bound in Malyutov and Tsitovich (2000).

Additional regularity conditions for constructing an optimal strategy are:
C2. There exist $t > 0$ and $f > 0$ such that for all $u \in U$ and $P \in \mathcal{P}$

$$\mathbf{E}_P \left(\sup_{Q \in \mathcal{P}} \exp(-tz(P, Q, X))\right) \leq f.$$

C3. $z(P, Q, x)$ is differentiable w.r.t. x and

$$D = \int_X z_1(x) \left(a(x)b(x)\right)^{1/2} dx < \infty,$$

where

$$z_1(x) = \sup_{Q \in \mathcal{P}} \left| \frac{\partial z(P, Q, x)}{\partial x} \right|,$$

$$\sup_{P \in \mathcal{P}} \int_{-\infty}^{x} p(t)\mu(dt) \leq a(x) \quad \text{and} \quad \sup_{P \in \mathcal{P}} \int_{x}^{\infty} p(t)\mu(dt) \leq b(x).$$

C4. There exist $b > 0$ and $K_1 = K_1(b)$ such that for every n the estimate $\hat{p} = \hat{p}_n$ for the density function of i.i.d.(P) observations $X_1, \ldots, X_n$ exists with

$$\mathbf{E}_P(I(P, \hat{P})) \leq K_1 n^{-b}. \tag{7.11}$$

Remark 2. If, for example, the set X is the interval $[0, 1]$ and, for $P \in \mathcal{P}$, the function $\log p : \mathbf{R} \to \mathbf{R}$ having period 1, belongs to the Sobolev space W_2^r on X, $r \geq 1$, then (Barron and Sheu, 1991) condition C4 is valid with $b = \frac{2r}{1+2r}$. Therefore, if additionally to C3 we assume

$$\int_X \left(\frac{\partial z(P, Q, x)}{\partial x}\right)^2 dx \leq c < \infty$$

with the boundary condition $z(P, Q, 0) = z(P, Q, 1)$, then C4 is valid for $b = \frac{2}{3}$.

If $\mathcal{P}$ is a quasi-homogeneous family (Centsov, 1982) then C4 is verified using the methodology of Centsov (1982) for density estimation with $b \geq \frac{1}{2}$ depending on the smoothness of $\mathcal{P}$.

Remark 3. Usually the estimate $\hat{P}$ is constructed by approximating $\mathcal{P}$ by a parametric exponential family of distributions $\mathcal{A}_m$ of dimension m and finding the ML-estimate as if $P \in \mathcal{A}_m$. Then

$$I(P, \hat{P}) \leq \gamma_1 m^{-r_1} + \gamma_2 \frac{m^{r_2}}{n} \tag{7.12}$$

for some $\gamma_1 > 0$ and $\gamma_2 > 0$, r_1 depends on the smoothness of $\mathcal{P}$, while r_2 depends on a basis for the family $\mathcal{A}_m$. Optimization of (7.12) over m gives (7.11).

Now we introduce our strategy $s^* = s^*(n)$ depending on a parameter n. Procedure s^* consists of conditionally i.i.d. loops. The loop terminating with the event (7.13) is the final loop of s^*. Every loop contains two phases.

Based on the first $L = [\sqrt{|\ln \alpha|}] + 1$ observations of a loop we estimate the density function P.

Let us enumerate measurements of the second phase anew and introduce $L_k(\hat{P}, Q) = \sum_{i=1}^{k} z(\hat{P}, Q, x_i)$, where $\hat{P}$ is the estimate of P in the first phase. We stop observations at the first moment M such that

$$\inf_{Q \in A_n(\hat{P})} L_M(\hat{P}, Q) > 1 + |\log \alpha| \tag{7.13}$$

or

$$M > 2k(\hat{P})^{-1} |\log \alpha| \tag{7.14}$$

and accept the hypothesis H_r (i.e. $\delta = r$) if (7.13) holds and $1 - r$ is the index of the set $A(\hat{P})$. In all other cases we begin a new loop.

Theorem 2. *For every $P \in \mathcal{P}$ under the conditions C1-C4 and appropriate parameters, s^* is an α-strategy and*

$$\mathbf{E}_P^{s^*} N \leq \frac{|\log \alpha|}{k(P)} + O\left(|\log \alpha|^{1-b/2}\right) + O\left(\sqrt{|\log \alpha|}\right). \tag{7.15}$$

The proof is similar to that of Theorem 3 in Malyutov and Tsitovich (2000). The difference is in estimating the mean length $\mathbf{E}_P^{s^*} N_2$ of the second phase which is the principal part of the total mean length.

Let $\mathcal{F}_0$ be the σ-algebra generated by all observations before the second phase starts and let $\mathcal{T}$ be an event $\{A(\hat{P}) = A(P)\}$. Introduce

$$\xi(k) = L_k(\hat{P}, Q) - L_{k-1}(\hat{P}, Q) - I(P, Q) + I(P, \hat{P}), \quad \Xi(k) = \sum_{l=1}^{k} \xi(l).$$

It can be shown that $\Xi(k)$ is a martingale and, under the validity of the event $\mathcal{T}$, C3 implies the bound

$$\mathbf{E}_P^{s^*}\left(\sup_{Q\in A(\hat{P}),k\le l}|\Xi(k)|\,\middle|\,\mathcal{F}_0,\mathcal{T}\right)=O\left(\sqrt{l}\right). \tag{7.16}$$

From C4, as in the proof of Theorem 3 in Malyutov and Tsitovich (1997b),

$$\mathbf{E}_P^{s^*}\left(N_2|\mathcal{F}_0,\mathcal{T}\right)\le\frac{|\log\alpha|}{k(P)-\varepsilon}+O\left(\sqrt{|\log\alpha|}\right), \tag{7.17}$$

where $\varepsilon=K_1|\log\alpha|^{-b/2}$.

We see that the second term in (7.15) is generally of a larger order of magnitude than that in (7.10). It is remedied by a strategy with more than two phases. The number of phases depends on b.

Theorem 3. *For every $P\in\mathcal{P}$ under the conditions C1-C4 with $b\ge\frac{1}{2}$ and appropriate parameters, the multi-phased strategy s^* is an α-strategy and*

$$\mathbf{E}_P^{s^*}N\le\frac{|\log\alpha|}{k(P)}+O\left(\sqrt{|\log\alpha|}\right).$$

The proof is based on making $\varepsilon\le|\log\alpha|^{-1/2}$ in (7.17) by more precise estimation of P in an appropriate multi-phased procedure.

3. NON-PARAMETRIC TESTING WITH CONTROL

Let the following regularity conditions hold:

C5. i. For every $u\in U$ the condition C4 is valid with the same values b and $K_1(b)$. ii. There exists a sequence of mixed controls $\mathbf{u}_n(P)$, $c>0$ and $K_2=K_2(c)$, such that $\mathbf{u}_n(\hat{P}_n)$ is a measurable control for every n and i.i.d.(P) observations $X_1,\dots,X_n$, $P\in\mathcal{P}$, such that

$$\mathbf{E}_P|\inf_{Q\in A(P)}(I_{u_n(\hat{P}_n)}(P,Q)-k(P))|\le K_2n^{-c}.$$

Remark 4. C5 can be verified by approximating $\mathcal{P}$, as in Remark 3, by exponential families using the methodology of Centsov (1982).

Our procedure $S^*=S^*(n)$, as before, has conditionally i.i.d. loops until the final success, almost as in (7.13). If an analogue of (7.14) holds, a new loop begins. Every loop contains two phases.

In the first phase we obtain the estimate $\hat{P}^u$ of P^u for every $u\in U$ based on $L=[\sqrt{|\ln\alpha|}]+1$ independent observations with this control.

We use the control $u(i)=u_L(\hat{P})$ for the i-th measurement of the second phase, stop observations and take the decision δ as in the strategy s^*.

Theorem 4. *Under the conditions C1-C5 for every $P \in \mathcal{P}$ and appropriate parameters, S^* is an α-strategy and*

$$\mathbf{E}_P^{S^*} N \le \frac{|\log \alpha|}{k(P)} + O\left(|\log \alpha|^{1-d/2}\right) + O\left(\sqrt{|\log \alpha|}\right)$$

where $d = \min(b, c)$.

The proof is similar to that of Theorem 2. Only, since the control $u_L(\hat{P})$ replaces $u^*(P)$, the estimate (7.17) is replaced by

$$\mathbf{E}_P^{S^*} (N_2 | \mathcal{F}_0, \mathcal{T}) \le \frac{|\log \alpha|}{k(P) - \varepsilon} + O\left(\sqrt{|\log \alpha|}\right),$$

where $\varepsilon = K_1 |\log \alpha|^{-\frac{b}{2}} + K_2 |\log \alpha|^{-\frac{c}{2}}$.

A further strengthening of the upper bound can be obtained by studying multi-phased strategies as in Section 2

4. EXTENSIONS AND APPLICATIONS

4.1. GENERAL RISK

Our Theorems 1 and 2 can be extended to allow a general loss function of the strategy length with power growth, as in Malyutov and Tsitovich (2000). Let $g : \mathbf{R}^+ \to \mathbf{R}^+$ be continuous, increasing to infinity and
i. $g(an) \le K_g a^k g(n)$ for all $a > 1$, large n, and some positive K_g and k;
ii. $1 - c_g \frac{m}{n} \le \frac{g(n+m)}{g(n)} \le 1 + C_g \frac{m}{n}$, with the same constants c_g and C_g for all n and k.

Let $\mathbf{E}_P^s g(N)$ be the risk of a strategy s.

Theorem 5. *Under condition C1 every α-admissible strategy s satisfies*

$$\mathbf{E}_P^s g(N) \ge g\left(\frac{|\log \alpha|}{k(P)}\right)\left(1 + O\left(|\log \alpha|^{-1/2}\right)\right).$$

Under the conditions C1-C4 and appropriate parameters, s^* is an α-strategy and

$$\mathbf{E}_P^{s^*} g(N) \le g\left(\frac{|\log \alpha|}{k(P)}\right)\left(1 + O\left(|\log \alpha|^{-b/2}\right) + O\left(|\log \alpha|^{-\frac{1}{2}}\right)\right).$$

The proof is a rather straightforward generalization of that for Theorems 1 and 2.

4.2. CHANGE-POINT DETECTION

We use a modification of our strategy s^* for the sequential detection of a change-point in the setting outlined in Section 1.3.

The modified strategy s^* is the following: let $\mathbf{X}_t = (x_t, x_{t+1}, \dots)$, where x_i is the i-th measurement. For every $t \geq 1$ under the sequence of measurements $\mathbf{X}_t$ we construct one loop of s^*, which is denoted by s_t^*, for estimating ν.

If (7.14) holds then we take $N_t = \infty$.

Theorem 6. *Under condition C1 for every α-strategy and sufficiently small α we have the lower bound*

$$\bar{\mathbf{E}}^s(N) \geq |\log\alpha| I_1^{-1} + O\left(\log|\log\alpha|\right).$$

Under conditions C1-C4 and appropriate parameters, s^* is an α-strategy and

$$\bar{\mathbf{E}}^s(N) \leq |\log\alpha| I_1^{-1} + O\left(|\log\alpha|^{1-b/2}\right) + O\left(\sqrt{|\log\alpha|}\right). \tag{7.18}$$

Proof. The lower bound follows from the proof in Lorden (1971). Condition (7.9) follows from Theorem 2 in Lorden (1971) and (7.18) follows from our Theorem 2. The risk $\bar{\mathbf{E}}^{s^*} g((N-\nu+1)^+)$ is bounded from above by Theorem 4.

4.3. SEQUENTIAL TESTING OF HOMOGENEITY

Let independent observations be taken from $m \geq 2$ populations with distributions $P_1, \dots, P_m$ on $(X, \mathcal{B}, \mu)$, $P_i \in \mathcal{Q}$, when conditions C1-C4 hold for $\mathcal{Q}$. We test $H_0 : P_1 = \cdots = P_m$ versus $H_1 : \max_{i,j} d(P_i, P_j) \geq \Delta > 0$, where d is an I-uniformly continuous distance on $\mathcal{Q}$. Application of control u means that we take an observation from the u-th population.

This is the problem of controlled experiments for testing the hypothesis $H_0 : P \in \mathcal{P}_0$ versus $H_1 : P \in \mathcal{P}_1$ with the indifference zone $P \in \mathcal{P}_+$, where $P = (P_1, \dots, P_m), P_i \in \mathcal{P}$, $\mathcal{P}_0 = \{P : P_1 = \cdots = P_m\}$, $\mathcal{P}_1 = \{P : \max_{i,j} d(P_i, P_j) \geq \Delta > 0\}$, $\mathcal{P}_+ = \{P : 0 < \max_{i,j} d(P_i, P_j) < \Delta\}$.

Let $X = [0,1]$, $\mathcal{Q}$ be such that $\log p \in W_2^r, ||D^r \log p||_2 \leq K$, $r > 1$, $p : \mathbf{R} \to \mathbf{R}$ have period 1 and let μ be Lebesgue measure. It can be deduced from Barron and Sheu (1991) that condition C5 is valid with $c = \frac{r-1}{2r}$. Therefore it follows from Remark 2 that the conditions of Theorem 4 are valid and our general strategy from the previous section is applicable for testing homogeneity.

Theorem 7. *Under conditions C1-C3 and appropriate parameters, S^* is an α-strategy and*

$$\mathbf{E}_P^{S^*} N \leq \frac{|\log\alpha|}{k^*(P)} + O\left(|\log\alpha|^{-d}\right) + O\left(\sqrt{|\log\alpha + |}\right),$$

where $d = \frac{r-1}{4r}$.

Acknowledgement. The authors are grateful to Prof. H. Chernoff for improving our language and style and for making our historical remarks more precise.

References

Barron, A.R. and Sheu, C.-H. (1991). Approximation of density functions by sequences of exponential families. *Ann. Statist.* **19**, 1347–1369.

Centsov, N.N. (1982). *Statistical Decision Rules and Optimal Inference.* Amer. Math. Soc. Transl. **53**. Providence RI.

Chernoff, H. (1959). Sequential design of experiments. *Ann. Math. Statist.* **30**, 755–770.

Chernoff, H. (1997). *Sequential Analysis and Optimal Design*. Philadelphia: SIAM.

Keener, R. (1984). Second order efficiency in the sequential design of experiments. *Ann. Statist.* **12**, 510–532.

Lai, T.L. (1995). Sequential change-point detection in quality control and dynamical systems. *J. Roy. Statist. Soc.* B **57**, 613–658.

Lalley, S.P. and Lorden, G. (1986). A control problem arising in the sequential design of experiments. *Ann. Probab.* **14**, 136–172.

Lorden, G. (1971). Procedures for reacting to a change in distribution. *Ann. Math. Statist.* **42**, 1897–1908.

Malyutov, M.B. (1983). Lower bounds for the mean length of sequentially designed experiments. *Soviet Math. (Izv. VUZ.)* **27**, 21–47.

Malyutov, M.B. and Tsitovich, I.I. (1997a). Asymptotically optimal sequential testing of hypotheses. In *Proc. Internat. Conf. on Distributed Computer Communication Networks: Theory and Applications*, pp. 134–141. Tel-Aviv.

Malyutov, M.B. and Tsitovich, I.I. (1997b). Sequential search for significant variables of an unknown function. *Problems of Inform. Transmiss.* **33**, 88–107.

Malyutov, M.B. and Tsitovich, I.I. (2000). Asymptotically optimal sequential testing of hypotheses. *Problems of Inform. Transmiss.* (To appear).

Schwarz, G. (1962). Asymptotic shapes of Bayes sequential testing regions. *Ann. Math. Statist.* **33**, 224–236.

Tsitovich, I.I. (1984). On sequential design of experiments for hypothesis testing. *Theory Probab. and Appl.* **29**, 778–781.

Tsitovich, I.I. (1990). Sequential design and discrimination. In *Models and Methods of Information Systems*, pp. 36–48. Moscow: Nauka (in Russian).

Tsitovich, I.I. (1993). *Sequential Discrimination*. D.Sci. thesis. Moscow (in Russian).

Chapter 8

VARIATIONAL CALCULUS IN THE SPACE OF MEASURES AND OPTIMAL DESIGN

Ilya Molchanov
Department of Statistics
University of Glasgow
ilya@stats.gla.ac.uk www.stats.gla.ac.uk/~ilya

Sergei Zuyev
Department of Statistics and Modelling Science
University of Strathclyde
sergei@stams.strath.ac.uk www.stams.strath.ac.uk/~sergei

Abstract The paper applies abstract optimisation principles in the space of measures within the context of optimal design problems. It is shown that within this framework it is possible to treat various design criteria and constraints in a unified manner providing a "universal" variant of the Kiefer-Wolfowitz theorem and giving a full spectrum of optimality criteria for particular cases. The described steepest descent algorithm uses the true direction of steepest descent and descends faster than the conventional sequential algorithms that involve renormalisation at every step.

Keywords: design of experiments, gradient methods, optimal design, regression, space of measures

Introduction

Consider a standard linear optimal design problem on a design space X (assumed to be a locally compact separable metric space) with regressors $f(x) = (f_1(x), \dots, f_k(x))$ and unknown coefficients $\beta = (\beta_1, \dots, \beta_k)$. The optimal design is described as a probability measure μ on X that determines the frequencies of making observations at particular points. The information

matrix is given by

$$M(\mu) = \int f(x)^\top f(x)\, \mu(dx)\,, \tag{8.1}$$

see, e. g., Atkinson and Donev (1992) for details. The measure μ that minimises $\det M^{-1}(\mu)$ is called the D-*optimal design measure*. To create a *convex* optimisation criterion it is convenient to minimise $\Psi(M) = -\log(\det M)$, which is a convex functional on the space of matrices. In general, a measure μ that minimises a given differentiable functional $\Psi(M(\mu))$ is called a Ψ-*optimal design measure*.

It is common to handle the above optimisation problem by first finding an optimal information matrix from the class defined by all admissible design measures and afterwards by relating it to the corresponding design measure. Schematically, such methods of finding the optimal solution involve the following two stages:

$$\Psi(M(\mu)) \quad \overset{\text{Step I}}{\longrightarrow} \quad M(\mu) \quad \overset{\text{Step II}}{\longrightarrow} \quad \mu\,. \tag{8.2}$$

Step I concerns optimisation in the Euclidean space of dimension $k(k+1)/2$, while Step II aims to identify the measure from the information matrix obtained in the previous step. However, the chain approach based on (8.2) cannot be easily amended when the formulation of the problem changes, e.g., when constraints on μ appear. This usually calls for major changes in proofs, since a new family of matrices has to be analysed in Step I. Therefore, for each new type of constraint on μ not only Step II, but also Step I, has to be reworked, as in Cook and Fedorov (1995) and Fedorov and Hackl (1997), for example, where a number of different types of constraint are analysed.

Although the family of non-negative measures is not a linear space, many objective functionals can be extended onto the linear space of *signed* measures as the definition (8.1) of the information matrix $M(\mu)$ applies literally to signed measures. This allows us to treat the design problems described above as a special case of the optimisation of functionals defined on signed measures. We show that this abstract setting naturally incorporates many specific optimal design problems and leads to new (apparently more efficient) steepest descent algorithms for the numerical computation of optimal designs.

Recall that the *directional* (or *Gateaux*) derivative of a real-valued functional ψ in a Banach space $\mathbb{E}$ is defined by

$$D\psi(x)[v] = \lim_{t \downarrow 0} t^{-1}(\psi(x+tv) - \psi(x))\,, \tag{8.3}$$

where $v \in \mathbb{E}$ is a direction (Hille and Phillips (1957)). The functional ψ is said to have a *Fréchet* (or strong) derivative if

$$\psi(x+v) = \psi(x) + D\psi(x)[v] + o(\|v\|) \quad \text{as } \|v\| \to 0\,,$$

where $D\psi(x)[v]$ is a bounded linear continuous functional of v. In the design context, ψ is a functional on the space of measures, such as $\psi(\mu) = \Psi(M(\mu))$, and $\|\cdot\|$ is a norm defined for signed measures, e. g. the total variation norm. Then the Fréchet derivative is written as $D\psi(\mu)[\eta]$, where η is a signed measure.

Clearly, if μ is a probability measure, then $\mu + t\eta$ as in (8.3) above is not necessarily a probability measure or even a positive measure. Therefore, $\psi(\mu + t\eta)$ may not have a direct interpretation within the design framework for an arbitrary η. To circumvent this difficulty, it is quite typical in the optimal design literature to replace (8.3) by the following definition of the directional derivative:

$$\tilde{D}\psi(\mu)[\eta] = \lim_{t\downarrow 0} t^{-1}(\psi((1-t)\mu + t\eta) - \psi(\mu)) \tag{8.4}$$

(Atkinson and Donev (1992); Wynn (1972)). Now $(1-t)\mu + t\eta$ is a probability measure for $t \in [0,1]$ if both μ and η are probability measures. This definition of the directional derivative is used to construct the steepest (with respect to the differential operator $\tilde{D}$) descent algorithms for finding optimal designs when it is not possible to obtain analytical solutions (Wynn (1972)). However, the descent related to (8.4) is only *asymptotically* equivalent to the true steepest descent as defined by D in (8.3). Indeed, it is easy to see from the definitions that

$$\tilde{D}\psi(\mu)[\eta] = D\psi(\mu)[\eta - \mu]\,.$$

Thus, the descent direction $\tilde{\eta}^*$ determined by minimising $\tilde{D}\psi(\mu)[\eta]$ over all η with the given norm is not the *true* steepest descent direction η^* obtained by minimising $D\psi(\mu)[\eta]$.

The steepest descent algorithm described in Section 3 emerges from our theoretical results on constrained optimisation in the space of measures presented in Section 1. In contrast to the classical sequential algorithms in the optimal design literature (Wu, 1978a, 1978b; Wu and Wynn (1978); Wynn (1970)), we do not renormalise the design measure obtained on each step. Instead, the algorithm adds a *signed* measure chosen to minimise the Fréchet derivative of the target function among all measures satisfying the imposed constraints. This extends the ideas of Atwood (1973, 1976) on the conventional optimisation framework. Working in the linear space of signed measures rather than on the cone of positive measures makes this algorithm just a special case of the general steepest descent algorithm known from the optimisation literature. To establish its required convergence properties it is now sufficient to refer to the general results for the steepest descent method (see Polak, 1997).

Section 1 surveys several necessary concepts of abstract optimisation for functionals defined on the cone of positive measures, describing a number of results adapted from Molchanov and Zuyev (1997), where proofs and further

generalisations can be found. Applications to optimal designs are discussed in Section 2. Section 3 is devoted to steepest descent type algorithms.

1. OPTIMISATION IN THE SPACE OF MEASURES

Let $\mathbb{M}$ be the family of all non-negative finite measures on a Polish space X and $\widetilde{\mathbb{M}}$ be the linear space of all signed measures with bounded total variation on its Borel sets. In numerical implementations X becomes a grid in a Euclidean space $\mathbb{R}^d$ and $\mu \in \mathbb{M}$ is a non-negative array indexed by the grid's nodes. For every $\eta \in \widetilde{\mathbb{M}}$, denote by η^+ (resp. η^-) the positive (resp. negative) part in the Jordan decomposition of η.

Consider the following optimisation problem with finitely many equality and inequality constraints:

$$\psi(\mu) \to \inf \tag{8.5}$$

subject to

$$\begin{cases} H_i(\mu) = 0\,, & i = 1, \dots, m\,; \\ H_j(\mu) \le 0\,, & j = m+1, \dots, k\,. \end{cases} \tag{8.6}$$

It is always assumed that $f : \widetilde{\mathbb{M}} \mapsto \mathbb{R}$ and $H : \widetilde{\mathbb{M}} \mapsto \mathbb{R}^k$ are Fréchet differentiable functions. Most differentiable functionals of measures met in practice have derivatives which can be represented in the integral form. We consider here only this common case, so that there exist measurable real-valued functions $d_\psi(x, \mu)$ and $h_i(x, \mu)$, $i = 1, \dots, k$, such that for all $\eta \in \widetilde{\mathbb{M}}$

$$D\psi(\mu)[\eta] = \int d_\psi(x, \mu)\eta(dx) \quad \text{and} \quad DH(\mu)[\eta] = \int h(x, \mu)\eta(dx)\,, \tag{8.7}$$

where $h = (h_1, \dots, h_k)$. Note that all integrals are over X unless otherwise specified. Furthermore, assume that the solution μ of Problem (8.5) is *regular*, i.e. the functions $h_1(\cdot, \mu), \dots, h_m(\cdot, \mu)$ are linearly independent and there exists $\eta \in \widetilde{\mathbb{M}}$ such that

$$\begin{cases} \int h_i(x, \mu)\eta(dx) = 0 & \text{for all } i = 1, \dots, m\,, \\ \int h_j(x, \mu)\eta(dx) < 0 & \text{for all } j \in \{m+1, \dots, k\} \text{ verifying } H_j(\mu) = 0. \end{cases} \tag{8.8}$$

(Without the inequality constraints, (8.8) holds trivially with η being the zero measure). The regularity condition above guarantees the existence and boundedness of the Lagrange multipliers (Zowe and Kurcyusz (1979)) for the problem (8.5).

Theorem 1 (see Molchanov and Zuyev (1997)**).** *Let μ be a regular local minimum of ψ subject to (8.6). Then there exists $u = (u_1, \dots, u_k)$ with $u_j \le 0$*

if $H_j(\mu) = 0$ *and* $u_i = 0$ *if* $H_j(\mu) < 0$ *for* $j \in \{m+1, \dots, k\}$*, such that*

$$\begin{cases} d_\psi(x,\mu) = u\, h(x,\mu)^\top & \mu - \textit{almost everywhere}\,, \\ d_\psi(x,\mu) \geq u\, h(x,\mu)^\top & \textit{for all } x \in X\,. \end{cases} \tag{8.9}$$

Example 1. The simplest (but extremely important) example concerns minimisation of $\psi(\mu)$ over all $\mu \in \mathbb{M}$ with the fixed total mass $\mu(X) = 1$. In the above framework, this corresponds to the case when $k = 1$ and the only constraint is $H_1(\mu) = \int \mu(dx) - 1 = 0$. Then (8.9) becomes

$$\begin{cases} d_\psi(x,\mu) = u & \mu - \text{almost everywhere}\,, \\ d_\psi(x,\mu) \geq u & \text{for all } x \in X\,. \end{cases} \tag{8.10}$$

In the sequel we call (8.10) a necessary condition for a minimum in the fixed total mass problem.

2. APPLICATIONS TO OPTIMAL DESIGN

Optimal design problems can be naturally treated within the general framework described above of optimisation of functionals defined on finite measures. Theorem 1 directly applies to functionals of measures typical in the optimal design literature and under quite general differentiable constraints. Although we do not make any convexity assumptions on ψ and work exclusively with necessary optimal conditions, convexity of ψ immediately ensures the existence of the optimal design.

Theorem 1 can easily be specialised to functionals ψ that effectively depend on the information matrix. Given a function $\Psi : \mathbb{R}^{k^2} \mapsto \mathbb{R}$, a *$\Psi$-optimal design measure* μ minimises $\psi(\mu) = \Psi(M(\mu))$ over all probability measures on X with possibly some additional constraints. Direct computation of the derivative of ψ leads to the following result that can be considered as a generalised Kiefer–Wolfowitz theorem.

Corollary 2. *Let* μ *provide a* Ψ*-optimal design subject to constraints (8.6). Then (8.9) holds with*

$$d_\psi(x,\mu) = f(x) D\Psi(M)(\mu) f^\top(x)$$

and

$$D\Psi(M)(\mu) = \left\| \frac{\partial \Psi(M)}{\partial m_{ij}} \right\|_{ij} (\mu)\,,$$

where $m_{ij} = \int f_i(x) f_j(x)\, \mu(dx)$ *is the* (i,j)*-th entry of the information matrix.*

Example 2. *(*D*-optimal design)* Put $\Psi(M) = -\log\det M$. Then $D\Psi(M) = -M^{-1}$, the only constraint if $H(\mu) = \mu(X) - 1$ has the derivative $h = 1$. Since $d_\psi(x,\mu) = -f(x)M^{-1}(\mu)f^\top(x) = -d(x,\mu)$, Theorem 1 yields the necessary condition in the Kiefer-Wolfowitz characterisation of D-optimal designs.

Example 3. *(*D*-optimal design with fixed moments)* Let $X = \mathbb{R}^d$ and assume that along with the constraint on the total mass $\mu(X) = 1$ we fix the expectation of μ, which is a vector $m = \int x\mu(dx)$. These constraints can be written as $H(\mu) = (m, 0)$, where $H(\mu)$ is a $(d+1)$-dimensional vector function with the components $H_i(\mu) = \int x_i\mu(dx)$ for $i = 1, \ldots, d$ and $H_{d+1}(\mu) = \mu(X) - 1$. Clearly, (8.7) holds with $h(x,\mu) = (x, 1)$. By Theorem 1, if μ minimises $\Psi(\mu)$ under the condition $H(\mu) = (m, 0)$, then

$$\begin{cases} d_\psi(x,\mu) = v\ x^\top + u & \mu - \text{almost everywhere}\,, \\ d_\psi(x,\mu) \geq v\ x^\top + u & \text{for all } x \in X\,, \end{cases}$$

for some $v \in \mathbb{R}^d$. In other words, $d(x,\mu)$ is *affine* for μ-almost all x.

Example 4. *(*D*-optimal designs with a limited cost)* Let $h(x)$ determine the cost of making an observation at point x. If the expected total cost is bounded by C, then the design measure μ should, in addition to $\mu(X) = 1$, satisfy the constraint $\int h(x)\mu(dx) \leq C$. If μ provides a Ψ-optimal design under such constraints, then

$$\begin{cases} d_\psi(x,\mu) = u + wh(x) & \mu - \text{a.e.}\,, \\ d_\psi(x,\mu) \leq u + wh(x) & \text{for all } x \in X\,, \end{cases}$$

for some $u \in \mathbb{R}$ and $w \leq 0$ if $\int h(x)\mu(dx) = C$ and $w = 0$ if $\int h(x)\mu(dx) < C$.

Since the function $\psi(\mu) = \Psi(M(\mu))$ in the D-optimality problem is a convex function of μ, the above necessary conditions become necessary and sufficient conditions, thus providing a full characterisation of the optimal designs. Other examples, like D-optimal designs with bounded densities, design on product-spaces, *etc.*, can be treated similarly.

3. GRADIENT METHODS

Direction of steepest descent. The gradient descent method relies on knowledge of the derivative of the target functional. As before, it is assumed that $\psi(\mu)$ is Fréchet differentiable and its derivative has the integral representation (8.7). We first consider the general setup, which will be specialised later for particular constraints and objective functionals.

The most basic method of the gradient descent type used in optimal design suggests moving from μ_n (the approximation at step n) to $\mu_{n+1} = (1-\alpha_n)\mu_n + \alpha_n\zeta_n$, where $0 < \alpha_n < 1$ and ζ_n minimises $\tilde{D}\Psi(\mu)[\zeta]$ over all probability measures ζ (Wynn (1970)). It is easy to see that such ζ_n is concentrated at the points where the corresponding gradient function $d_\psi(x,\mu)$ is minimised. Rearranging the terms, we obtain

$$\mu_{n+1} = \mu_n + \alpha_n(\zeta_n - \mu_n)\,. \tag{8.11}$$

In this form the algorithm looks like a conventional descent algorithm that moves along the direction $\eta_n = \alpha_n(\zeta_n - \mu_n)$. In this particular case, the step size is $\|\eta_n\| \le 2\alpha_n$, and $\eta_n = \eta_n^+ - \eta_n^-$ with

$$\eta_n^+ = \alpha_n\zeta_n\,, \quad \eta_n^- = \alpha_n\mu_n\,.$$

Such a choice of η_n ensures that $\mu_n + \eta_n$ remains a probability measure since the negative part of η_n is proportional to μ_n with $\alpha_n < 1$.

However, if we do not restrict the choice for the descent direction to measures with the negative part proportional to μ_n, then it is possible to find a steeper descent direction $\eta_n = \eta$ than the one suggested by (8.11). If the current value $\mu_n = \mu$, then the steepness of the descent direction η is commonly characterised by the value of the derivative

$$D\psi(\mu)[\eta] = \int d_\psi(x,\mu)\eta(dx)\,.$$

The true steepest descent direction must be chosen to minimise $D\psi(\mu)[\eta]$ over all signed measures η with total variation $\|\eta\| \le \varepsilon = 2\alpha_n$ such that $\mu + \eta$ belongs to the family $\mathbb{M}$ of non-negative measures and satisfies all specified constraints. For instance, if the only constraint is that the total mass $\mu(X) = 1$, then any signed measure η with $\eta(X) = 0$ and negative part η^- dominated by μ would ensure that $\mu + \eta$ stays in $\mathbb{M}$ and maintains the required total mass.

Consider the problem with only linear constraints of equality type:

$$H_i(\mu) = \int h_i(x)\mu(dx) = a_i\,, \quad i = 1,\dots,k\,, \tag{8.12}$$

where $a_1,\dots,a_k$ are given real numbers. In vector form, $H(\mu) = \int h(x)\mu(dx) = A$, where $H = (H_1,\dots,H_k)$, $h = (h_1,\dots,h_k)$ and $A = (a_1,\dots,a_k)$ implying that $DH(\mu)[\eta] = \int h(x)\eta(dx)$.

For a $\mu \in \mathbb{M}$ denote by Υ_μ the family of all signed measures $\eta \in \widetilde{\mathbb{M}}$ such that $\mu + \eta \in \mathbb{M}$ and $\mu + \eta$ satisfies the constraints (8.12). The family Υ_μ represents admissible directions of descent. Let $\mu|_B$ denote the restriction of a measure μ onto a Borel set B, i.e. $\mu|_B(\cdot) = \mu(\cdot \cap B)$. The following results, given without

proof, characterise the steepest direction. Recall that vectors $w_1, \ldots, w_{k+1}$ are called *affinely independent* if $w_2 - w_1, \ldots, w_{k+1} - w_1$ are linearly independent.

Theorem 3 (see Molchanov and Zuyev (2000)**).** *The minimum of $D\psi(\mu)[\eta]$ over all $\eta \in \Upsilon_\mu$ with $\|\eta\| \le \varepsilon$ is achieved on a signed measure $\eta = \eta^+ - \eta^-$ such that η^+ has at most k atoms and $\eta^- = \sum_{i=1}^{k+1} t_i \, \mu|_{B_i}$ for some $0 \le t_i \le 1$ with $t_1 + \cdots + t_{k+1} = 1$ and some measurable sets B_i such that vectors $H(\mu|_{B_i}),\ i = 1, \ldots, k+1$, are affinely independent.*

Corollary 4. *If the only constraint is $\mu(X) = 1$, then the minimum of $D\psi(\mu)[\eta]$ over all $\eta \in \Upsilon_\mu$ with $\|\eta\| \le \varepsilon$ is achieved on a signed measure η such that η^+ is the positive measure of total mass $\varepsilon/2$ concentrated on the points of the global minima of $d_\psi(x, \mu)$ and $\eta^- = \mu|_{M(t_\varepsilon)} + \kappa\mu|_{M(s_\varepsilon) \setminus M(t_\varepsilon)}$, where*

$$M(p) = \{x \in X \,:\, d_\psi(x, \mu) \ge p\}\,,$$

and

$$t_\varepsilon = \inf\{p :\ \mu(M(p)) < \varepsilon/2\}\,, \tag{8.13}$$

$$s_\varepsilon = \sup\{p :\ \mu(M(p)) \ge \varepsilon/2\}\,. \tag{8.14}$$

The factor κ is chosen to ensure that $\mu(M(t_\varepsilon)) + \kappa\mu(M(s_\varepsilon) \setminus M(t_\varepsilon)) = \varepsilon/2$.

It is interesting to note that, without the constraint $\|\eta\| \le \varepsilon$ on the total variation norm of the increment measure, the steepest direction preserving the total mass is the measure $\eta = \delta_{x_0} - \mu$, where δ_{x_0} is the unit measure concentrated on a global minimum point x_0 of $d_\psi(x, \mu)$. Thus the classical algorithm based on the modified directional derivative (8.4) uses, in fact, the scaled variant of this direction.

Steepest descent in the fixed total mass problem. The necessary condition given in Corollary 2 can be used as a stopping rule for descent algorithms. Corollary 4 above describes the steepest descent direction in the minimisation problem with a fixed total mass, justifying the following algorithm:

Procedure `go.steep`

Input. Initial measure μ.
Step 0. Compute $\psi \leftarrow \psi(\mu)$.
Step 1. Compute $g \leftarrow d_\psi(x, \mu)$. If `is.optim`$(\mu,\ g)$, stop.
Otherwise, choose the step size ε.
Step 2. Compute $\mu_1 \leftarrow$ `take.step`(ε, μ, g).
Step 3. If $\psi_1 \leftarrow \psi(\mu_1) < \psi$, then $\mu \leftarrow \mu_1$; $\psi \leftarrow \psi_1$; and go to Step 2.
Otherwise, go to Step 1.

The necessary condition for the optimum, that is used as a stopping rule in Step 1 above, is given by (8.10): the function $g(x) = d_\psi(x, \mu)$ is constant and takes its minimal value on the support of an optimal μ. Strictly speaking, the support of μ on the *discrete* space X is the set $S = \{x \in X : \ \mu(x) > 0\}$. But in practice one may wish to ignore the atoms of mass less than a certain small threshold `supp.tol`. The boolean procedure `is.optim` has the following structure:

Procedure `is.optim`

Input. Measure μ, gradient function $g(x)$, tolerance `tol`, `supp.tol`.
Step 1. Determine support S of μ up to tolerance `supp.tol`.
Step 2. If $\max_{x \in S} g(x) - \min g(x) <$`tol` return `TRUE`, and otherwise `FALSE`.

The procedure `take.step` returns the updated measure $\mu_1 = \mu + \eta$, where η is an increment measure with total mass 0 and variance ε along the steepest direction described in Corollary 4.

Procedure `take.step`

Input. Step size ε, measure μ, gradient function $g(x)$.
Step 0. Put $\mu_1(\{x\}) \leftarrow \mu(\{x\})$ for each point $x \in X$.
Step 1. Find a point x_0 of the global minimum of $g(x)$ and put $\mu_1(\{x_0\}) \leftarrow \mu(\{x_0\}) + \varepsilon/2$.
Step 2. Find t_ε and s_ε from (8.13) and (8.14), put $\mu_1(\{x\}) = 0$ for all $x \in M(t_\varepsilon)$, decrease the total μ_1-mass of the points from $M(s_\varepsilon) \setminus M(t_\varepsilon)$ by value $\varepsilon/2 - \mu(M(t_\varepsilon))$ and return the obtained μ_1.

In Step 1 above the mass $\varepsilon/2$ can also be spread uniformly or in any other manner over the points of the global minimum of $g(x)$. The described algorithm leads to a global minimum when applied to convex objective functions. In the general case it may stuck in a local minimum, the feature common for gradient algorithms applied in the context of global optimisation.

There are many possible methods suitable to choose the step size ε in Step 1 of procedure `go.steep`. Many aspects can be considered: the previous step size and/or difference between the supremum and infimum of $d_\psi(x, \mu)$ over the support of μ. The Armijo method widely used for general gradient descent algorithms (Polak (1997)) defines the new step size to be $\beta^m \varepsilon$, the integer m is such that

$$\psi(\mu + \eta_m) - \psi(\mu) \le \alpha \int d_\psi(x, \mu) \eta_m(dx)\,,$$
$$\psi(\mu + \eta_{m-1}) - \psi(\mu) > \alpha \int d_\psi(x, \mu) \eta_{m-1}(dx)\,,$$

where $0 < \alpha < 1$ and η_m is the steepest descent measure with the total variation $\beta^m \varepsilon$ described in Corollary 4.

The corresponding steepest descent algorithms are applicable for general differentiable objective functions and realised in `Splus/R` library `mefista` (for MEasures with FIxed mass STeepest Ascent/descent) from bundle `mesop` that can be obtained from the authors' web-pages. The chosen optimality criterion and the regressors should be passed as arguments to the descent procedure using (appropriately coded) functions ψ and d_ψ.

A numerical example. *Cubic regression through the origin.* Let $X = [0, 1]$ and let $f = (x, x^2, x^3)$. For the D-optimal design in this model an exact solution is available that assigns equal weights to three points $(5 \pm \sqrt{5})/10$ and 1 (Atkinson and Donev (1992), p.119). The initial measure is taken to be uniform on $[0, 1]$, discretised with grid size 0.01. In the algorithm described in (Atkinson and Donev (1992)) and further referred to as A0, a mass equal to the step size is added into the point of minimum of the gradient function $d_\psi(x, \mu)$, and the measure is then rescaled to the original total mass. The algorithm A1 is based on Corollary 4 and described above.

We run both algorithms A0 and A1 until either the objective function ψ no longer decreases or the stopping condition is satisfied: the standardised response function attains its minimum and is constant (within the predetermined tolerance level of 0.01) on the support of the current measure.

It takes 28 steps for our algorithm A1 to obtain the following approximation to the optimal measure: $\mu(0.27) = 0.2088645$, $\mu(0.28) = 0.1247447$, $\mu(0.72) = 0.2239635$, $\mu(0.73) = 0.1090888$, $\mu(1) = 0.3333385$. Note that the distribution of the mass over the nearest grid points corresponds exactly to the true atoms positioned at irrational points $(5 \pm \sqrt{5})/10$, since $\mu(0.27) + \mu(0.28) = 0.3336092$ and $\mu(0.72) + \mu(0.73) = 0.3330523$. The final step size is 0.00078125, $\psi(\mu) = -1.184291 \cdot 10^{-5}$, and the range of $d_\psi(x, \mu)$ on the support of μ is $0.006407306 <$`tol`$= 0.01$. Thus, algorithm A1 achieves the optimal design measure. In contrast, it takes 87 steps for A0 to arrive at a similar result (which still has a number of support points with negligible mass, but not true zeroes). Despite the fact that our algorithm makes more calculations at each step; it took 1.96 seconds of system time on a Pentium-II PC with 256 Mb RAM running at 450 MHz under `Linux` to finalise the task, compared to 3.70 seconds for A0. The difference becomes even more spectacular for smaller tolerance levels or finer grids.

Multiple linear constraints. Although the steepest direction for optimisation with many linear constraints given by (8.12) is characterised in Theorem 3, its practical determination becomes a difficult problem. Indeed, it is easy to see that for a discrete space X (used in numerical methods) minimisation of $D\psi(\mu)[\eta]$ over all signed measures $\eta \in \Upsilon_\mu$ with $\|\eta\| = \varepsilon$ is a linear programming problem of dimension equal to the cardinality of X. Therefore, in the

presence of many constraints, it might be computationally more efficient to use an approximation to the exact steepest direction.

For instance, it is possible to fix the negative component of the increment η at every step and vary its positive part η^+. Due to Theorem 3, the positive part $\nu = \eta^+$ of the steepest increment measure consists of at most k atoms. With the negative part η^- being proportional to μ, we suggest moving from the current measure μ to $\mu + \eta$ where $\eta = \nu - \gamma\mu$ for some $\gamma > 0$. This is equivalent to renormalisation $c(\mu + \nu')$ with $c = (1 - \gamma)$ and $\nu' = c^{-1}\nu$. The measure ν has k atoms of masses $p_1, \ldots, p_k$ located at points $x_1, \ldots, x_k$ chosen to minimise the directional derivative $D\psi(\mu)[\eta]$ (or, equivalently, $D\psi(\mu)[\nu]$) and to maintain the imposed constraints (8.12). The value of γ characterises the size of the step, although it is not equal to the total variation of η.

To satisfy the linear constraints $H(\mu + \nu - \gamma\mu) = A = (a_1, \ldots, a_k)$ we need to have

$$H(\nu) = \sum_{j=1}^{k} p_j h(x_j) = \gamma A\,,$$

which can be written in matrix form as

$$\mathbb{H}(x_1, \ldots, x_k)p^\top = \gamma A^\top$$

with $p = (p_1, \ldots, p_k)$ and $\mathbb{H}(x_1, \ldots, x_k) = [h_i(x_j)]_{i,j=1}^k$. This implies

$$p^\top = \gamma \mathbb{H}(x_1, \ldots, x_k)^{-1} A^\top\,. \tag{8.15}$$

Since $\eta = \nu - \gamma\mu$, the directional derivative $D\psi(\mu)[\eta]$ is minimised if ν minimises

$$D\psi(\mu)[\nu] = \gamma d(x_1, \ldots, x_k)\mathbb{H}(x_1, \ldots, x_k)^{-1} A^\top\,, \tag{8.16}$$

where $d(x_1, \ldots, x_k) = (d_\psi(x_1, \mu), \ldots, d_\psi(x_k, \mu))$. The right-hand side of (8.16) is a function of k variables $x_1, \ldots, x_k$ that should be minimised to find the 'optimal' locations of the atoms. Their masses $p_1, \ldots, p_k$ are determined by (8.15). Note that minimisation is restricted to only those k-tuples $x_1, \ldots, x_k$ that provide $p^\top$ with all non-negative components.

Since this descent differs from the *steepest* descent given in Theorem 3, an additional analysis is necessary to ensure that the algorithm described does converge to the desired solution. By using the same arguments as in Wu and Wynn (1978), it is possible to prove the dichotomous theorem for this situation.

The descent algorithm for many constraints based on renormalisation procedure has been programmed in the `Splus` and `R` languages. The corresponding library `medea` (for MEasure DEscent/Ascent) from bundle `mesop` and related examples can again be obtained from the authors' web-pages.

In the case of a single constraint on the total mass the described approach turns into a renormalisation method: ν has a single atom that is placed at the point of the global minimum of $d_\psi(\cdot, \mu)$ to minimise $D\psi(\mu)[\nu]$.

References

Atkinson, A. C. and Donev, A. N. (1992). *Optimum Experimental Designs*. Oxford: Clarendon Press.

Atwood, C. L. (1973). Sequences converging to D-optimal designs of experiments. *Ann. Statist* **1**, 342–352.

Atwood, C. L. (1976). Convergent design sequences, for sufficiently regular optimality criteria. *Ann. Statist.* **4**, 1124–1138.

Cook, D. and Fedorov, V. (1995). Constrained optimization of experimental design. *Statistics* **26**, 129–178.

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

Fedorov, V. V. and Hackl, P. (1997). *Model-Oriented Design of Experiments*, Volume 125 of *Lect. Notes Statist.* New York: Springer-Verlag.

Ford, I. (1976). *Optimal Static and Sequential Design: a Critical Review*. PhD Thesis, Department of Statistics, University of Glasgow.

Hille, E. and Phillips, R. S. (1957). *Functional Analysis and Semigroups*, Volume XXXI of *AMS Colloquium Publications*. Providence, RI: American Mathematical Society.

Molchanov, I. and Zuyev, S. (1997). Variational analysis of functionals of a Poisson process. *Rapport de Recherche* 3302, INRIA, Sophia-Antipolis. To appear in *Math. Oper. Res.*

Molchanov, I. and Zuyev, S. (2000). Steepest descent algorithms in space of measures. To appear.

Polak, E. (1997). *Optimization*. New York: Springer-Verlag.

Wu, C.-F. (1978a). Some algorithmic aspects of the theory of optimal design. *Ann. Statist.* **6**, 1286–1301.

Wu, C.-F. (1978b). Some iterative procedures for generating nonsingular optimal designs. *Comm. Statist. Theory Methods* A7 **14**, 1399–1412.

Wu, C.-F. and Wynn, H. P. (1978). The convergence of general step-length algorithms for regular optimum design criteria. *Ann. Statist.* **6**, 1273–1285.

Wynn, H. P. (1970). The sequential generation of D-optimum experimental designs. *Ann. Math. Statist.* **41**, 1655–1664.

Wynn, H. P. (1972). Results in the theory and construction of D-optimum experimental designs. *J. Roy. Statist. Soc.* B **34**, 133–147.

Zowe, J. and Kurcyusz, S. (1979). Regularity and stability for the mathematical programming problem in Banach spaces. *Appl. Math. Optim.* **5**, 49–62.

Chapter 9

ON THE EFFICIENCY OF GENERALLY BALANCED DESIGNS ANALYSED BY RESTRICTED MAXIMUM LIKELIHOOD

H. Monod
Unité de Biométrie, INRA-Versailles
F78026 Versailles Cedex, France
monod@versailles.inra.fr

Abstract Restricted maximum likelihood (reml) is commonly used in the analysis of incomplete block designs. With this method, treatment contrasts are estimated by generalized least squares, using an estimated variance-covariance matrix of the observations as if it were known. This leads to under-estimation of the variance of treatment contrasts, because uncertainty on the variance components is not adequately taken into account. To correct for this bias, Kackar and Harville (1984) and Kenward and Roger (1997) propose adjusted estimators of the treatment variance-covariance matrix, based on Taylor series expansions.
We consider small experiments with an orthogonal block structure. The adjusted estimator of Kenward and Roger (1997) is calculated when the design is generally balanced. A small modification is proposed that leads to a simple expression for the adjustment, as a function of the efficiency factors of the design, the variance components and the dimensions of the block strata. The behaviour of the adjusted estimator is assessed through a simulation study based on a semi-Latin square for twelve treatments.

Keywords: generally balanced design, restricted maximum likelihood

Introduction

Residual maximum likelihood (REML) is widely used for the analysis of designed experiments with one or more blocking factors. For these applications, block effects are usually considered as random, whereas treatment effects are considered as fixed. When treatments are non-orthogonal to blocks, REML provides an efficient and convenient way to estimate the variance components

and to combine information from several block strata to estimate treatment effects.

When the REML method is applied to a mixed model, variance components are estimated first (by residual maximum likelihood), and then fixed treatment effects are estimated by generalized least squares, using the estimated variance-covariance matrix of the observations as if it were known. Usually, the bias and variance of this matrix are not taken into account, causing under-estimation of the variance of fixed treatment effects (see Searle, Casella and McCulloch, 1992, p.320).

A recent paper by Kenward and Roger (1997) presents a coherent framework for drawing inferences on fixed effects, when using REML with small samples. In particular, Kenward and Roger (1997) give an adjusted formula for the estimation of the variance of fixed effects. In this paper, we apply this general formula to designed experiments. We consider designs with an orthogonal block structure, and we show that the properties of these designs lead to a simplified formula. Additional simplifications occur when the design is generally balanced in the sense of Houtman and Speed (1985).

The results have several applications. Firstly, they give a better understanding of how the bias depends on the size of the experiment, the degree of imbalance and the actual values of the variance components. Secondly, they can help to compare more accurately small-size experimental designs with different block structures. The work presented in this paper originated from a study to compare the efficiency of complete block designs, incomplete block designs and semi-Latin squares for sunflower variety trials. To assess the behaviour of the adjusted variance estimator for such applications, a simulation study is performed on a semi-Latin square for twelve treatments in four replications.

1. SMALL-SAMPLE INFERENCE FOR FIXED EFFECTS USING REML

In this section, we give the results in Kenward and Roger (1997) which are of interest for the present paper, making only a slight change for the design context. We consider a vector Y of N observations which follows a multivariate Gaussian distribution,

$$Y \sim \mathcal{N}(X\theta; \Sigma), \tag{9.1}$$

where X $(N \times v)$ is the design matrix for the treatments, θ $(v \times 1)$ is the vector of fixed treament effects and Σ $(N \times N)$ is the variance-covariance matrix which depends on a vector ξ of s unknown parameters. It is necessary to assume that the first two partial derivatives of Σ with respect to ξ exist. In this paper, Σ depends on ξ linearly, so that the second-order partial derivatives are zero.

The REML estimator $\widehat{\xi}$ maximizes that part of the (log-)likelihood which does not depend on the fixed effects, and which can be written, up to a constant,

as

$$-\log(|\Sigma|) - \log\left(\left|X'\Sigma^{-1}X\right|\right) - Y'\left(\Sigma^{-1} - \Sigma^{-1}X(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}\right)Y. \tag{9.2}$$

Let $\widehat{\Sigma}$ denote the corresponding estimate of Σ. Then the REML estimator $\widehat{\theta}$ of the fixed effects is equal to

$$\widehat{\theta} = \widehat{\Phi}X'\widehat{\Sigma}^{-1}Y, \tag{9.3}$$

with

$$\Phi(\xi) = (X'\Sigma(\xi)^{-1}X)^{-1} \text{and } \widehat{\Phi} = \Phi(\widehat{\xi}) = (X'\widehat{\Sigma}^{-1}X)^{-1}. \tag{9.4}$$

The matrix $\widehat{\Phi}$ is commonly used as an estimate of the variance-covariance matrix of $\widehat{\theta}$. However, it is biased for small samples and tends to overestimate the precision on the fixed effects. From previous results by Kackar and Harville (1984) based on Taylor series expansions, Kenward and Roger (1997) give the following approximate expression for the variance-covariance matrix of treatment effects,

$$\Phi_A = \Phi + 2\Lambda, \tag{9.5}$$

where

$$\Lambda = \Phi\left\{\sum_{i=1}^{r}\sum_{j=1}^{r} W_{ij}(Q_{ij} - P_i\Phi P_j)\right\}\Phi, \tag{9.6}$$

$$P_i = X'\frac{\partial\Sigma^{-1}}{\partial\xi_i}X, \quad Q_{ij} = X'\frac{\partial\Sigma^{-1}}{\partial\xi_i}\Sigma\frac{\partial\Sigma^{-1}}{\partial\xi_j}X \text{ and} \tag{9.7}$$

$$W = I_E^{-1} \text{ with } 2(I_E)_{ij} \tag{9.8}$$

$$= \operatorname{tr}\left(\frac{\partial\Sigma^{-1}}{\partial\xi_i}\Sigma\frac{\partial\Sigma^{-1}}{\partial\xi_j}\Sigma\right) - \operatorname{tr}\left(2\Phi Q_{ij} - \Phi P_i\Phi P_j\right). \tag{9.9}$$

The matrix W corresponds to the variance-covariance matrix of $\widehat{\xi}$, approximated by the inverse of the expected information matrix I_E. In the calculations, possible bias in the estimation of ξ is neglected. The adjusted estimator $\widehat{\Phi_A}$ is then obtained by plugging $\widehat{\xi}$ in the expressions above.

2. GENERALLY BALANCED DESIGNS

2.1. ORTHOGONAL BLOCK STRUCTURE

We consider experimental designs with an orthogonal block structure (Bailey, 1984). This covers a large proportion of the designs used in practice. In particular, this includes the block structures made of orthogonal factors with crossing and nesting relationships (Nelder, 1965).

The orthogonal block structure of a design determines an orthogonal direct-sum decomposition of $\Re^N$ into s subspaces V_i, $i = 0, \ldots, s-1$, called strata. One stratum, V_0 say, is generated by the all-one vector. The associated model is (9.1), with

$$\Sigma = \sum_{i=0}^{s-1} \xi_i S_i, \tag{9.10}$$

where S_i is the orthogonal projection matrix on V_i and the ξ_is are unknown positive variance parameters. The S_i matrices are idempotent and mutually orthogonal. Thus, $S_i S_j = \delta_{i=j} S_i$, where $\delta_{i=j}$ equals 1 if $i = j$ and equals 0 otherwise.

The mixed model defined by equations (9.1) and (9.10) is partly validated by randomization. Indeed, adequate randomization guarantees that the first and second moments of Y are as given by this model (Bailey, 1991; Bardin et Azaïs, 1991). The only constraint on the ξ_i parameters is that they be positive.

Alternatively, when the block structure is factorial, equation (9.10) can be obtained by assuming that blocks have random factorial effects. However, this introduces more restrictive constraints on the variance parameters. Consider for instance a block design with blocks of size k ; in that case, there are two strata (in addition to the stratum associated with the all-one vector), the within-block (V_1, say) and the between-block (V_2) strata. If block effects are assumed to be random with variance σ_B^2 and if σ_E^2 denotes the residual variance, then $\xi_1 = \sigma_E^2$ and $\xi_2 = k\sigma_B^2 + \sigma_E^2$, which implies that $\xi_2 \geq \xi_1$.

2.2. GENERAL BALANCE

We now consider designs which are equireplicate and generally balanced in the sense of Houtman and Speed (1983) and Bailey (1993). For such designs, the treatment space can be decomposed into the orthogonal direct sum of n subspaces $\mathcal{T}_a$ which are common eigenspaces of the matrices $X'S_iX$. Thus, there exist scalars e_{ia} such that

$$X'S_iX = r \sum_{a=0}^{n-1} e_{ia} T_a, \tag{9.11}$$

where r is the number of replicates and T_a is the orthogonal projector on $\mathcal{T}_a$. One treatment subspace, $\mathcal{T}_0$ say, is generated by the all-one vector and is associated with the treatment mean, whereas the other strata are associated with treatment contrasts.

The coefficient e_{ia} is called the efficiency factor of the design for treatment subspace $\mathcal{T}_a$ in the stratum V_i, with the efficiency factors satisfying $\sum_i e_{ia} = 1$. In addition, $e_{0,0} = 1$ and $e_{ia} = 0$ if one and only one of i or a is zero.

Example 1: In a balanced incomplete block design, there are three strata (0: mean; 1: within-block and 2: between-block), two treatments subspaces $\mathcal{T}_a$ (0: mean; 1: contrasts) and $e_{1,1}$ is the classical efficiency factor of the design.

Example 2: In an orthogonal design, for example a complete block design, a Latin square or an orthogonal split-plot design, each treatment subspace $\mathcal{T}_a$ is estimated in a single stratum, so that for each a, all the efficiency factors e_{ia} are zero except one which is equal to one.

Less trivial examples are given by Bailey (1993) and in Section 4

3. VARIANCE OF ESTIMATED TREATMENT EFFECTS IN A GENERALLY BALANCED DESIGN

We can now combine the results in the previous two sections, to get a bias-corrected estimator $\widehat{\Phi_A}$ of the variance of treatment effects, when the design is generally balanced. This will generalize the results given by Kackar and Harville (1984) for balanced incomplete block designs.

The equations (9.5) to (9.9) simplify when the design has an orthogonal block structure and is generally balanced. Indeed, the properties of the S_i matrices imply that $\Sigma^{-1} = \sum_i \xi_i^{-1} S_i$, $P_i = -\xi_i^{-2} X' S_i X$, and $Q_{ij} = \delta_{i=j} \xi_i^{-3} X' S_i X$. In addition, from equation (9.11) we get

$$X'\Sigma^{-1}X = r \sum_{a=0}^{n-1} \left(\sum_{i=0}^{s-1} \xi_i^{-1} e_{ia} \right) T_a \quad \text{and} \tag{9.12}$$

$$\Phi = (1/r) \sum_{a=0}^{n-1} \left(\sum_{i=0}^{s-1} \xi_i^{-1} e_{ia} \right)^{-1} T_a. \tag{9.13}$$

We will use the following notation: $f_{ia} = \xi_i^{-1} e_{ia}$ and $h_{ia} = f_{ia}/f_{+a}$, where $f_{+a} = \sum_i f_{ia}$. Thus, $r f_{+a}$ is the overall information on any normed contrast in $\mathcal{T}_a$ and h_{ia} is the proportion of that information coming from stratum V_i.

Let $\nu_i, i = 0, \cdots, s-1$, denote the strata dimensions, with $\mu_a, a = 0, \cdots, n-1$, denoting the dimensions of the $\mathcal{T}_a$ subspaces. One intermediate result gives

$$\Phi = (1/r) \sum_a f_{+a}^{-1} T_a,$$

$$\Lambda = (1/r) \sum_a f_{+a}^{-1} \left(\sum_i W_{ii} \xi_i^{-2} h_{ia} - \sum_i \sum_j W_{ij} \xi_i^{-1} \xi_j^{-1} h_{ia} h_{ja} \right) T_a$$

and

$$(I_E)_{ij} = (1/2) \xi_i^{-1} \xi_j^{-1} \left(\delta_{i=j} (\nu_i - 2 \sum_a h_{ia} \mu_a) + \sum_a h_{ia} h_{ja} \mu_a \right).$$

Let M denote a normalized version of I_E, that is, the matrix such that $(M)_{ij} = (1/2)\left(\delta_{i=j}(\nu_i - 2\sum_a h_{ia}\mu_a) + \sum_a h_{ia}h_{ja}\mu_a\right)$. It may happen that M is not invertible, if one or more strata are saturated by treatment effects and all the information on these treatment effects is concentrated in these strata. However, M will usually be invertible and we assume it is in the sequel. Let W^* denote the inverse of M. Then we obtain the following expression for the approximate variance-covariance matrix of treatment effects

$$\Phi_A = \frac{1}{r}\sum_{a=0}^{n-1} f_{+a}^{-1}(1+d_a)T_a, \tag{9.14}$$

$$\text{with } d_a = 2\left(\sum_{i=1}^{s} W_{ii}^* h_{ia} - \sum_{i=1}^{s}\sum_{j=1}^{s} W_{ij}^* h_{ia}h_{ja}\right). \tag{9.15}$$

Thus the approximate variance-covariance matrix of treatment contrasts is a linear combination of the T_a matrices. The usual estimate is obtained when the d_as are neglected, so d_a represents the small-sample adjustment for the treatment subspace $\mathcal{T}_a$ as a proportion of the usual estimate. If $\mathcal{T}_a$ is estimated in a single stratum, then there is no recovery of information between strata and we find, logically, that $d_a = 0$.

A simpler expression is obtained for d_a if the non-diagonal terms of W^* are neglected, which yields

$$d_a = 4\sum_{i=0}^{s-1} \frac{h_{ia}(1-h_{ia})}{\nu_i^*}, \tag{9.16}$$

where $\nu_i^* = \nu_i - \sum_a \mu_a h_{ia}(2-h_{ia})$ represents the effective number of degrees of freedom in V_i which are available for the estimation of ξ_i.

Notice that, for a given design with $e_{ia} \in (0,1)$ and depending on the values of the ξ_i parameters, $4h_{ia}(1-h_{ia})$ can vary within $(0,1)$, whereas ν_i^* can vary within $(\nu_i - \mu_{(i)}, \nu_i)$, where $\mu_{(i)}$ is the number of treatment degrees of freedom which are estimable in stratum V_i. An extreme situation occurs when a stratum is saturated by treatment effects ($\nu_i = \mu_{(i)}$) and provides all the information on these treatment effects ($h_{ia} = 1$). Then, the variance component ξ_i is confounded with treatment effects and is not estimable. This is always the case for the parameter ξ_0, which is confounded with the treatment mean.

4. EXAMPLE

A semi-Latin square of size $(n \times n)/k$ is a design for $v = nk$ treatments in n rows and n columns, with k plots at the intersection of each row and each plot (Bailey, 1992). In a semi-Latin square, each row and each column is a

complete replicate. Trojan designs are optimal semi-Latin squares constructed by juxtaposing k ($k \leq n-1$) mutually orthogonal Latin squares of order n.

We consider a trojan design of size $(4 \times 4)/3$ for 12 treatments (Fig. 9.1). The efficiency factors are given in Table 9.1. To illustrate the behaviour of the adjusted estimator given by (9.14) and (9.16), we present the results of simulations performed for this Trojan design, using Splus.

A	α	a	B	β	b	C	γ	c	D	δ	d
B	γ	d	A	δ	c	D	α	b	C	β	a
C	δ	b	D	γ	a	A	β	d	B	α	c
D	β	c	C	α	d	B	δ	a	A	γ	b

Figure 9.1 Trojan design of size $(4 \times 4)/3$ for 12 treatments. The three alphabets correspond to three mutually orthogonal Latin squares

Strata		$\mathcal{W}_0$	$\mathcal{W}_1$	$\mathcal{W}_2$
	dim	1	2	9
mean	1	1	0	0
rows	3	0	0	0
columns	3	0	0	0
blocks	9	0	0	1/3
plots	32	0	1	2/3

Table 9.1 Efficiency factors for the design in Figure 9.1. Treatment subspace $\mathcal{W}_1$ corresponds to the contrasts between alphabets, and $\mathcal{W}_2$ corresponds to the contrasts within alphabets

Datasets were simulated as samples of independent normal variables with variance one, plus block effects with variance $\sigma_B^2 \in \{0, 0.5, 1\}$. A set of simulations was also performed with $\xi_{\text{plot}} = \sigma_E^2 = 1$ and $\xi_{\text{block}} = 0.4$, which corresponds to a "negative block variance" equal to -0.2 when the formula $\xi_{\text{block}} = k\sigma_B^2 + \sigma_E^2$ is considered.

Two REML analyses were performed: in the non-restricted analysis, the $\widehat{\xi_i}$ are only constrained to be positive; in the restricted analysis, they are constrained according to the model with random block, row and column effects (assuming positive or null variances for these factors).

For each value of σ_B^2, 500 datasets were simulated and analysed. For both REML analyses, the following quantities were calculated (see Azaïs, Monod and Bailey, 1998): the empirical average pairwise variance of treatment effects

(APV), the predicted average pairwise variance using the usual $\widehat{\Phi}$ estimator (PAPV), and the predicted average pairwise variance using the adjusted $\widehat{\Phi}_A$ estimator (PAPV-adj). The means, over the 500 simulations, of APV, PAPV and PAPV-adj were divided by the mean APV value obtained by an ANOVA analysis with treatment and row effects only. Thus the quantities reflect the increase or decrease in the variance of treatment effects, when a semi-Latin square is used instead of a complete block design.

σ_B^2	REML method	APV	PAPV	PAPV-adj
-0.2	non-restricted	1.03	0.72	1.17
-0.2	restricted	1.01	0.97	1.17
0	non-restricted	1.13	0.88	1.16
0	restricted	1.02	0.95	1.14
0.5	non-restricted	0.93	0.83	0.94
0.5	restricted	0.91	0.82	0.94
1	non-restricted	0.75	0.70	0.76
1	restricted	0.75	0.69	0.77

Table 9.2 Simulation results for the $(4 \times 4)/3$ semi-Latin square. APV: empirical average variance; PAPV: usual predicted average pairwise variance; PAPV-adj: adjusted predicted average pairwise variance

The results are presented in Table 9.2. For $\sigma_B^2 \in \{0.5, 1\}$, PAPV-adj is successful in reducing the variance under-estimation observed for PAPV. However, for $\sigma_B^2 = -0.2$ and for the restricted analysis with $\sigma_B^2 = 0$, the adjustment is too strong and PAPV-adj overestimates the variance. Thus the adjustment performs well, except when the "true" variance in the block stratum is close to the lower limit used for REML estimation. In that case, there is indeed a non-negligible probability that the variance estimates will fall on a boundary and the Taylor series expansions are no longer valid. This was already remarked on by Kackar and Harville (1984) in their simulations for balanced incomplete block designs and by Kenward and Roger (1997).

References

Azaïs, J.-M., Monod, H. and Bailey, R.A.B. (1998). The influence of design on validity and efficiency of neighbour methods. *Biometrics* **54**, 1374–1387.

Bailey, R.A.B. (1984). Discussion of the paper by T. Tjur "Analysis of variance models in orthogonal designs". *International Statistical Review* **52**, 65–76.

Bailey, R.A.B. (1991). Strata for randomized experiments. *Journal of the Royal Statistical Society* B **53**, 27–66.

Bailey, R.A.B. (1992). Efficient semi-Latin squares. *Statistica Sinica* **2**, 413–437.

Bailey, R.A.B. (1993). General balance: artificial theory or practical relevance ? In *Proceedings of the International Conference on Linear Statistical Inference LINSTAT'93* Eds T. Calinsky and R. Kala, pp.171–184. Dordrecht: Kluwer.

Bardin, A. and Azaïs, J.-M. (1991). Une hypothèse minimale pour une théorie des plans d'expérience randomisés. *Revue de Statistique Appliquée* **38**, 5–20.

Houtman, A.M. and Speed, T.P. (1985). Balance in designed experiments with orthogonal block structure. *Annals of Statistics* **11**, 1069–1085.

Kackar, A.N. and Harville, D.A. (1984). Approximations for standard errors of estimators of fixed and random effects in mixed linear models. *Journal of the American Statistical Association* **79**, 853–862.

Kenward, M. and Roger, J.H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics* **53**, 983–997.

Nelder, J.A. (1965). The analysis of randomized experiments with orthogonal block structure: block structure and the null analysis of variance. *Proceedings of the Royal Society* A **283**, 147–162.

Searle, S.R., Casella, G. and McCulloch, C.E. (1992). *Variance Components*. New York: Wiley.

Chapter 10

CONCENTRATION SETS, ELFVING SETS AND NORMS IN OPTIMUM DESIGN

Andrej Pázman
Faculty of Mathematics and Physics, Comenius University,
SK-842 48 Bratislava, Slovak Republic
pazman@fmph.uniba.sk

Abstract The aim of the paper is to compare different notions which have appeared in the development of optimum experimental design (concentration (or confidence) ellipsoid, Elving set, lower bounds for variances) and to show that they are connected via certain norms. This helps in using them for a graphical representation of the properties of experimental designs in linear models, or for comparing possibilities of optimum designs in different models.

Keywords: design of experiments, confidence ellipsoid, minimal variance, visualization of design

Introduction

A certain lack of robustness of optimum designs with respect to the choice of optimality criteria and/or the choice of the regression model are obstacles for a wider use of optimum design among practitioners. One can scarcely avoid this lack of robustness (which is in the substance of classical regression models), but one can try to give to the experimentalists simple tools to see quickly how a change of the criteria and/or of the model will change the results. In the present time of excellent graphical possibilities of computers, such tools should be graphical. From 2- or 3-dimensional pictures the experimenter should be able to grasp intuitively the influence of criteria or of model choice.

Here we want to contribute to this direction by reconsidering two well-known graphical presentations of designs, and extending their possibilities. One of them is the concentration ellipsoid (or the confidence ellipsoid), which has a very nice interpretation in terms of different optimality criteria. Its origin is in multivariate statistical analysis; in the context of experimental design it was

presented, probably for the first time, in Fedorov & Pázman (1968) (see also Fedorov, 1972). Recently, it was reconsidered in Fedorov & Hackl (1997). The concentration ellipsoid can be constructed straightforwardly, as it can for parameter subsets, which is its big advantage when the number of parameters is large. However, it always represents some fixed design, so it does not reflect the possibilities of optimizing designs.

On the other hand, as is well known, optimum designs are graphically well represented by the Elving set (Elfving, 1952). Recently, large extensions of the Elfving set have been presented in a series of papers by Dette (e.g. Dette, 1993).

One reason why we can easily read the variance of an estimator from the ellipsoid of concentration is that the square root of this variance is a norm. In Pázman (1974, 1978) further norms in experimental design are studied. More recent results of this kind have been presented in e.g. Pukelsheim (1981) or Dette (1993) as duality relations.

All this theoretical background may sound very abstract, but in fact we shall show that this is the way to relate the Elfving set to the concentration ellipsoid. We want to use these norms to extend the graphical possibilities of the ellipsoid of concentration (the "butterfly" defined below) and to consider projections and cuts of the Elfving set.

1. PROPERTIES OF THE CONCENTRATION ELLIPSOID AND OF THE ELFVING SET

1.1. THE LINEAR MODEL

Here we consider the usual linear regression model

$$y(x) = f^T(x)\,\theta + \varepsilon(x), \tag{10.1}$$

where $f(x) = (f_1(x), ..., f_m(x))^T$ and where x is the explanatory or control variable which takes values in a compact set $\mathcal{X} \subset \mathbb{R}$. The theory presented here holds for any compact $\mathcal{X}$. However, the suggested constructions are relatively easy to perform only when the dimension of $\mathcal{X}$ is small, which is often the case in applications. On the other hand, we make *a priori* no restrictions on the dimension of the unknown vector of parameters $\theta \in \mathbb{R}^m$. The functions $f_1(.), ..., f_m(.)$ are assumed to be linearly independent and continuous on $\mathcal{X}$. At each $x \in \mathcal{X}$ the random variable $y(x)$ can be observed, different or repeated observations being uncorrelated. It is assumed that $E[\varepsilon(x)] = 0, Var[\varepsilon(x)] = \sigma^2\lambda(x)$, with $\lambda(.)$ known. Without loss of generality one takes $\sigma^2\lambda(x) \equiv 1$. We consider here the approximate design theory (in the sense of Kiefer), i.e. a design ξ is a probability measure on $\mathcal{X}$. Without restrictions on generality one can suppose that the support of each ξ, which is the set $\{x \in \mathcal{X} : \xi(x) > 0\}$, is finite.

The matrix

$$M(\xi) = \sum_{x \in \mathcal{X}} f(x) f^T(x) \xi(x)$$

is the information matrix of the design ξ in the model (2). When the model is regular, $M^{-1}(\xi)$ is equal to the variance matrix of the least-squares estimator $\hat{\theta}$ of θ,

$$Var_\xi\left(\hat{\theta}\right) = M^{-1}(\xi).$$

In the general case one has for any $c \in \mathbb{R}^m$

$$Var_\xi\left(c^T\hat{\theta}\right) = \begin{cases} c^T M^-(\xi)\, c & \text{if c is in the range of } M(\xi) \\ \infty & \text{otherwise.} \end{cases}$$

Here $M^-(\xi)$ denotes an arbitrary g-inverse of $M(\xi)$.

1.2. THE CONCENTRATION ELLIPSOID

When ξ is a design, its concentration ellipsoid is the set

$$\mathcal{E}_\xi = \left\{z \in \mathbb{R}^m : z^T M(\xi) z \leq 1\right\}.$$

Note that it is an ellipsoidal cylinder when $M(\xi)$ is singular. The set $\mathcal{E}_\xi$ specifies how concentrated is the probability distribution of $\hat{\theta}$ around the true θ. A small $\mathcal{E}_\xi$ means a very concentrated distribution. Another interpretation of $\mathcal{E}_\xi$ is that a similar expression yields the confidence region for θ, when the number 1 in the inequality is changed to another number, depending on the required confidence level, and the vector $\theta - \hat{\theta}$ is substituted for the vector z.

Properties. One can compare very well any two designs, say ξ and η, according to the shape of their confidence ellipsoids. It is known that:

I) ξ is uniformly not worse than η, i.e. $M(\xi) \geq M(\eta)$ in the Loewner ordering, if and only if $\mathcal{E}_\xi \subseteq \mathcal{E}_\eta$.

II) ξ is better than η in the sense of the D-optimality criterion if and only if Volume$[\mathcal{E}_\xi] <$ Volume$[\mathcal{E}_\eta]$.

III) ξ is better than η in the sense of the E-optimality criterion if and only if $\rho(\xi) < \rho(\eta)$, where $\rho(\xi)$ is the radius of the smallest ball containing $\mathcal{E}_\xi$.

IV) ξ is better than η in the sense of the A-optimality criterion if and only if $d(\xi) < d(\eta)$, where $d(\xi)$ is the length of the main diagonal of the parallelepiped which surrounds $\mathcal{E}_\xi$.

V) ξ is better than η in the sense of the c-optimality criterion, i.e. $Var_\xi\left(c^T\hat{\theta}\right) < Var_\eta\left(c^T\hat{\theta}\right)$

$$\Leftrightarrow \quad \max_{z \in \mathcal{E}_\xi} \left[z^T c\right]^2 < \max_{z \in \mathcal{E}_\eta} \left[z^T c\right]^2$$

VI) The square root of the variance of the estimator $\hat{\theta}_i$ is equal to the length of the orthogonal projection of $\mathcal{E}_\xi$ onto the ith coordinate axis. More generally, the square root of the variance of any $c^T\hat{\theta}$ (with $\|c\| = 1$) is the length of the projection of $\mathcal{E}_\xi$ onto the straight line defined by the vector c.

A conclusion of all these nice properties is: Do we need any optimality criteria for comparing designs? Is it not better to compare visually the concentration ellipsoids and so to compare designs "globally" (due to the excellent properties of our eye-brain connections), instead of looking just for one property expressed by an optimality criterion?

The butterfly. The last statement is emphasized by considering two other sets obtained directly from $\mathcal{E}_\xi$. The set

$$\mathcal{W}_\xi = \left\{ \sqrt{Var_\xi\left(c^T\hat{\theta}\right)} \times c : c \in \mathbb{R}^m, \|c\| \leq 1 \right\},$$

which can be constructed graphically from $\mathcal{E}_\xi$ (see Fig. 10.1), since, according to the property VI), we can write

$$\mathcal{W}_\xi = \left\{ \left(\max_{z \in \mathcal{E}_\xi} \left| z^T c \right| \right) c : \|c\| \leq 1 \right\}$$

and the set

$$\mathcal{S}_\xi = \left\{ c \in \mathbb{R}^m : Var_\xi\left(c^T\hat{\theta}\right) \leq 1 \right\} = \left\{ c \in \mathbb{R}^m : \max_{z \in \mathcal{E}_\xi} \left(z^T c\right)^2 \leq 1 \right\}.$$

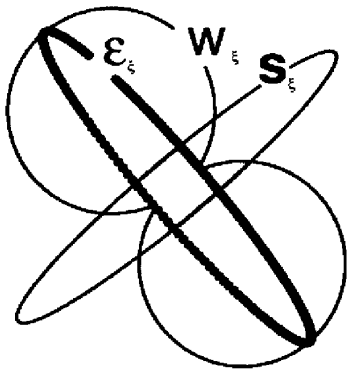


Figure 10.1 The butterfly when $M_{11}(\xi) = 2.1, M_{12}(\xi) = M_{21}(\xi) = 1.9, M_{22}(\xi) = 1.5$

The last expression suggests in addition a graphical construction from $\mathcal{E}_\xi$ (see Fig. 10.1). Both $\mathcal{W}_\xi$ and $\mathcal{S}_\xi$ characterize the quality of the design ξ. In

particular, $\mathcal{W}_\xi$ allows us to see how good the design is for different choices of $c^T\theta$. Graphically, $\mathcal{W}_\xi$ has the form of "wings" and $\mathcal{S}_\xi$ may have the form of the "body" in a picture, which, with poetical exaggeration, can be called "the butterfly", as in Fig. 10.1.

Nevertheless, we obtain such a nice "butterfly" only if the variance matrix of $\hat{\theta}$ is neither "too large" nor "too small", since, in the first case, $\mathcal{S}_\xi \in \mathcal{W}_\xi$ (so we have "an egg in a butterfly") and, in the second case $\mathcal{W}_\xi \in \mathcal{S}_\xi$ (and we have "a butterfly in an egg").

The "magic" norm. All the properties of $\mathcal{E}_\xi$ given above, apart from II), and consequently also the "butterfly", can be derived directly from the known equality

$$Var_\xi\left(c^T\hat{\theta}\right) = \max_{\alpha\in\mathbb{R}^m, M(\xi)\alpha\neq 0} \frac{\left(c^T\alpha\right)^2}{\alpha^T M\left(\xi\right)\alpha} = \max\left\{\left(c^T\alpha\right)^2 : \alpha\in\mathcal{E}_\xi\right\}, \tag{10.2}$$

(Rao (1973), eq. 1f.1.1). The equality also holds when $M\left(\xi\right)$ is singular, or when the variance is infinite. The expression in (10.2) is a squared norm of a functional defined on the set of all possible response functions in the model (2). To see this, denote by $\mathcal{H}$ the set of all possible response functions $\eta_\theta\left(x\right) = f^T\left(x\right)\theta$ in (2) and denote by $g_c\left(.\right)$ the functional defined on $\mathcal{H}$ by the formula

$$g_c\left(\eta_\theta\right) = c^T\theta; \qquad \text{when } \eta_\theta\left(x\right) = f^T\left(x\right)\theta.$$

Then, according to (10.2), the variance of the least squares estimate of the functional is

$$\begin{aligned} Var_\xi\left(\hat{g}_c\left(.\right)\right) &= Var_\xi\left(c^T\hat{\theta}\right) \\ &= \max_{\eta\in\mathcal{H}} \frac{g_c^2\left(\eta\right)}{\int_{\mathcal{X}}\eta^2\left(x\right)d\xi\left(x\right)}, \end{aligned} \tag{10.3}$$

which evidently is the $L_2\left(\xi\right)$ norm of the functional $g_c\left(.\right)$ (cf. Pázman (1974) for an extension to infinite dimensional models).

Graphs for parameter subsets. The basic difficulty is that we cannot draw pictures of the ellipsoid or of the butterfly for more than 3 parameters. But when the parameter dimension is large, one can construct them for parameter subsets.

Suppose that $\theta_1, ..., \theta_s$ is such a subset, with $s < m$. The concentration ellipsoid for these parameters is the set

$$\mathcal{E}_\xi^{(s)} = \left\{z\in\mathbb{R}^s : z^T M^{(s)}\left(\xi\right)z \leq 1\right\}.$$

Here $M^{(s)}(\xi)$ is the information matrix for these parameters,

$$M^{(s)}(\xi) = M_{11}(\xi) - M_{12}(\xi) M_{22}^{-}(\xi) M_{21}(\xi), \tag{10.4}$$

where

$$M(\xi) = \begin{pmatrix} M_{11}(\xi) & M_{12}(\xi) \\ M_{21}(\xi) & M_{22}(\xi) \end{pmatrix}$$

is the decomposition into blocks, and $M_{22}^{-}(\xi)$ is an arbitrary g-inverse of $M_{22}(\xi)$. As follows from property V) or VI), $\mathcal{E}_{\xi}^{(s)}$ is the orthogonal projection of the ellipsoid $\mathcal{E}_{\xi}$ onto the coordinate plane defined by the first s coordinates. The components of the "butterfly" are now the sets

$$\mathcal{W}_{\xi}^{(s)} = \left\{ \sqrt{Var_{\xi}\left(c^T \hat{\theta}^{(s)}\right)} \times c : c \in \mathbb{R}^s, \|c\| \leq 1 \right\}$$

and

$$\mathcal{S}_{\xi}^{(s)} = \left\{ c \in \mathbb{R}^s : Var_{\xi}\left(c^T \hat{\theta}^{(s)}\right) \leq 1 \right\} = \left\{ c \in \mathbb{R}^s : \max_{z \in \mathcal{E}_{\xi}^{(s)}} \left(z^T c\right)^2 \leq 1 \right\}.$$

The interpretation and the properties of $\mathcal{E}_{\xi}^{(s)}$ are exactly the same as the properties I) to VI) stated for $\mathcal{E}_{\xi}$, except that $M^{(s)}(\xi)$ is substituted for $M(\xi)$, and we consider only optimality criteria related to $\theta_1, ..., \theta_s$. So, we can have full visual information about the properties of the design ξ by taking simultaneous pictures for any couple (θ_i, θ_j), $i < j$. And we have much more information about the properties of the design than just from the value of an optimality criterion.

1.3. THE ELFVING SET AND THE ELFVING THEOREM

By definition, the Elfving set is equal to the set

$$S = conv\left(\{f(x) : x \in \mathcal{X}\} \cup \{-f(x) : x \in \mathcal{X}\}\right)$$

where $conv$ denotes the convex hull. It is related to the model (2) and not just to a particular design. So, in a certain sense, it compares models, not designs. The famous Elfving theorem (Elfving, 1952) is

Theorem 1. *We have*

$$\mu \in \arg\min_{\xi} Var_{\xi}\left(c^T \hat{\theta}\right)$$

if and only if there is a number $\gamma > 0$, such that γc is a boundary point of S, and that there is a function $\delta(x)$, with values equal either to 1 or to -1, so that $\gamma c = \sum_{x \in \mathcal{X}} \delta(x) f(x) \mu(x)$. Then the equality $\gamma^{-2} = Var_\mu\left(c^T \hat{\theta}\right)$ holds.

Properties of S. The properties of S can be expressed in a form which is analogous (or dual) to the properties of $\mathcal{E}_\xi$ in Section 1.2. For this purpose, let us consider a model alternative to model (2), having the same vector of parameters θ,

$$y^*(x) = f^{*T}(x)\theta + \varepsilon^*(x); \quad x \in \mathcal{X}^* \tag{10.5}$$

and let us denote by S^* the Elfving set corresponding to model (10.5). The following is known:

i) We have $S^* \subset S$ if and only if model (2) is uniformly not worse than model (10.5) in the sense that for every $c \in \mathbb{R}^m$ we have

$$\min_\xi Var_\xi\left(c^T \hat{\theta}\right) \leq \min_\mu Var^*_{\mu,}\left(c^T \hat{\theta}\right).$$

Here $Var^*_\mu\left(c^T \hat{\theta}\right)$ denotes the variance in (10.5) under the design μ on $\mathcal{X}^*$.

ii) Model (2) is better than model (10.5) in the sense of the achievable minimal value of the E-optimality criterion if and only if $\tau(S^*) < \tau(S)$ where $\tau(S)$ denotes the radius of the in-ball of S (Pukelsheim and Studden, 1993) when the minimal eigenvalue of the E-optimal design has multiplicity 1. See Dette and Studden (1993) for the general case, where however a generalized Elfving set must be used.

iii) Model (2) is better than model (10.5) in the sense of the c-optimality criterion if and only if $\{c\} \cap S^* \subset \{c\} \cap S$, where $\{c\}$ denotes the straight line defined by the vector c.

iv) Model (2) is uniformly better than model (10.5) with respect to the parameter subset $\theta_1, ..., \theta_s$ if and only if

$$\left\{e^{(1)}, ..., e^{(s)}\right\} \cap S^* \subset \left\{e^{(1)}, ..., e^{(s)}\right\} \cap S,$$

where $\left\{e^{(1)}, ..., e^{(s)}\right\}$ denotes the coordinate plane in $\mathbb{R}^m$ given by the canonical vectors $e^{(1)}, ..., e^{(s)}$ (i.e. vectors on coordinate axes). This means that, in contrast to the properties of the concentration ellipsoid, Elving sets for parameter subsets are given not by projections, but by cuts, of S.

2. MINIMAL VARIANCES AND NORMS

We present here an equality similar to (10.2), but for the minimal variance in model (2).

Theorem 2. *For any $c \in \mathbb{R}^m$ we can write*

$$\min_{\xi} Var_{\xi}\left(c^T\hat{\theta}\right) = \max_{\alpha\in\mathbb{R}^m\setminus\{0\}} \frac{\left(c^T\alpha\right)^2}{\max_{x\in\mathcal{X}}\left[\alpha^T f(x)\right]^2}. \tag{10.6}$$

Proof. The theorem appeared in Pázman (1978) in the form given below in the Remark following Theorem 3. It was shown there that the Elfving theorem is a direct consequence of (10.6). Here we present a short proof of (10.6), but based on the Elfving theorem.

Exchanging the order of max and min we obtain

$$\min_{\xi}\max_{\alpha} \frac{\left(c^T\alpha\right)^2}{\sum_x\left(\alpha^T f(x)\right)^2\xi(x)} \geq \max_{\alpha} \frac{\left(c^T\alpha\right)^2}{\max_x\left(\alpha^T f(x)\right)^2}.$$

Hence, according to (10.2) we obtain the inequality $\geq$ instead of the equality in (10.6). To obtain the converse inequality, define $\tilde{c} = c \Big/ \left[\min_{\xi} Var_{\xi}\left(c^T\hat{\theta}\right)\right]^{1/2}$. From the Elfving theorem it follows that $\tilde{c}$ is a boundary point of S. Hence, there is a tangent hyperplane to S at that point, i.e. there is an $\alpha_o \in \mathbb{R}^m$, such that

$$\max_x\left[\alpha_o^T f(x)\right]^2 = \left(\tilde{c}^T\alpha_o\right)^2.$$

Here we used the fact that $\{f(x) : x\in\mathcal{X}\}\cup\{-f(x) : x\in\mathcal{X}\}$ is the generating set of S. Hence

$$\begin{aligned}\max_{\alpha}\frac{\left(c^T\alpha\right)^2}{\max_x\left(\alpha^T f(x)\right)^2} &\geq \frac{\left(\tilde{c}^T\alpha_o\right)^2}{\max_x\left(\alpha_o^T f(x)\right)^2}\min_{\xi} Var_{\xi}\left(c^T\hat{\theta}\right)\\ &= \min_{\xi} Var_{\xi}\left(c^T\hat{\theta}\right)\end{aligned}$$

■

For the parameter subset $\theta_1, ..., \theta_s$ we have the following statement:

Theorem 3. *For any $c \in \mathbb{R}^s$ we can write*

$$\min_{\xi} Var_{\xi}\left(\sum_{i=1}^{s} c_i\hat{\theta}_i\right) = \max_{\alpha\in\mathbb{R}^s, M^{(s)}(\xi)\alpha\neq 0} \frac{\left(c^T\alpha\right)^2}{\max_{\xi}\alpha^T M^{(s)}(\xi)\alpha}. \tag{10.7}$$

Proof. We have

$$\begin{aligned}Var_{\xi}\left(\sum_{i=1}^{s} c_i\hat{\theta}_i\right) &= c^T\left[M^{(s)}(\xi)\right]^{-1}c\\ &= \max_{\alpha\in\mathbb{R}^s}\frac{\left(c^T\alpha\right)^2}{\alpha^T M^{(s)}(\xi)\alpha}.\end{aligned}$$

So, as in the first part of the proof of Theorem 2, we obtain

$$\min_{\xi} Var_{\xi}\left(\sum_{i=1}^{s} c_i\hat{\theta}_i\right) \geq \max_{\alpha\in\mathbb{R}^s} \frac{\left(c^T\alpha\right)^2}{\max_{\xi}\alpha^T M^{(s)}(\xi)\,\alpha}.$$

Write

$$Q(\xi) = \begin{pmatrix} M^{(s)}(\xi) & 0 \\ 0 & 0 \end{pmatrix}.$$

According to (10.4), the matrix $M(\xi) - Q(\xi)$ is positive semidefinite, i.e. $\alpha^T M(\xi)\,\alpha - \alpha^T Q(\xi)\,\alpha \geq 0$ for every $\alpha \in \mathbb{R}^m$. Now let $d^T = \left(c^T, 0\right)$. We have, according to Theorem 2,

$$\begin{aligned}\min_{\xi} Var_{\xi}\left(\sum_{i=1}^{s} c_i\hat{\theta}_i\right) &= \min_{\xi} Var_{\xi}\left(d^T\hat{\theta}\right) = \max_{\alpha\in\mathbb{R}^m} \frac{\left(d^T\alpha\right)^2}{\max_{\xi}\alpha^T M(\xi)\,\alpha} \\ &\leq \max_{\alpha\in\mathbb{R}^m} \frac{\left(d^T\alpha\right)^2}{\max_{\xi}\alpha^T Q(\xi)\,\alpha} = \max_{\alpha\in\mathbb{R}^s} \frac{\left(c^T\alpha\right)^2}{\max_{\xi}\alpha^T M^{(s)}(\xi)\,\alpha}.\end{aligned}$$

Remark. The right-hand side of (10.6) is again a squared norm of the functional $g_c(.)$ considered in Section 1.2, but a different norm, namely

$$\max_{\eta\in\mathcal{H}} \frac{|g_c(\eta)|}{\max_{x\in\mathcal{X}}|\eta(x)|}.$$

(See Pázman (1978) for the infinite-dimensional case). ■

3. THE MINIMAL CONCENTRATION SET - A WAY TO CONSTRUCT THE ELFVING SET

3.1. CONSTRUCTIONS FOR THE FULL SET OF PARAMETERS

Consider the set

$$\mathcal{E} = \left\{ z \in \mathbb{R}^m : \max_{x\in\mathcal{X}} \left[z^T f(x)\right]^2 \leq 1 \right\},$$

which can be called the **minimal concentration set**, since we have

$$\mathcal{E} = \left\{ z \in \mathbb{R}^m : \forall_{\xi} \sum_{x} \left[z^T f(x)\right]^2 \xi(x) \leq 1 \right\} = \cap_{\xi}\mathcal{E}_{\xi}.$$

On the other hand, from Theorem 2, it follows that the Elfving set is equal to

$$S = \left\{ c \in \mathbb{R}^m : \max_{z\in\mathcal{E}} \left[c^T z\right]^2 \leq 1 \right\}, \tag{10.8}$$

which relates the minimal concentration set to the Elfving set. This suggests a good graphical construction of the Elfving set (when $\dim(\theta)$ is small). Namely, we can relatively easy construct the set $\mathcal{E}$ "point by point", since the required maximization is over the set $\mathcal{X}$ which is low-dimensional. Then, having the picture of the boundary of $\mathcal{E}$, one can easily obtain the set S graphically from (10.8).

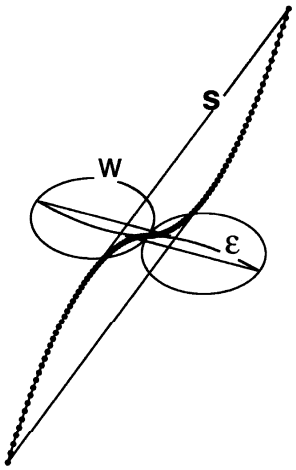


Figure 10.2 Minimal butterfly with the Elfving set in the model $f^T(x) = (x, x^2), 0 < x < 2.5$

We can also construct the "minimal butterfly" as in Section 1.2, starting from $\mathcal{E}$ instead of $\mathcal{E}_\xi$ (see Fig. 10.2). Then the Elfving set S plays the role of S_ξ (i.e. the "body" of the butterfly) and the "wings" are

$$\mathcal{W} = \left\{ \sqrt{\min_{\xi} Var_\xi \left(c^T \hat{\theta} \right)} \times c : \|c\| \leq 1 \right\}$$

$$= \left\{ \left[\max_{u \in \mathcal{E}} \left| c^T u \right| \right] c : c \in \mathbb{R}^m, \|c\| \leq 1 \right\}.$$

As written above, in contrast to Section 1.2, these pictures are not aimed at comparing designs, but may help to compare models.

3.2. THE PROJECTIONS OF THE ELFVING SET

When $\dim(\theta)$ is large, one cannot construct the Elfving set or the butterfly graphically. However, one can relatively easily construct the projections of these sets.

For each $x \in \mathcal{X}$ let

$$f^{(s)}(x) = (f_1(x), ..., f_s(x))^T.$$

Then the set

$$S_{(s)} = conv\left(\left\{f^{(s)}(x) : x \in \mathcal{X}\right\} \cup \left\{-f^{(s)}(x) : x \in \mathcal{X}\right\}\right)$$

is evidently the projection of the Elfving set onto the plane of the first s coordinates. In full mathematical analogy with the case $s = m$ (without any statistical interpretation) we obtain

$$S_{(s)} = \left\{ c \in \mathbb{R}^s : \max_{\alpha \in \mathbb{R}^s} \frac{\left(c^T \alpha\right)^2}{\max_{x \in \mathcal{X}} \left[\left(f^{(s)}(x)\right)^T \alpha\right]^2} \leq 1 \right\}.$$

However, notice that the expression

$$\max_{\alpha \in \mathbb{R}^s} \frac{\left(c^T \alpha\right)^2}{\max_{x \in \mathcal{X}} \left[\left(f^{(s)}(x)\right)^T \alpha\right]^2} \tag{10.9}$$

is not the minimal variance of $c^T \hat{\theta}^{(s)}$. So, one cannot obtain optimum variances for parameter subsets by considering projections of the Elfving set and $S_{(s)}$ is not the Elfving set for the parameter subset $\theta^{(s)}$. Still, these projections may be useful for giving an idea of the shape of the set S.

As stated in Section 1.3, Property iv), the Elfving set for a parameter subset is not the projection but the cut of S. But to construct such cuts (without having the set S beforehand) is much more difficult than to construct projections.

3.3. ELFVING SETS FOR PARAMETER SUBSETS

The easiest way is to start by constructing the minimal concentration set for the parameter subset $\theta_1, ..., \theta_s$,

$$\mathcal{E}^{(s)} = \cap_\xi \mathcal{E}_\xi^{(s)} = \left\{ z \in \mathbb{R}^s : \max_\xi z^T M^{(s)}(\xi)\, z \leq 1 \right\},$$

with $M^{(s)}(\xi)$ given by (10.4). Here the maximization cannot be reduced to a maximization over the set $\mathcal{X}$, but requires a program for maximization of the concave function $\xi \to z^T M^{(s)}(\xi)\, z$. The Elfving set for the parameter subset (hence a cut of S) is equal to

$$S^{(s)} = \left\{ c \in \mathbb{R}^s : \max_{z \in \mathcal{E}^{(s)}} \left(c^T z\right)^2 \leq 1 \right\},$$

which can be constructed graphically from the set $\mathcal{E}^{(s)}$. The wings of the "minimal butterfly" for parameter subsets are then

$$\mathcal{W}^{(s)} = \left\{ \max_{z \in \mathcal{E}^{(s)}} \left|c^T z\right| \times c : c \in \mathbb{R}^s, \ \ \|c\| \leq 1 \right\}.$$

4. A NOTE ON POSSIBLE EXTENSIONS

The ideas presented here can also be generalized to larger setups, by using extensions of the Elfving theorem. The idea is that, in principle, the D-optimality criterion, or another criterion, behaves similarly to the criterion of c-optimality, which, in the present paper, has been related to several norms. So, again, some norms can be defined. However, this leads to the use of very complex Elfving sets in rather structured spaces, as in Dette (1993), which cannot be presented graphically.

Acknowlegdement. Research supported by the Slovak VEGA grant agency, project No. 1/7295/20.

References

Dette, H. (1993). Elfving's theorem for D-optimality. *Ann. Statist.* **21**, 753–776.

Dette, H., Studden, W.J. (1993). Geometry of E-optimality. *Ann. Statist.* **21**, 416–433.

Elfving, G. (1952). Optimum allocation in linear regression theory. *Ann. Math. Statist.* **23**, 255–262.

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

Fedorov, V.V. and Hackl, P. (1997). *Model Oriented Design of Experiments.* Lecture Notes in Statistics No. 125. New York: Springer-Verlag.

Fedorov, V.V. and Pázman, A. (1968). Design of physical experiments. *Fortschritte der Physik* **16**, 325–355.

Pázman, A. (1974). The ordering of experimental designs (A Hilbert space approach). *Kybernetika* **10**, 373–388.

Pázman, A. (1978). Hilbert-space methods in experimental design. *Kybernetika* **14**, 73–84.

Pukelsheim, F. (1981). On c-optimal design measures. *Mathem. Operationsforschung und Statistik* ser. *Statistics* **12**, 13–20.

Pukelsheim, F. and Studden, W.J. (1993). E-optimal designs for polynomial regression. *Ann. Statist.* **21**, 402–415.

Rao, C.R. (1973). *Linear Statistical Inference and Its Applications (2nd edition)*. New York: Wiley.

Chapter 11

SEQUENTIAL CONSTRUCTION OF AN EXPERIMENTAL DESIGN FROM AN I.I.D. SEQUENCE OF EXPERIMENTS WITHOUT REPLACEMENT

Luc Pronzato
Laboratoire I3S, CNRS–UNSA, Sophia-Antipolis, France
pronzato@unice.fr

Abstract We consider a regression problem, with observations $y_k = \eta(\boldsymbol{\theta}, \xi_k) + \epsilon_k$, where $\{\epsilon_k\}$ is an i.i.d. sequence of measurement errors and where the experimental conditions ξ_k form an i.i.d. sequence of random variables, independent of $\{\epsilon_k\}$, which are observed sequentially. The length of the sequence $\{\xi_k\}$ is N but only $n < N$ experiments can be performed. As soon as a new experiment ξ_k is available, one must decide whether to perform it or not. The problem is to choose the n values $\xi_{k_1}, \ldots, \xi_{k_n}$ at which observations $y_{k_1}, \ldots, y_{k_n}$ will be made in order to estimate the parameters $\boldsymbol{\theta}$. An optimal rule for the on–line selection of $\xi_{k_1}, \ldots, \xi_{k_n}$ is easily determined when $p = \dim \boldsymbol{\theta} = 1$. A suboptimal open–loop feedback–optimal rule is suggested in Pronzato (1999b) for the case $p > 1$. We propose here a different suboptimal solution, based on a one–step–ahead optimal approach. A simple procedure, derived from an adaptive rule which is asymptotically optimal, Pronzato (1999a), when $p = 1$ ($N \to \infty$, n fixed), is presented. The performances of these different strategies are compared on a simple example.

Keywords: Sequential design, random experiments, expected determinant

Introduction

Consider a regression model, with observations

$$y_k = \eta(\bar{\boldsymbol{\theta}}, \xi_k) + \epsilon_k \,, \tag{11.1}$$

where $\{\epsilon_k\}$ is an i.i.d. sequence of measurement errors, assumed for simplicity to be distributed $\mathcal{N}(0, \sigma^2)$, and $\bar{\boldsymbol{\theta}} \in \Theta$ is the unknown true value of the model parameters to be estimated, with Θ an open subset of $\mathbb{R}^p$. The function $\eta(\boldsymbol{\theta}, \xi)$

is assumed continuously differentiable in $\boldsymbol{\theta}$, uniformly in ξ. The experimental conditions $\xi_k \in \mathcal{X}$ form an i.i.d. sequence, independent of $\{\epsilon_k\}$, of length N. Only $n < N$ experiments can be made. As soon as a new experiment ξ_k becomes available, one must decide whether to perform it or not, and one wishes to select n values $\xi_{k_1}, \ldots, \xi_{k_n}$ at best.

Let ν denote the probability measure of the ξ_k's, assumed to be known, and let $\{u_k\}$ denote the decision sequence: $u_k = 1$ if y_k is observed, with experimental conditions ξ_k, and $u_k = 0$ otherwise. We denote the random sequence $\xi_1, \ldots, \xi_N$ by $\boldsymbol{\xi}_1^N$ and the decision sequence $u_1, \ldots, u_N$ by $\boldsymbol{u}_1^N$, with, for any admissible policy,

$$u_j \in \mathcal{U}_j \subseteq \{0,1\}\,,\ j = 1, \ldots, N\,,\ \sum_{j=1}^{N} u_j = n\,. \tag{11.2}$$

We consider design criteria $\Phi(\cdot)$ for the estimation of $\boldsymbol{\theta}$ that are increasing functions of the Fisher information matrix, evaluated at a prior value $\hat{\boldsymbol{\theta}}^0$ for $\boldsymbol{\theta}$ if $\eta(\boldsymbol{\theta}, \xi)$ is nonlinear in $\boldsymbol{\theta}$ (local design). We define the rank–one matrix $\mathcal{M}(\xi)$ as

$$\mathcal{M}(\xi) = \frac{1}{\sigma^2} \frac{\partial \eta(\boldsymbol{\theta}, \xi)}{\partial \boldsymbol{\theta}} |_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}^0} \frac{\partial \eta(\boldsymbol{\theta}, \xi)}{\partial \boldsymbol{\theta}^\top} |_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}^0}$$

and we write $J(\boldsymbol{u}_1^N, \boldsymbol{\xi}_1^N) = \Phi[\sum_{k=1}^N u_k \mathcal{M}(\xi_k)]$. The measure ν for ξ induces a measure μ for $\mathcal{M}(\xi)$; we shall denote $\mathcal{M}_k = \mathcal{M}(\xi_k)$, $\mathbf{M}_j = \sum_{k=1}^j u_k \mathcal{M}_k$ and we assume that $\bar{\mathbf{M}} = E\{\mathcal{M}_k\}$ and $E\{\Phi(\mathbf{M}_N)\}$ exist. Note that ξ_k, and thus $\mathcal{M}_k$, is known when u_k is chosen. The problem is then to maximise

$$EJ(\boldsymbol{u}_1^N) = E\{\Phi(\mathbf{M}_N)\} = E\{\Phi(\sum_{k=1}^{N} u_k \mathcal{M}_k)\} \tag{11.3}$$

w.r.t. $\boldsymbol{u}_1^N$ satisfying (11.2), where the expectation $E\{\cdot\}$ is w.r.t. the product measure $\mu^{\otimes N}$ of $\mathcal{M}_1, \ldots, \mathcal{M}_N$. For any sequence $\boldsymbol{u}_1^N$ and any step j, $1 \le j \le N$, a_j will denote the number of observations already made; that is,

$$a_j = \sum_{k=1}^{j-1} u_k\,, \tag{11.4}$$

with $a_1 = 0$. The problem corresponds to a discrete–time stochastic control problem, where j represents time, $\mathcal{S}_j = (a_j, \mathbf{M}_{j-1}, \mathcal{M}_j)$ and $u_j \in \mathcal{U}_j \subseteq \{0,1\}$ respectively represent the state and control at time j. For each $j \in \{1, \ldots, N\}$ and each policy $\boldsymbol{u}_j^N$, the optimal decisions are obtained by solving the following

problem:

$$\max_{u_j \in \mathcal{U}_j} [E_{\mathcal{M}_{j+1}} \{ \max_{u_{j+1} \in \mathcal{U}_{j+1}} [E_{\mathcal{M}_{j+2}} \{ \max_{u_{j+2} \in \mathcal{U}_{j+1}} [\dots$$
$$E_{\mathcal{M}_{N-1}} \{ \max_{u_{N-1} \in \mathcal{U}_{N-1}} [E_{\mathcal{M}_N} \{ \max_{u_N \in \mathcal{U}_N} [\Phi(\sum_{k=1}^{N} u_k \mathcal{M}_k)]\}]\} \dots]\}]\}] \quad (11.5)$$

where $E_{\mathcal{M}_k}\{.\}$ denotes the expectation w.r.t. $\mathcal{M}_k$, distributed with the measure μ, and, for any $k = j, \dots, N$,

$$\mathcal{U}_k = \mathcal{U}_k(a_k) = \begin{cases} \{0\} & \text{if } a_k = n\,, \\ \{1\} & \text{if } a_k + N - k + 1 \leq n\,, \\ \{0,1\} & \text{otherwise.} \end{cases} \quad (11.6)$$

One can notice that the formulation of the problem would be much more complicated for a criterion function evaluated at the expected value of $\mathbf{M}_N$ or $\mathbf{M}_N^{-1}$, that is, when the problem is to maximize $\Phi(E\{\mathbf{M}_N\})$, or to minimize $\Phi(E\{\mathbf{M}_N^{-1}\})$, hence the choice (11.3).

The case $p = \dim \boldsymbol{\theta} = 1$ is considered in Section 1 The optimal (closed–loop) solution is given by a backward recurrence equation. A simple open–loop solution, asymptotically optimal for $N \to \infty$ with n fixed, is presented. Section 2 concerns the multidimensional case $p > 1$, for which the optimal solution cannot be obtained in close form. Several suboptimal solutions (open–loop feedback–optimal, one–step–ahead optimal) are proposed. Finally, the different strategies are compared on an illustrative example in Section 3

1. ESTIMATION OF A SCALAR PARAMETER

When $p = 1$, $\mathcal{M}$ is scalar, and we simply take $J(\boldsymbol{u}_1^N, \boldsymbol{\xi}_1^N) = \sum_{k=1}^N u_k \mathcal{M}_k$, that is, we maximize the expected information (note that the expected variance is not additive, which makes its minimisation more difficult). The problem to be solved at step j then becomes

$$\max_{u_j \in \mathcal{U}_j} [u_j \mathcal{M}_j + E_{\mathcal{M}_{j+1}} \{ \max_{u_{j+1} \in \mathcal{U}_{j+1}} [u_{j+1} \mathcal{M}_{j+1} + \cdots$$
$$E_{\mathcal{M}_{N-1}} \{ \max_{u_{N-1} \in \mathcal{U}_{N-1}} [u_{N-1} \mathcal{M}_{N-1} + E_{\mathcal{M}_N} \{ \max_{u_N \in \mathcal{U}_N} u_N \mathcal{M}_N \}] \} \cdots]\} \quad (11.7)$$

with the constraints (11.6) on the sets $\mathcal{U}_k$. It presents some similarities with the well-known secretary problem, see Pronzato (1999a) for a discussion. The state $\mathcal{S}_j$ at step j reduces to $\mathcal{S}_j = (a_j, \mathcal{M}_j)$, with a_j given by (11.4). Let $C(j, a_j, \mathcal{M}_j, \boldsymbol{u}_j^N)$ denote the optimal conditional expected gain–to–go at step j given $\mathcal{S}_j$, $C(j, a_j, \mathcal{M}_j, \boldsymbol{u}_j^N) = E\{\sum_{k=j}^N u_k \mathcal{M}_k | a_j, \mathcal{M}_j\}$, let $\tilde{C}(j, a_j, \mathcal{M}_j)$ denote its value when $\boldsymbol{u}_j^N$ is chosen optimally, and $c(j, a_j, \boldsymbol{u}_j^N)$, $\tilde{c}(j, a_j)$ respectively denote the expected values of $C(j, a_j, \mathcal{M}_j, \boldsymbol{u}_j^N)$ and $\tilde{C}(j, a_j, \mathcal{M}_j)$ w.r.t.

$\mathcal{M}_j$. One has, for $n-N+j-1 < a_j < n$, $\tilde{C}(j, a_j, \mathcal{M}_j) = \max_{u_j \in \{0,1\}}[u_j \mathcal{M}_j + \tilde{c}(j+1, a_j + u_j)]$. The optimal decision is thus

$$\tilde{u}_j(a_j, \mathcal{M}_j) = \begin{cases} 0 & \text{if } \mathcal{M}_j \leq \tilde{s}(j, a_j) = \tilde{c}(j+1, a_j) - \tilde{c}(j+1, a_j+1)\,, \\ 1 & \text{otherwise,} \end{cases} \tag{11.8}$$

which gives $\tilde{C}(j, a_j, \mathcal{M}_j) = \max[\mathcal{M}_j + \tilde{c}(j+1, a_j+1), \tilde{c}(j+1, a_j)]$ and the following backward recurrence equation for $\tilde{c}(j, a_j)$:

$$\tilde{c}(j, a_j) = E\{\max[\mathcal{M} + \tilde{c}(j+1, a_j+1), \tilde{c}(j+1, a_j)]\}\,. \tag{11.9}$$

The constraints (11.2) give $\tilde{c}(j, j+n-N-1) = (N-j+1)E\{\mathcal{M}\}$ and $\tilde{C}(j, n, \mathcal{M}_j) = 0$ for all $\mathcal{M}_j$, which initialises the recurrence (11.9) used to compute the optimal thresholds $\tilde{s}(j, a_j)$ in (11.8). The asymptotic properties ($N \to \infty$, n fixed) of this optimal solution are considered in Pronzato (1999a). In particular, it is shown that when the support of μ is unbounded, for any $a_j < n$, $\tilde{s}(j, a_j)$ and $\tilde{c}(j, a_j)$ tend to infinity as $N \to \infty$, and, when the tail of the distribution function of $\mathcal{M}$ decreases fast enough, a simple open–loop rule is asymptotically optimal.

Assume that μ is absolutely continuous w.r.t. the Lebesgue measure, with $\varphi(\cdot)$ its density, let $F(\cdot)$ denote the distribution function of $\mathcal{M}$, and consider the following assumption on the tail of $F(\cdot)$.

H1: $\bar{F}(\cdot) = 1 - F(\cdot)$ is twice differentiable, the density $\varphi(\cdot)$ is such that $\varphi(s) > 0$ and its derivative $\varphi'(\cdot)$ satisfies $\varphi'(s) < 0$ for s larger than some s_1. Moreover, $\bar{F}(\cdot)$ has the representation

$$\bar{F}(s) = \bar{F}(s_0) \exp\left[- \int_{s_0}^{s} \frac{1}{a(t)} dt \right]\,,\ s \geq s_0\,,$$

where the auxiliary function $a(t) > 0$ is absolutely continuous w.r.t. Lebesgue measure, with derivative $a'(t)$ having limit $\lim_{t\to\infty} a'(t) = \bar{a} \in [0, 1)$.

Note that $a(s) = \bar{F}(s)/\varphi(s)$ for $s > s_0$. When $\bar{a} = 0$, $F(\cdot)$ is a von Mises function, see Embrechts *et al.* (1997), p. 138, a class which contains for instance the exponential, normal, lognormal, Weibull and Gamma distributions, all with a tail decreasing faster than any power law $s^{-\alpha}$. In that case, $\lim_{t\to\infty} \bar{F}(t)\varphi'(t)/[\varphi(t)]^2 = -1$ and $\lim_{t\to\infty} a(t)/t = 0$, see Embrechts *et al.* (1997), p. 140. When $a(t) = t/b(t)$ with $b(t) \to \alpha \in (1, \infty)$ as $t \to \infty$, $\bar{a} = 1/\alpha$ and $\bar{F}(\cdot)$ is regularly varying with index $-\alpha$; that is (see Embrechts *et al.* (1997), p. 566),

$$\bar{F}(\cdot) \in \mathcal{R}_{-\alpha}\ :\ \lim_{s\to\infty} \frac{\bar{F}(ts)}{\bar{F}(s)} = t^{-\alpha}\,,\ t > 0\,.$$

We define

$$A(s) = \frac{h(s)\varphi(s)}{[\bar{F}(s)]^2}\,, \quad \bar{A} = \lim_{s\to\infty} A(s)\,,$$

where $h(s) = E\{\max(\mathcal{M} - s, 0)\}$. When $F(\cdot)$ is a von Mises function, direct application of L'Hôpital's rule shows that $\bar{A} = 1$. When $\bar{F}(\cdot) \in \mathcal{R}_{-\alpha}$, $\alpha \in (1, \infty)$, $\varphi(\cdot) \in \mathcal{R}_{-(\alpha+1)}$ and from Feller (1966), vol. 2, p. 281, $\bar{A} = \alpha/(\alpha-1)$.

The following property is proved in Pronzato (1999a).

Theorem 1. *Assume that $F(\cdot)$ satisfies* **H1**. *Any admissible open–loop decision rule $\boldsymbol{u}_1^N$, defined by $u_j(a_j, \mathcal{M}_j) = 0$ if $\mathcal{M}_j \le s(j, a_j)$, $u_j(a_j, \mathcal{M}_j) = 1$ otherwise, with thresholds $s(j, a_j)$ such that $(N-j)\bar{A}\bar{F}[s(j, a_j)] \to n - a_j$ when $N \to \infty$, is asymptotically optimal in the following sense:*
when $n - a_j = 1$,

(i) $c(j, n-1, \boldsymbol{u}_j^N)/\tilde{c}(j, n-1) \to 1$ as $N \to \infty$ if $\liminf_{s\to\infty} a(s) > c > 0$;

(ii) $\tilde{c}(j, n-1) - c(j, n-1, \boldsymbol{u}_j^N) \to 0$ as $N \to \infty$ if $\limsup_{s\to\infty} a(s) < C < \infty$;

when $n - a_j > 1$,

(iii) $c(j, a_j, \boldsymbol{u}_j^N)/\tilde{c}(j, a_j) \to 1$ as $N \to \infty$ if $\bar{a} = 0$ and $\liminf_{s\to\infty} a(s) > c > 0$;

(iv) $\tilde{c}(j, a_j) - c(j, a_j, \boldsymbol{u}_j^N) \to 0$ as $N \to \infty$ if $\bar{a} = 0$ and $a(s) \to 0$ as $s \to \infty$.

In particular, the open–loop rule defined by

$$u_j(a_j, \mathcal{M}_j) = \begin{cases} 0 & \text{if } \bar{F}(\mathcal{M}_j) \ge \frac{n-a_j}{\bar{A}(N-j+1)+(1-\bar{A})(n-a_j)-\epsilon}\,, \\ 1 & \text{otherwise,} \end{cases} \tag{11.10}$$

with $0 < \epsilon \ll 1$, satisfies (11.6) and the conditions in Theorem 1, and is thus asymptotically optimal.

When the distribution of the $\mathcal{M}_k$'s is unknown, their empirical distribution can be plugged in to the construction of the optimal rule (11.8) or the open-loop rule (11.10). This (suboptimal) approach corresponds to Certainty Equivalence (CE) control. First, we delay the decision for a few steps in order to initialise the construction of the empirical distribution of the $\mathcal{M}_k$'s, see Pronzato (1999a) for a discussion; then, at each step j, we substitute $\hat{F}_j(\cdot)$ for $F(\cdot)$, with $\hat{F}_j(\cdot)$ the empirical distribution function of $\mathcal{M}$ based on previous observed values, including $\mathcal{M}_j$. In the case of (11.10), this CE open–loop rule can be expressed directly in terms of the order statistics $\{\mathcal{M}^*_{i,j-1}\}$, ($\{\mathcal{M}^*_{i,j-1}\}$ denotes the sequence obtained by ordering the $\mathcal{M}_k$'s, $k \le j-1$, by decreasing values, with for any j, $\mathcal{M}^*_{i,j-1} = -\infty$ for $i > j-1$ and $\mathcal{M}^*_{0,j-1} = \infty$). The rule (11.10) then becomes

$$u_j^{CE}(a_j, \mathcal{M}_j) = \begin{cases} 0 & \text{if } \mathcal{M}_j < \mathcal{M}^*_{l_j,j-1}\,, \\ 1 & \text{otherwise,} \end{cases} \tag{11.11}$$

with

$$l_j = \left\lceil \frac{j(n-a_j)}{\bar{A}(N-j+1)+(1-\bar{A})(n-a_j)-\epsilon} \right\rceil , \qquad (11.12)$$

where $\lceil x \rceil$ rounds x to the nearest larger integer.

2. ESTIMATION OF A PARAMETER VECTOR

When $p = \dim(\boldsymbol{\theta}) > 1$, the problem (11.5) cannot be solved analytically, and we restrict our attention to suboptimal solutions.

2.1. OPEN–LOOP FEEDBACK–OPTIMAL DECISIONS

A suboptimal approach, called Open–Loop Feedback–Optimal (OLFO) in control theory, assumes that at each step j, all the decisions $\boldsymbol{u}_j^N$ may only depend on the current state $\mathcal{S}_j$. The OLFO solution is then obtained by solving the optimisation problem:

$$\max_{\boldsymbol{u}_j^N} [E_{\mathcal{M}_{j+1},\ldots,\mathcal{M}_N}\{\Phi(\sum_{k=j}^{N} u_k \mathcal{M}_k)|\mathcal{S}_j\}] ,$$

with $\boldsymbol{u}_j^N$ satisfying the constraints (11.2) and being a function of $\mathcal{S}_j = (a_j, \mathbf{M}_{j-1}, \mathcal{M}_j)$ only. It appears as a deterministic quantity in the conditional expectation above, with $\sum_{k=j}^{N} u_k = \max(0, n-a_j)$. For $n-N+j-1 < a_j < n$, the decision rule is thus:

$$\hat{u}_j(\mathcal{S}_j) = \begin{cases} 0 & \text{if } E_{\mathcal{M}_{j+1},\ldots,\mathcal{M}_{n+j-a_j}}\{\Phi(\mathbf{M}_{j-1}+\mathcal{M}_j+\sum_{k=j+1}^{n+j-a_j}\mathcal{M}_k)\} \leq \\ & E_{\mathcal{M}_{j+1},\ldots,\mathcal{M}_{n+j-a_j+1}}\{\Phi(\mathbf{M}_{j-1}+\sum_{k=j+1}^{n+j-a_j+1}\mathcal{M}_k)\} , \\ 1 & \text{otherwise} . \end{cases} \qquad (11.13)$$

Note that $\hat{u}_j$ is independent of N. It can be easily determined when the criterion takes the form $\Phi(\mathbf{M}) = \det(\boldsymbol{\Omega}^{-1}+\mathbf{M})$, with $\boldsymbol{\Omega}$ a positive–definite matrix. Indeed, expectations of determinants are obtained analytically as follows, see Pronzato (1999b).

Theorem 2. *Assume that the vectors $\mathbf{z}_i \in \mathbb{R}^p$ are i.i.d., with a probability measure μ_z such that $\bar{\mathbf{M}} = E\{\mathbf{z}\mathbf{z}^\top\}$ exists, then for any regular matrix $\mathbf{Q}$,*

$$E_{\mathbf{z}_1,\ldots,\mathbf{z}_k}\{\det[\mathbf{Q}+\sum_{l=1}^{k}\mathbf{z}_l\mathbf{z}_l^T]\} = \det[\mathbf{Q}] \times \left[1+\sum_{l=1}^{\min(k,p)} \mathtt{C}_k^l P_l(\mathbf{t}_1^l)\right] ,$$

where $\mathtt{C}_k^l = k!/[l!(k-l)!]$, $\mathbf{t}_1^l = (t_1,\ldots,t_l)$, with $t_l = tr[(\mathbf{Q}^{-1}\bar{\mathbf{M}})^l]$, and

$$P_l(\mathbf{t}_1^l) = l! \sum_{1^{n_1}\ldots l^{n_l}} \frac{(-1)^{l+\sum_{i=1}^{l} n_i}}{(1^{n_1}\ldots l^{n_l})(n_1!\ldots n_l!)} t_1^{n_1}\ldots t_l^{n_l} ,$$

where the summation $\sum_{1^{n_1}\ldots l^{n_l}}$ *is over all non-negative integers* $n_1, \ldots, n_l$ *such that* $\sum_{i=1}^{l} i n_i = l$*, that is, over all partitions of* l *which have* n_i *parts of size* i*.*

The proof is by induction on k. One has for $l = 1, \ldots, 4$:

$$P_1(t_1) = t_1\,, \qquad P_2(\mathbf{t}_1^2) = t_1^2 - t_2\,,$$
$$P_3(\mathbf{t}_1^3) = t_1^3 - 3t_1t_2 + 2t_3\,, \quad P_4(\mathbf{t}_1^4) = t_1^4 - 6t_1^2t_2 + 8t_1t_3 + 3t_2^2 - 6t_4\,.$$

The OLFO thresholds in (11.13) are obtained by applying Theorem 2 two times, with $\mathbf{Q} = \mathbf{\Omega}^{-1} + \mathbf{M}_{j-1} + \mathcal{M}_j$ and $\mathbf{Q} = \mathbf{\Omega}^{-1} + \mathbf{M}_{j-1}$ respectively. Note that this approach is not restricted to parameter estimation in regression models (11.1); the only requirement for Theorem 2 to apply is that the information matrix can be written as a sum of i.i.d. rank–one matrices.

2.2. ONE–STEP–AHEAD DECISIONS

If the experimental conditions could be chosen freely in a given set $\mathcal{X}$, at step j a steepest ascent algorithm, see Fedorov (1972); Wynn (1970), would select $\xi_j^* = \arg\max_{\xi_j \in \mathcal{X}} d(\mathbf{M}_{j-1}, \mathcal{M}_j)$, with

$$d(\mathbf{M}_{j-1}, \mathcal{M}_j) = \frac{\partial \Phi[(1-\alpha)\mathbf{M}_{j-1} + \alpha \mathcal{M}_j]}{\partial \alpha}_{|\alpha=0}\,.$$

When $\Phi(\mathbf{M}) = \det(\mathbf{\Omega}^{-1} + \mathbf{M})$, choosing ξ_j^* produces the maximal increase in Φ, since

$$\Phi(\mathbf{M}_{j-1} + \mathcal{M}_j) = \Phi(\mathbf{M}_{j-1})\{1 + \text{trace}[\mathcal{M}_j(\mathbf{\Omega}^{-1} + \mathbf{M}_{j-1})^{-1}]\}$$

and

$$\begin{aligned} d(\mathbf{M}_{j-1}, \mathcal{M}_j) \;\; = \;\; & \Phi(\mathbf{M}_{j-1})\{\text{trace}[\mathcal{M}_j(\mathbf{\Omega}^{-1} + \mathbf{M}_{j-1})^{-1}] \\ & - \text{trace}[\mathbf{M}_{j-1}(\mathbf{\Omega}^{-1} + \mathbf{M}_{j-1})^{-1}]\}\,. \end{aligned}$$

Assume that at step j there is only one additional observation to make, that is, $a_j = n - 1$. The state $\mathcal{S}_j$ is known, and the optimal solution of (11.5) is obtained as in Section 1, by solving a problem similar to (11.7) with the i.i.d. scalar variables $\mathcal{M}_k$, $k \geq j$, replaced by $\mathcal{M}'_k = d(\mathbf{M}_{j-1}, \mathcal{M}_k)$. In particular, Theorem 1 applies, and, under the conditions of the theorem, the open–loop rule (11.10) is asymptotically optimal.

When more observations are to be made, we still define $\mathcal{M}'_k = d(\mathbf{M}_{j-1}, \mathcal{M}_k)$ and maximise at step j:

$$EJ'(\boldsymbol{u}_j^N) = E_{\mathcal{M}'_{j+1}, \ldots, \mathcal{M}'_N}\{\sum_{k=j}^{N} u_k \mathcal{M}'_k | \mathcal{S}_j\}\,. \tag{11.14}$$

The optimal solution is given by (11.8), which provides a suboptimal solution for (11.5). The strategy can be further simplified by using (11.10). The OLFO rule can also be used to get a suboptimal solution for the maximisation of (11.14): ξ_j is accepted if $\mathcal{M}'_j > E\{\mathcal{M}'_k|\mathcal{S}_j\}$ and rejected otherwise.

Note that the distribution of $\mathcal{M}'_k$ may be difficult to derive from that of ξ_k, whereas using CE and empirical distributions is straightforward, see in particular (11.11,11.12).

3. EXAMPLE

Consider the following linear regression model, quadratic in the design variable ξ_k : $\eta(\boldsymbol{\theta}, \xi_k) = \theta_0 + \theta_1 \xi_k + \theta_2 \xi_k^2$, with $\Theta = I\!R^3$, $\mathcal{X} = I\!R$, $\bar{\boldsymbol{\theta}} = (-4, 4, -1)^\top$, so that the response $\eta(\bar{\boldsymbol{\theta}}, \xi)$ reaches its maximum at $\xi^* = 2$. The observation errors ϵ_k are i.i.d. $\mathcal{N}(0, \sigma^2)$, with $\sigma = 0.1$, and the ξ_k's are i.i.d. $\mathcal{N}(1, 1)$. $N = 100$ values of ξ_k are proposed, and $n = 10$ observations can be made.

Four strategies are compared, all based on $\Phi(\mathbf{M}) = \det(\boldsymbol{\Omega}^{-1} + \mathbf{M})$, with $\boldsymbol{\Omega} = 10^6 \mathbf{I}$:

- S_1 corresponds to the OLFO rule (11.13);
- S_2 corresponds to the OLFO rule for the solution of the one–step–ahead problem (11.14);
- S_3 corresponds to the optimal rule (11.8) for the maximisation of (11.14);
- S_4 corresponds to the open–loop rule (11.11,11.12) for the maximisation of (11.14), with $\epsilon = 0.01$ and $\bar{A} = 1$.

All the strategies use CE: in S_1, $\bar{\mathbf{M}}$ is replaced by the empirical mean of $\mathcal{M}_k$, see Theorem 2; S_2 (resp. S_3) uses the empirical mean (resp. empirical distribution) of $\mathcal{M}'_k$. Ten random samples $\xi_{-9}, \ldots, \xi_0$ are used to initialise the construction of the empirical quantities. Two criteria are used for the comparison: the determinant $\mathcal{D}$ of the information matrix per sample: $\mathcal{D}(\boldsymbol{u}_1^N) = \det[\mathbf{M}_N/n]$, and the squared error $\mathcal{E}$ for the estimation of ξ^*,

$$\mathcal{E}(\boldsymbol{u}_1^N) = [\bar{\theta}_1/(2\bar{\theta}_2) - \hat{\theta}_1^N/(2\hat{\theta}_2^N)]^2 \,,$$

where $\hat{\boldsymbol{\theta}}^N$ is the maximum *a posteriori* estimator, with a normal prior $\mathcal{N}(\hat{\boldsymbol{\theta}}^0, \boldsymbol{\Omega})$, $\hat{\boldsymbol{\theta}}^0 = (0, 3, -0.5)^\top$.

The results obtained for $q = 1000$ independent repetitions of the experiment are given in Table 11.1. From a result in Pronzato (1998a), taking n experiments at random gives on the average

$$E\{\det[\sum_{k=1}^{N} \mathcal{M}_k/n]\} = \frac{n!}{n^p(n-p)!} \det[\bar{\mathbf{M}}] \simeq 1.44\ 10^6 \,,$$

and the four strategies above thus yield an important improvement in terms of $\mathcal{D}$. Note that the formula above shows that the approximate design problem defined by the maximisation of $E\{\det[\mathbf{M}_N/n]\}$ with $n \to \infty$ (for instance, $n = \gamma N$, $0 < \gamma < 1$, $N \to \infty$) corresponds to the determination of an optimal constrained measure, a problem for which one can refer to Wynn (1982), Fedorov(1989) and in this volume, to Sahm and Schwabe (2000).

	$\mathcal{D}$	$\mathcal{E}$
S_1	8.92 10^6 (9.5 10^6)	5.09 10^{-5} (1.0 10^{-4})
S_2	8.71 10^6 (9.3 10^6)	5.03 10^{-5} (1.3 10^{-4})
S_3	1.82 10^7 (1.3 10^7)	2.84 10^{-5} (4.3 10^{-5})
S_4	1.73 10^7 (1.3 10^7)	2.87 10^{-5} (4.7 10^{-5})

Table 11.1 Average performances of strategies S_1 to S_4 ($N = 100$, $n = 10$, 1000 repetitions, standard deviations are in brackets)

Since the standard deviations in Table 1 are rather large, we perform a statistical analysis of the results, based on the method of paired comparisons. First, we compute the differences between the performances of each pair of strategies, the same ξ_k's and ϵ_k's being used for all strategies in each experiment. For comparing S_j with S_k, we compute $\Delta^{j,k}_{\mathcal{D},i} = [\mathcal{D}(S_j)]_i - [\mathcal{D}(S_k)]_i$, with $[\mathcal{D}(S)]_i$ the value of $\mathcal{D}$ obtained for strategy S in the i-th experiment. This gives $q = 1000$ independent realisations of $\Delta^{j,k}_{\mathcal{D},i}$, with empirical mean $E(\Delta^{j,k}_{\mathcal{D}})$ and standard deviation $\sigma(\Delta^{j,k}_{\mathcal{D}})$. The same is done with the criterion $\mathcal{E}$, which yields $E(\Delta^{j,k}_{\mathcal{E}})$ and $\sigma(\Delta^{j,k}_{\mathcal{E}})$. Then, we test if S_j performs significantly better or worse than S_k in terms of $\mathcal{D}$ and $\mathcal{E}$ by computing the ratios $\rho^{j,k}_{\mathcal{D}} = \sqrt{q}E(\Delta^{j,k}_{\mathcal{D}})/\sigma(\Delta^{j,k}_{\mathcal{D}})$ and $\rho^{j,k}_{\mathcal{E}} = \sqrt{q}E(\Delta^{j,k}_{\mathcal{E}})/\sigma(\Delta^{j,k}_{\mathcal{E}})$. If the two decision rules have similar average performances, which corresponds to the null hypothesis, then $\rho^{j,k}_{\mathcal{D}}$ and $\rho^{j,k}_{\mathcal{E}}$ approximately follow Student's t-distribution with $q-1$ degrees of freedom. For large values of q, which is the case here, the distribution is approximately $\mathcal{N}(0,1)$, and the critical value, for a level of significance 0.5% (one sided–test), is 2.576. Values $\rho^{j,k}_{\mathcal{D}}$ larger than 2.576 (resp. $\rho^{j,k}_{\mathcal{E}}$ smaller than -2.576) thus indicate that S_j performs significantly better than S_k in terms of $\mathcal{D}$ (resp. $\mathcal{E}$). The values of $\rho^{j,k}_{\mathcal{D}}$ and $\rho^{j,k}_{\mathcal{E}}$ are given in Table 11.2, with indices j, k corresponding respectively to lines and columns.

Table 11.2 indicates that the performances of S_1 and S_2 are not significantly different, and that S_3 and S_4 perform significantly better than S_1 and S_2, both in terms of $\mathcal{D}$ and $\mathcal{E}$. We conclude this example by indicating the average computing time for the four strategies. They are respectively 2.3, 0.14, 27.9

$\rho_{\mathcal{D}}^{j,k}$	S_1	S_2	S_3	S_4	$\rho_{\mathcal{E}}^{j,k}$	S_1	S_2	S_3	S_4
S_1	·	1.45	-23.12	-23.70	S_1	·	0.12	6.89	6.53
S_2	-1.45	·	-24.14	-24.64	S_2	-0.12	·	5.32	5.15
S_3	23.12	24.14	·	4.05	S_3	-6.89	-5.32	·	-0.16
S_4	23.70	24.64	-4.05	·	S_4	-6.53	-5.15	0.16	·

Table 11.2 Values of $\rho_{\mathcal{D}}^{j,k}$ and $\rho_{\mathcal{E}}^{j,k}$ (method of paired comparisons) for strategies S_1 to S_4 ($N = 100$, $n = 10$, 1000 repetitions)

and 0.35 seconds. The evaluation of expected values of determinants in S_1, see Theorem 2, explains the larger computing time than for S_2. Taking the small computing time of S_4 into account, together with the fact that the performances of S_3 and S_4 are very close, we conclude that S_4 is the most attractive strategy.

References

Embrechts, P., Klüppelberg, C. and Mikosch T. (1997). *Modelling Extremal Events*. Berlin: Springer-Verlag.

Fedorov, V.V. (1972). *Theory of Optimal Experiments*. N.Y.: Academic Press.

Fedorov, V.V. (1989). Optimal design with bounded density: optimization algorithms of the exchange type. *JSPI* **22**, 1–13.

Feller, W. (1966). *An Introduction to Probability Theory and Its Applications*. N.Y.: Wiley.

Pronzato, L. (1998a). On a property of the expected value of a determinant. *Statistics & Probability Letters* **39**, 161–165.

Pronzato, L. (1998b). Optimal selection of information with restricted storage capacity. In *Proc. ICASSP'98* **4**, pp. 2285–2288. Seattle.

Pronzato, L. (1999a). Optimal and asymptotically optimal decision rules for sequential screening and resource allocation. Technical Report 99–19, Laboratoire I3S, CNRS–UPRES-A 6070, Sophia Antipolis, France, `http://www.i3s.unice.fr/~pronzato/` .

Pronzato, L. (1999b). Sequential selection of observations in randomly generated experiments. *Tatra Mountains Math. Pub.* **17**, 167–175.

Wynn, H.P. (1970). The sequential generation of D-optimum experimental designs. *Ann. Math. Stat.* **41**, 1655–1664.

Wynn, H.P. (1982). Optimum submeasures with applications to finite population sampling. In *Statistical Decision Theory and Related Topics III. Proc. 3rd Purdue Symp* **2** (Eds S.S. Gupta and J.O. Berger), 485-495. N.Y.: Academic Press.

Chapter 12

OPTIMAL CHARACTERISTIC DESIGNS FOR POLYNOMIAL MODELS

J.M. Rodríguez-Díaz and J. López-Fidalgo
Department of Statistics of the University of Salamanca
Plaza de los Caídos, s/n, 37008-Salamanca, Spain
juanmrod@gugu.usal.es and fidalgo@gugu.usal.es

Abstract Using the characteristic polynomial coefficients of the inverse of the information matrix, design criteria can be defined between A- and D-optimality (López-Fidalgo and Rodríguez-Díaz, 1998). With a slight modification of the classical algorithms, the gradient expression allows us to find some optimal characteristic designs for polynomial regression. We observe that these designs are a smooth transition from A- to D-optimal designs. Moreover, for some of these optimal designs, the efficiencies for both criteria, A- and D-optimality, are quite good.

Nice relationships emerge when plotting the support points of these optimal designs against the number of parameters of the model. In particular, following the ideas developed by Pukelsheim and Torsney (1991), we have considered A-optimality. Another mathematical expression can be given for finding A-optimal support points using nonlinear regression. This could be very useful for obtaining optimal designs for the other characteristic criteria.

Keywords: A-optimality, characteristic criteria, D-optimality, polynomial regression

Introduction

Two of the best known and most used criteria in optimal experimental design are A- and D-optimality. Both are based on the eigenvalues of the inverse of the information matrix, or more precisely on the first and the last coefficients of the characteristic polynomial of this matrix. Paying attention to the rest of the coefficients it is possible to develop new criteria, called the *characteristic criteria* (López-Fidalgo and Rodríguez-Díaz, 1998). They can be defined as

$$\Phi_{Ch_k}(\xi) = \varphi_k \left[M(\xi)^{-1}\right], \quad k = 1, \ldots, m,$$

where m is the number of parameters of the model, $M(\xi)$ is the information matrix of the design ξ, defined as usual, and φ_k is the k-th coefficient of the

characteristic polynomial of the matrix, that is

$$\varphi_k(A) = \sum_{i_1 < \cdots < i_k} \lambda_{i_1} \cdots \lambda_{i_k},$$

with $\lambda_1, \ldots, \lambda_m$ the eigenvalues of A.

López-Fidalgo and Rodríguez-Díaz (1998) give some interesting properties for these new criteria. There are strong relationships between them. They are also differentiable and their gradients are

$$\bigtriangledown \Phi_k(M) = \sum_{i=k}^{m} \varphi_i(M^{-1})\ (-M)^{i-k-1}.$$

Now, in order to obtain optimal designs for the Ch_k-criteria, we will use the following gradient algorithm described in Pazman (1986), p. 157. Once the model and the design interval are chosen the method takes the following steps:

- The initial design is any ξ_0 such that $\det M(\xi_0) \neq 0$.
- At each stage, the previous design is modified as follows:

  $$\xi_{n+1} = (1 - \beta_n)\xi_n + \beta_n \xi_{x_n},$$

  where the β_n have to satisfy the conditions

  $$\beta_n \in (0, 1), \quad \sum_{n=0}^{\infty} \beta_n = \infty, \quad \lim_{n \longrightarrow \infty} \beta_n = 0$$

  and x_n is the point where the minimum of the following function

  $$x \in X \longrightarrow f^t(x) \bigtriangledown \Phi[M(\xi_n)] f(x)$$

  is achieved. Here we have called X the design space, $f(x)$ the model and $\bigtriangledown \Phi[M(\xi)]$ the gradient of the criterion function Φ applied to the information matrix of the design ξ.

In this paper we will focus on the polynomial model

$$E[y(x)] = a_0 + a_1 x + a_2 x^2 + \cdots + a_d x^d, \quad Var[y(x)] = \sigma^2,$$

which is very much used due to its simplicity and the fact that it can be seen as an approximation of more complicated functions. In the same sense the design interval will be $[-1, 1]$. Both the model and the design interval have been widely used, and many times together, for different criteria. Pukelsheim and Studden (1993) and Dette and Studden (1994) concentrate on E-optimality. Dette and Wong (1995) pay attention to the G-criterion while Studden (1980)

focuses on D_s-optimality, which aims to optimize the estimation of the s highest coefficients of the model. This was a generalization of the work of Kiefer and Wolfowitz (1959) for best estimation of a_d. Dette (1994) focuses on the problem of discrimination between different polynomial models for a general interval. New criteria are defined for that. Kiefer and Studden (1976) study large degree models for different criteria. Hoel and Levine (1964) and Dette and Wong (1996) try to get the best designs for extrapolation. Bellhouse and Herzberg (1984) consider uniform designs, with applications in cartography.

Eventually we will introduce a procedure for determining approximate support for Φ_{Ch_k}-optimal designs. We will use the method proposed by Pukelsheim and Torsney (1991) in order to get the optimal weights for this support.

1. DESIGNS AND EFFICIENCIES

The algorithm described previously has been implemented on the computer using the mathematical calculus program Mathematica. The results shown in Table 12.1 were achieved after 10.000 iterations on a Pentium II 166 computer. For this procedure we set $\beta_n = 1/(n+2)$ and obtained Φ_{Ch_k}-optimal designs for polynomial models of degree two to six. Note that the cases $k = 1$ and $k = d+1$ correspond respectively to A- and D-optimality, which can be found in profusion in the literature. We have included them in order to show the behaviour of the general algorithm for these particular cases and to compare the results with the rest of the optimum designs. In particular, the Φ_{Ch_m}- or D-optimum design is easy to obtain exactly for the polynomial model, for it is supported at the roots of the Legendre polynomials and all the support points have the same weight (see e.g. Pazman (1986), pp 177, 179).

The first thing one can observe is that we have found Φ_{Ch_k}-optimal designs supported on m points, the number of parameters of the model. This is well known for $k = 1$ and $k = m$ (A- and D-optimality). Furthermore, it is easy to see that the first optimal designs (low values of k for each model) put less weight at the extreme points of the interval, but their importance is increasing with k to the Φ_{Ch_m}-optimum, for which the m chosen points are equally important. Moreover, it seems there is a smooth transition between the A- and the D-optimum designs as we pass through the intermediate criteria.

The efficiencies of these designs for the characteristic functions are shown in Table 12.2. Cell (i, j) contains the efficiency of the Φ_{Ch_i}-optimal design for the criterion function Φ_{Ch_j} for different degrees.

As might be expected, for each model the efficiency of the Φ_{Ch_k}-optimum design decreases when we move away from k, but for the optimum designs corresponding to low values of k this reduction is more noticeable than for values near m. When we compute the average efficiency, or the quadratic mean of the efficiencies, of each optimal design, either for all the criteria or even only

d	k	Φ_{Ch_k}-optima for the polynomial model of degree d			
	1	±1. (.250)	0 (.500)		
2	2	±1. (.297)	0 (.407)		
	3	±1. (.333)	0 (.333)		
	1	±1. (.150)	±0.464 (.350)		
3	2	±1. (.173)	±0.424 (.327)		
	3	±1. (.215)	±0.435 (.285)		
	4	±1. (.250)	±0.447 (.250)		
	1	±1. (.105)	±0.677 (.250)	0 (.290)	
	2	±1. (.116)	±0.643 (.256)	0 (.256)	
4	3	±1. (.139)	±0.633 (.232)	0 (.257)	
	4	±1. (.170)	±0.643 (.216)	0 (.228)	
	5	±1. (.200)	±0.655 (.200)	0 (.200)	
	1	±1. (.080)	±0.789 (.187)	±0.291 (.232)	
	2	±1. (.086)	±0.767 (.198)	±0.267 (.216)	
5	3	±1. (.099)	±0.749 (.195)	±0.28 (.206)	
	4	±1. (.116)	±0.747 (.179)	±0.278 (.205)	
	5	±1. (.141)	±0.756 (.173)	±0.281 (.186)	
	6	±1. (.167)	±0.765 (.167)	±0.285 (.167)	
	1	±1. (.065)	±0.853 (.148)	±0.479 (.185)	0 (.205)
	2	±1. (.069)	±0.839 (.156)	±0.451 (.182)	0 (.186)
	3	±1. (.076)	±0.824 (.162)	±0.455 (.166)	0 (.191)
6	4	±1. (.086)	±0.815 (.155)	±0.461 (.171)	0 (.175)
	5	±1. (.100)	±0.817 (.145)	±0.46 (.168)	0 (.173)
	6	±1. (.121)	±0.823 (.144)	±0.464 (.156)	0 (.158)
	7	±1. (.143)	±0.83 (.143)	±0.469 (.143)	0 (.143)

Table 12.1 Φ_{Ch_k}-optimum designs for the first six polynomial models

d	$k\backslash h$	1	2	3	4	5	6	7
	1		98.5	94.5				
2	2	98.8		98.9				
	3	96.2	98.9					
	1		98.9	96.1	91.7			
3	2	99.2		98.5	95.0			
	3	97.9	98.8		99.0			
	4	96.1	96.0	99.0				
	1		99.0	97.4	94.6	90.7		
	2	99.3		99.2	96.9	93.2		
4	3	98.5	99.4		99.0	96.3		
	4	97.7	98.0	99.1		99.1		
	5	96.6	96.0	96.8	99.2			
	1		99.1	97.7	96.4	94.0	90.6	
	2	99.4		99.3	98.2	95.9	92.5	
5	3	98.7	99.4		99.6	97.9	94.9	
	4	98.2	98.8	99.6		99.2	97.0	
	5	97.7	97.5	98.5	99.3		99.3	
	6	97.0	96.3	96.6	97.3	99.3		
	1		99.2	98.0	96.8	95.8	93.8	90.7
	2	99.4		99.4	98.4	97.4	95.4	92.3
	3	98.8	99.5		99.6	98.8	97.1	94.2
6	4	98.5	98.9	99.6		99.7	98.3	95.8
	5	98.2	98.4	99.1	99.7		99.4	97.6
	6	97.8	97.7	98.2	98.7	99.4		99.4
	7	97.3	96.7	96.7	97.0	97.8	99.4	

Table 12.2 Entry (k, h) is the efficiency of the optimum design for Φ_{Ch_k} relative to the optimum design for Φ_{Ch_h}, $k, h = 1, 2, \ldots, d+1$.

for A- and D-optimality, we find that the highest efficiencies are reached for the highest values of k. But, surprisingly, the best value is not reached for $k = m$, but for $k = m - 1$ or $k = m - 2$. This means that if we consider, for instance, the quadratic model and we are interested in minimizing the average of the variances of the estimates of the parameters, we should choose the A-optimum design ($k = 1$). If we want to minimize the volume of the confidence ellipsoid of the estimates of the parameters, we have to choose the D-optimum design ($k = 3$). But if our intentions are, in fact, not well-defined, the best option would be to choose the Φ_{Ch_2}-optimal design, which is quite efficient for both interests. Similar comments can be made for other degrees.

2. APPROXIMATE SUPPORT POINTS

The Φ_{Ch_k}-optimum designs are strongly inter-related, either when we pay attention to the support or to the weights; also either when we focus on a fixed degree and let k vary or when we fix k and deal with different values of d. This

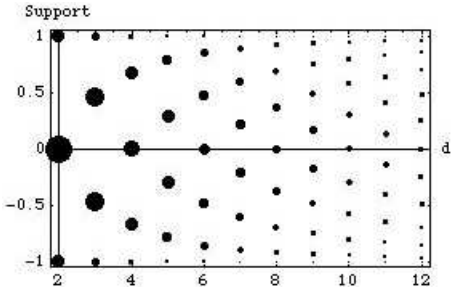


Figure 12.1 A-optimal designs for the polynomial models of degree 2 to 12

last representation has more keenly attracted our attention. Figure 12.1 shows a plot of the Φ_{Ch_k}-optimal designs (with k fixed equal to 1 in this case, that is A-optimality) against the degree of the model. Thus the support of the Φ_{Ch_1}-optimum for the model of degree d is placed in the vertical line at d, the diameter of each point being proportional to the weight it has in the optimum design. As can be seen, the supports form well-defined curves, while the weights decrease as d increases. We have shown the plot for A-optimality, since it is the first of the characteristic criteria and is very popular, but the rest of the criteria yield similar plots.

It is a natural thought to try to compute these curves in order to obtain the optimum design in a simple and easy way. From now on we will define Φ-weights for a given support to be the best weights for those points from the point of view of Φ-optimality. Pukelsheim and Torsney (1991) compute the best A-weights for the quantiles of the arcsin distribution first considered by Fedorov (1972) for the D-criterion. The A-efficiencies are quite good, even better than the D-efficiencies for the D-weights for the same support. In this first approach we consider two possible supports $\{\pm x_{i\,1}, i = 1, 2, \ldots, \left[\frac{m}{2}\right]\}$ and $\{\pm x_{i\,2}, i = 1, 2, \ldots, \left[\frac{m}{2}\right]\}$, where

$$x_{i\,1} = \left[1 - \left(\frac{2\ i}{m}\right)^{c_i}\right],\ x_{i\,2} = \left[1 - \left(\frac{2\ i}{m}\right)^{m^{d_i}}\right], \quad i = 0, 1, 2, \ldots, \left[\frac{m}{2}\right]$$

and c_i and d_i are to be determined using nonlinear regression techniques.

Each support depends only on one parameter, which can be estimated by a function depending on i. The sequences $\{c_i\}$ and $\{d_i\}$ are decreasing in i. We first use nonlinear regression to estimate c_i and d_i for support points up to degree 12. Once we have different estimates for c_i and d_i, $i = 1, \ldots, 5$ we employ linear regression to obtain formulae relating c_i and d_i with i and so obtain $c_i = 1.741 - 0.0297\ i$ and $d_i = 0.325 - 0.03\ i$. We call these DESIGN I and II respectively. A graphical representation (not given) of these curves

d	3	4	5	6	7	8	9	10	11	12
I	98.9	99.5	99.4	99.5	99.2	98.9	98.3	97.5	96.4	95.0
II	98.9	98.8	99.3	99.8	99.9	99.5	98.4	96.5	93.8	90.3
III	100.	100.	99.7	99.7	99.7	99.8	99.8	99.9	100.	99.9
IV	98.9	98.6	98.6	98.5	98.5	98.5	98.5	98.5	98.5	98.5

Table 12.3 Efficiencies (%) for A-optimality

d	3	4	5	6	7	8	9	10	15	20
I	99.2	99.3	99.2	99.4	99.5	99.6	99.6	99.6	99.5	99.0
II	99.4	99.1	99.0	99.2	99.3	99.5	99.7	99.9	99.6	98.0
III	100.	100.	99.9	99.9	99.9	99.9	99.9	99.9	99.8	99.8
IV	99.1	98.6	98.3	98.1	98.0	97.9	97.9	97.9	98.0	98.2

Table 12.4 Efficiencies (%) for D-optimality

shows that the first crosses the correct values from above at some point and then declines further and further below them. With the second the opposite happens. This suggests that a convex combination of the support points given by these two supports, $x_{i\,3} = (1-\alpha)x_{i\,1} + \alpha x_{i\,2}$, $0 \leq \alpha \leq 1$ (DESIGN III), should achieve good results. This method gives us approximate support for the optimum design. We can again use the procedure described by Pukelsheim and Torsney (1991) to obtain the optimal weights.

The A-efficiencies for the three designs when A-optimal weights for the given support points are used are shown in Table 12.3. DESIGN III has been chosen taking $\alpha = 0.5$ and clearly gives very good efficiencies for the first polynomial models. The arcsin support points of Pukelsheim and Torsney (1991) together with their optimal weights have been included in the table (DESIGN IV).

In principle, we cannot use the same argument for the D-optimum design. This is because while the A-optimum design is always Φ_{Ch_1}-optimum, the D-optimum design agrees with the Φ_{Ch_m}-optimum design and $m = d + 1$ varies with the degree of the polynomial model. However, making the same picture for $k = m$ we get a similar representation (in fact it is also similar for $k = m - 1$, $k = m - 2$,...) and similar patterns for the supports can be used to approximate the curves. This time we used support points up to degree 40 and $i = 1, 2, \ldots, 11$. The values of the parameters were computed as $c_i = 1.7$ and $d_i = 0.229 - 0.0086\ i$, with weights fixed to $1/m$. The efficiencies are shown in Table 12.4.

Remark: We have found a good approximation for the support points using regression techniques. Something similar could be done for the weights, for cases where the process described by Pukelsheim and Torsney (1991) becomes too complex.

References

Bellhouse, R. and Herzberg, A. M. (1984). Equally spaced design points in polynomial regression: a comparison of systematic sampling methods with the optimal design of experiments. *Canadian Journal of Statistics* **12**, 77–90.

Dette, H. (1994). Discrimination designs for polynomial regression on compact intervals. *Annals of Statistics* **22**, 890–903.

Dette, H. and Studden, W. J. (1994). Optimal designs with respect to Elfving's partial minimax criterion in polynomial regression. *Ann. Inst. Statist. Math.* **46**, 389–403.

Dette, H. and Wong, W.K. (1995). On G-efficiency calculation for polynomial models. *Annals of Statistics* **23**, 2081–2101.

Dette, H. and Wong, W.K. (1996). Robust optimal extrapolation designs. *Biometrika* **83**, 667–680.

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

Hoel, P. G. and Levine, A. (1964). Optimal spacing and weighting in polynomial prediction. *Ann. Math. Statist.* **35**, 1553–1560.

Kiefer, J. and Studden, W. J. (1976). Optimal designs for large degree polynomial regression. *Annals of Statistics* **4**, 1113–1123.

Kiefer, J. and Wolfowitz, J. (1959). Optimum designs in regression problems. *Ann. Math. Statist.* **30**, 271–294.

López-Fidalgo, J. and Rodríguez-Díaz, J.M. (1998). Characteristic Polynomial Criteria in Optimal Experimental Design. In *MODA 5–Advances in Model-Oriented Data Analysis and Experimental Design* Eds A.C. Atkinson, L. Pronzato and H.P. Wynn, pp. 31–38. Heidelberg: Physica-Verlag.

Pázman, A. (1986). *Foundations of Optimum Experimental Design*. Dordrecht: Reidel.

Pukelsheim, F. and Studden, W. J. (1993). E-optimal designs for polynomial regression. *Annals of Statistics* **21**, 402–415.

Pukelsheim, F. and Torsney, B. (1991). Optimal weights for experimental designs on linearly independent support points. *Annals of Statistics* **19**, 1614–1625.

Studden, W. J. (1980). D_s-optimal designs for polynomial regression using continued fractions. *Annals of Statistics* **8**, 1132–1141.

Chapter 13

A NOTE ON OPTIMAL BOUNDED DESIGNS

Michael Sahm
University of Heidelberg, Institute for Applied Mathematics,
Im Neuenheimer Feld 294, D-69 120 Heidelberg, Germany
msahm@statlab.uni-heidelberg.de

Rainer Schwabe
University of Tübingen, Institute of Medical Biometry,
Westbahnhofstr. 55, D-72 070 Tübingen, Germany
rainer.schwabe@uni-tuebingen.de

Abstract An equivalence theorem is formulated to characterize optimal designs with general prespecified direct constraints on the design intensity. As an application locally optimal designs are obtained with bounded intensity for the logistic regression model. Moreover, it is shown that, for additive models, optimal marginally bounded designs can be generated from their optimal counterparts in the corresponding marginal models.

Keywords: constrained design, equivalence theorem, optimal design, logistic regression, bounded design intensity, additive model, marginal design

Introduction

In many applications constraints are given on the possible settings for a design variable either by the availability of experimental units with given properties or by the need to include certain preselected subjects or sites in the study. The first constraint of availability provides an upper bound for the designs under consideration, whereas the second constraint of prespecified settings results in a lower bound for the designs.

For the continuous case where the design intensity is bounded by an atom free measure from above, equivalence theorems have been propagated by Wynn (1982) and Fedorov (1989). More recent references on this topic are Cook and Fedorov (1995) and Fedorov and Hackl (1997). For particular covariate models,

Wierich (1985) obtained characterizations for the finite sample optimal design when both upper and lower bounds are given for the competing designs.

In the present note we formulate an equivalence theorem for the situation of general constraints given by upper and lower bounds for the design intensity. These constraints can be specified as arbitrary measures and, hence, both the continuous case of Wynn and Fedorov and the discrete case of Wierich are simultaneously covered. As a nonstandard application we obtain locally optimal designs with bounded intensity for the logistic regression model. Finally, optimal marginally bounded designs are constructed for additive models as a product of their optimal counterparts in the corresponding marginal models.

1. GENERAL SETTING

We consider the general linear model

$$Y(x) = f(x)^\top \beta + Z(x),$$

where f is a set of known regression functions and β is the vector of unknown parameters. Repeated experiments result in uncorrelated and homoscedastic errors Z.

Due to the generality of the problem we restrict our attention to generalized designs ξ which are defined as probability measures on the design region $\mathcal{X}$ of possible settings for the design variable x. The set of all generalized designs on $\mathcal{X}$ will be denoted by Ξ. For the definition and properties of generalized (approximate) designs we refer to the seminal paper by Kiefer (1959) and the monograph by Fedorov (1972).

The quality of a design ξ is usually measured by its information matrix $M(\xi) = \int_{\mathcal{X}} f(x) f(x)^\top \xi(dx)$ which is essentially proportional to the inverse of the covariance matrix for the best linear unbiased estimator of the parameter vector β. For the comparison of different designs we make use of so–called design criteria Φ which are convex or pseudo–convex functions of ξ based on the corresponding information matrices $M(\xi)$. The most popular criterion is that of D–optimality which aims at minimizing the determinant of the inverse $M(\xi)^{-1}$ of the information matrix. This approach minimizes the volume of the confidence ellipsoid in the case of normal errors. Other convex criteria include those based on the variance function $f(x)^\top M(\xi)^{-1} f(x)$ of the prediction $\widehat{Y}(x)$ of the response $f(x)^\top \beta$ for various x (IMSE and minimax criteria).

A powerful tool for investigating convex criteria Φ is the concept of directional derivatives. The quantity $F_\Phi(\xi, \eta) = \lim_{\alpha \searrow 0} \frac{1}{\alpha}(\Phi(\alpha\eta + (1-\alpha)\xi) - \Phi(\xi))$ is called the directional derivative of Φ at ξ in the direction of η. Denote by $F_\Phi(\xi; x) = F_\Phi(\xi, \delta_x)$ the directional derivative in the direction of the one-point (Dirac) measure δ_x. Then Φ is called linearly differentiable if $F_\Phi(\xi, \eta) = \int_{\mathcal{X}} F_\Phi(\xi; x)\, \eta(dx)$.

Let Ξ' be a convex set of competing designs, $\Xi' \subset \Xi$. A by now well–known consequence of the theory of convex analysis is the following theorem by Whittle (1973):

Theorem A.
(i) *ξ^* minimizes $\Phi(\xi)$ in Ξ' if and only if $F_\Phi(\xi^*, \eta) \geq 0$ for every design $\eta \in \Xi'$.*
(ii) *If Φ is linearly differentiable, then ξ^* minimizes $\Phi(\xi)$ in Ξ if and only if $F_\Phi(\xi^*; x) \geq 0$ for every setting $x \in \mathcal{X}$. Moreover, if ξ^* minimizes $\Phi(\xi)$, then $F_\Phi(\xi^*; x) = 0$ for ξ^*–almost all x.*

2. AN EQUIVALENCE THEOREM FOR BOUNDED DESIGNS

We are interested in the general situation of explicitly specified upper and lower bounds μ and ν for the competing designs, $\nu \leq \xi \leq \mu$, where μ and ν are arbitrary measures on the design region $\mathcal{X}$ satisfying the natural conditions $\nu \leq \mu$ and $\nu(\mathcal{X}) \leq 1 \leq \mu(\mathcal{X})$. Then the set $\Xi_\nu^\mu = \{\xi;\, \nu \leq \xi \leq \mu\}$ of competing designs is non–empty and convex. In the following we will only consider linearly differentiable design criteria.

Theorem 1. *The following three statements are equivalent:*
(a) ξ^* minimizes $\Phi(\xi)$ in Ξ_ν^μ.
(b) There exists a number s such that $F_\Phi(\xi^*; x) \leq s$ for $(\xi^* - \nu)$–almost all x and, simultaneously, $F_\Phi(\xi^*; x) \geq s$ for $(\mu - \xi^*)$–almost all x.
(c) There are subsets $\mathcal{X}_1, \mathcal{X}_2 \subset \mathcal{X}$ and a number s such that

(i) $\xi^* = \nu$ *on* $\mathcal{X}_1$ *and* $\xi^* = \mu$ *on* $\mathcal{X}_2$,
(ii) $\sup_{x \in \mathcal{X}_2} F_\Phi(\xi^*; x) \leq s \leq \inf_{x \in \mathcal{X}_1} F_\Phi(\xi^*; x)$,
(iii) $F_\Phi(\xi^*; x) = s$ *on* $\mathcal{X} \setminus (\mathcal{X}_1 \cup \mathcal{X}_2)$.

The proof of the theorem is given in the Appendix. Statement (b) of Theorem 1 may be rewritten in terms of essential suprema and essential infima as $\operatorname{ess\,sup}_{\xi^* - \nu} \mathrm{F}_\Phi(\xi^*; \mathrm{x}) \leq \operatorname{ess\,inf}_{\mu - \xi^*} \mathrm{F}_\Phi(\xi^*; \mathrm{x})$. If we let $\nu = 0$, $\Xi^\mu = \{\xi;\, \xi \leq \mu\}$, we obtain a generalization of the theorem by Wynn (1982) and Fedorov (1989) for optimal designs bounded from above:

Corollary 1. *The following three statements are equivalent:*
(a) ξ^* minimizes $\Phi(\xi)$ in Ξ^μ.
(b) There exists a number s such that $F_\Phi(\xi^*; x) \leq s$ for ξ^*–almost all x and, simultaneously, $F_\Phi(\xi^*; x) \geq s$ for $(\mu - \xi^*)$–almost all x.
(c) There are subsets $\mathcal{X}_1, \mathcal{X}_2 \subset \mathcal{X}$ and a number s such that

(i) $\xi^* = 0$ *on* $\mathcal{X}_1$ *and* $\xi^* = \mu$ *on* $\mathcal{X}_2$,
(ii) $\sup_{x\in\mathcal{X}_2} F_\Phi(\xi^*; x) \le s \le \inf_{x\in\mathcal{X}_1} F_\Phi(\xi^*; x)$,
(iii) $F_\Phi(\xi^*; x) = s$ *on* $\mathcal{X} \setminus (\mathcal{X}_1 \cup \mathcal{X}_2)$.

For optimal designs bounded from below, $\xi^* \in \Xi_\nu = \{\xi;\ \xi \ge \nu\}$, a slightly more delicate consideration yields the following result which combines features of Theorem 1 and part (ii) of Theorem A.

Corollary 2. *The following three statements are equivalent:*
(a) ξ^* *minimizes* $\Phi(\xi)$ *in* Ξ_ν.
(b) There exists a number s *such that* $F_\Phi(\xi^*; x) = s$ *for* $(\xi^* - \nu)$*–almost all* x *and, simultaneously,* $F_\Phi(\xi^*; x) \ge s$ *for all* $x \in \mathcal{X}$.
(c) There is a subset $\mathcal{X}_1 \subset \mathcal{X}$ *and a number* s *such that*
(i) $\xi^* = \nu$ *on* $\mathcal{X}_1$,
(ii) $\inf_{x\in\mathcal{X}_1} F_\Phi(\xi^*; x) \ge s$,
(iii) $F_\Phi(\xi^*; x) = s$ *on* $\mathcal{X} \setminus \mathcal{X}_1$.

In fact, letting $\nu = 0$ in Corollary 2 we get a reformulation of Whittle's Theorem A (ii) where $s = 0$. The main feature of the above results is that the design region can be split up into two subsets $\mathcal{X}_1$ and $\mathcal{X}_2$ for which the optimal design ξ^* coincides either with the lower bound or with the upper bound and for which the directional derivative at ξ^* is constant on the remaining set $\mathcal{X}\setminus(\mathcal{X}_1\cup\mathcal{X}_2)$. In particular, in the continuous case this remaining set is typically a null set with respect to μ and we obtain a zero–one law in the sense that the optimal design ξ^* coincides with either the lower or upper bound.

3. APPLICATION: BOUNDED DESIGNS FOR LOGISTIC REGRESSION

As an illustration of the equivalence theorem we present an application to a logistic regression model where the intensity is bounded. As in all nonlinear regression models the information matrix depends on the unknown parameter β, and we consider locally optimal designs which solve the optimization problem for specified values of β. Note that the concept of local optimality can be understood as a local linearization of the model. Hence, linear methods can be applied.

In particular, we consider the binary response model

$$P_x(Y = 1) \ = \ F(b(x - a)) \ = \ \frac{1}{1 + e^{-b(x-a)}}$$

where the probability of a positive response at dose x is modeled by a logistic regression. Here the dose x may be given either on a linear or a logarithmic scale. The parameter $\beta = (a, b)$ consists of the effective mean dose (ED50)

a and the slope $b > 0$, and F is the standard logistic distribution function $F(x) = (1 + e^{-x})^{-1}$. For a design ξ the asymptotic variance of the maximum likelihood estimator for β is proportional to the inverse of the Fisher information matrix

$$M(\xi;\beta) \;=\; \int_{I\!R} \begin{pmatrix} b^2 & -b\,(x-a) \\ -b\,(x-a) & (x-a)^2 \end{pmatrix} h(b(x-a))\,\xi(dx)$$

where $h(x) = F'(x)^2/(F(x)(1-F(x)) = F(x)(1-F(x))$.

For the D–criterion the derivative of the criterion function is given by

$$F_{\Phi,\beta}(\xi;x) \;=\; 2 - h(b(x-a))\,\left((-b,(x-a))M(\xi;\beta)^{-1}(-b,(x-a))^{\top}\right.\,.$$

We consider the situation where the competing designs have a bounded density. To be more specific, we want to determine the D–optimal designs in Ξ^{μ}, where the upper bound μ is given by the density $g(x) = c \cdot 1_{[D_1,D_2]}(x)$ with respect to the Lebesgue measure, i. e. $g(x) = c$ for doses x which lie between a minimal possible dose D_1 and the maximal applicable dose D_2, $c(D_2 - D_1) > 1$, and $g(x) = 0$ otherwise. Usually, $D_1 = 0$ if the dose x is measured on a linear scale and $D_1 = -\infty$ if x is measured on a logarithmic scale. In total we have a compound constraint: neither too high nor too low doses are possible and, additionally, neither are replications of 'similar' doses.

By Corollary 1 there are disjoint sets $\mathcal{X}_1$ and $\mathcal{X}_2$ such that the density g^* of the optimal design ξ^* is equal to 0 on $\mathcal{X}_1$ and is equal to c on $\mathcal{X}_2$. Furthermore, there is a number s such that $\sup_{x\in\mathcal{X}_2} F_{\Phi,\beta}(\xi^*;x) \le s \le \inf_{x\in\mathcal{X}_1} F_{\Phi,\beta}(\xi^*;x)$ and $F_{\Phi,\beta}(\xi^*;x) = s$ on $\mathcal{X} \setminus (\mathcal{X}_1 \cup \mathcal{X}_2)$.

As for unrestricted optimal designs in regression models, a powerful tool is the investigation of the number of possible modes of the directional derivative $F_{\Phi,\beta}(\xi^*;x)$.

Proposition. *In the present setting, for every design ξ the directional derivative $F_{\Phi,\beta}(\xi^*;x)$ has ,at most, three local extrema as a function of x.*

This assertion is proved essentially by showing that the derivative of the function $u(x) = F_{\Phi}(\xi^*;x)$ has at most three changes of sign which, in turn, is achieved by expressing $u'(x)$ as a power series and bounding the number of sign changes of the coefficients. A similar result for the double exponential model can be found in Dette and Sahm (1997). The proposition is also related to the geometric condition in Sitter and Wu (1993).

The proposition, together with the fact that $\lim_{x\to\pm\infty} F_{\Phi,\beta}(\xi;x) = 2$, implies that $\mathcal{X} \setminus (\mathcal{X}_1 \cup \mathcal{X}_2)$ consists of not more than four points and, hence, the density g^* of the optimal design is constant to c on, at most, two intervals with total length $1/c$.

Result. *If the constraint μ is given by the density $g(x) = c \cdot 1_{[D_1, D_2]}(x)$ then, in a logistic regression model, the locally D–optimal design ξ^* is given by the density g^* which is either of the form*

(i) $g^*(x) = c \cdot 1_{[d_1, d_2]}(x)$, *where* $D_1 \le d_1 < d_2 \le D_2$, $d_2 - d_1 = 1/c$ *and* $F_{\Phi,\beta}(\xi^*; x) \le \max(F_{\Phi,\beta}(\xi^*; d_1), F_{\Phi,\beta}(\xi^*; d_2))$ *for* $d_1 \le x \le d_2$. *Moreover,* $F_{\Phi,\beta}(\xi^*; d_1) = F_{\Phi,\beta}(\xi^*; d_2)$ *in the case* $D_1 < d_1$ *and* $d_2 < D_2$.

(ii) $g^*(x) = c \cdot 1_{[d_1, d_2]}(x) + c \cdot 1_{[d_3, d_4]}(x)$ *where* $D_1 \le d_1 < d_2 < d_3 < d_4 \le D_2$, $d_2 - d_1 + d_4 - d_3 = 1/c$ *and* $F_\Phi(\xi^*, d_2) = F_\Phi(\xi^*, d_3) \ge \max(F_{\Phi,\beta}(\xi^*; d_1), F_{\Phi,\beta}(\xi^*; d_4))$, *with equality if either* $D_1 < d_1$ *or* $d_4 < D_2$.

From these two cases the optimal designs can be obtained numerically. In Table 13.1 we exhibit a few examples of locally optimal designs for different values of the slope parameter b and the mean effective dose (ED50) a in the standardized situation, $c = 1$, $D_1 = -\infty$ and $D_2 = 0$, when the dose x is reported on a logarithmic scale.

b	a	d_1	d_2	d_3	d_4
0.5	-2	-5.98	-5.47	-0.49	0
0.5	0	-5.23	-4.70	-0.47	0
0.5	2	-4.81	-4.27	-0.46	0
1	-2	-3.80	-3.30	-0.70	-0.20
1	0	-2.83	-2.28	-0.45	0
1	2	-2.55	-1.97	-0.42	0
5	-2	-2.57	-2.07	-1.93	-1.43
5	0	-1	0	—	—
5	2	-1	0	—	—

Table 13.1 Locally D–optimal designs for a logistic regression model; bounded intensity

Note that for small values of a the constraint of a maximal dose is of no importance and the resulting locally optimal designs are only constrained by the design density bounded by one. For smaller values of the slope parameter b, the optimal design consists of two separate intervals whereas for larger values these intervals are merged into one.

Remarks. 1. The result derived above does not depend on the particular shape of the bounding density, i.e. the constant density. It also holds in the general setting $g(z) = \varrho(z)\, 1_{[D_1, D_2]}(z)$ with $\varrho(z) > 0$.

2. By translation and rescaling, any restriction of the form $\xi \leq \mu$, μ given by the density $g(z) = c\,1_{[D_1,D_2]}$, $c > 0$, can be transformed to one of the form $\widetilde{g}(z) = 1_{[D,0]}$. More precisely, let $\xi^*_{a,b,g}$ denote the locally optimal design for parameters a and b under the constraint that the design density is bounded by $g(x) = c\,1_{[D_1,D_2]}$. Then, if $D_2 < \infty$, $\xi^*_{a,b,g}(x) = \xi^*_{\widetilde{a},\widetilde{b},\widetilde{g}}(c(x - D_2))$, where $\widetilde{a} = c(a - D_2)$, $\widetilde{b} = b/c$ and $\widetilde{g}(z) = 1_{[c(D_1-D_2),0]}$.
If $D_2 = \infty$ and $D_1 > -\infty$ then by an additional reflection, we have $\xi^*_{a,b,g}(x) = \xi^*_{\widetilde{a},\widetilde{b},\widetilde{g}}(c\,(D_1 - x))$, whith $\widetilde{a} = c(D_1 - a)$, $\widetilde{b} = b/c$ and $\widetilde{g}(z) = 1_{(\infty,0]}$.
Alternatively one could only consider the design problem in the standard form $a = 0$ and $b = 1$, since, again owing to the transformation equivariance, $\xi^*_{a,b,g}(x) = \xi^*_{0,1,\widetilde{g}}(b(x - a))$ where $\widetilde{g}(z) = c/b\ 1_{[b(D_1-a),bD_2]}(z)$.

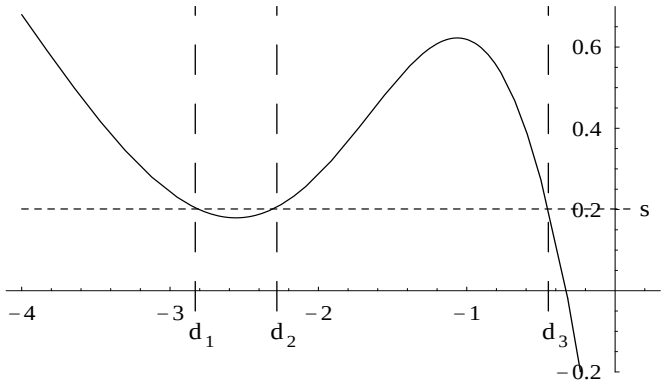


Figure 13.1 Directional derivative $F_{\Phi,\beta}(\xi; x)$ of the locally optimal design

Figure 13.1 exhibits the directional derivative $F_{\Phi,\beta}(\xi; x)$ of the locally optimal design for the parameter values $b = 1$ and $a = 0$. The reference lines indicate the support points $d_1, \ldots, d_4$ of the design (cf Table13.1) and the corresponding value for the bound s occuring in the equivalence theorem.

4. MARGINALLY BOUNDED DESIGNS

In this section we consider additive two–factor models

$$Y(x_1, x_2) = \beta_0 + f_1(x_1)^\top \beta_1 + f_2(x_2)^\top \beta_2 + Z(x_1, x_2)$$

with explicit constant term β_0, where the components x_1 and x_2 of the design variable $x = (x_1, x_2)$ may be adjusted independently, $x \in \mathcal{X}_1 \times \mathcal{X}_2$. Due to separate constraints on both variables we are looking for marginally bounded designs ξ where their one–dimensional projections ξ_1 and ξ_2 are bounded by $\nu_1 \leq \xi_1 \leq \mu_1$ and $\nu_2 \leq \xi_2 \leq \mu_2$, respectively. Note that the projections are

defined by $\xi_1(B_1) = \xi(B_1 \times \mathcal{X}_2)$ for every (measurable) subset B_1 of $\mathcal{X}_1$, and for ξ_2 accordingly. Denote by $\Xi_{\nu_1,\nu_2}^{\mu_1,\mu_2} = \{\xi;\, \nu_1 \le \xi_1 \le \mu_1,\, \nu_2 \le \xi_2 \le \mu_2\}$ the set of competing marginally bounded designs. Then we can characterize the optimal design for the D–criterion as the measure theoretic product of the optimal designs in the corresponding marginal models. This expands the decomposition results of Schwabe and Wierich (1995) and Schwabe (1996) to the situation of marginally bounded designs.

Theorem 2. *If the designs ξ_i^* are D–optimal in $\Xi_{\nu_i}^{\mu_i}$ for the marginal models given by the response functions $E(Y^{(i)}(x_i)) = \beta_0 + f_i(x_i)^\top \beta_i$, $x_i \in \mathcal{X}_i$, $i = 1,\, 2$, then the product design $\xi_1^* \otimes \xi_2^*$ is D–optimal in $\Xi_{\nu_1,\nu_2}^{\mu_1,\mu_2}$.*

Proof. As is shown by Schwabe and Wierich (1995), every design ξ is dominated by the product $\xi_1 \otimes \xi_2$ of its marginals with respect to the D–criterion, $\det M(\xi) \le \det M(\xi_1 \otimes \xi_2)$, and for product designs the determinant can be decomposed according to $\det M(\xi_1 \otimes \xi_2) = \det M_1(\xi_1) \det M_2(\xi_2)$, where $M_1(\xi_1)$ and $M_2(\xi_2)$ are the information matrices in the marginal models. Moreover, with $\xi \in \Xi_{\nu_1,\nu_2}^{\mu_1,\mu_2}$ its marginals satisfy the marginal bounds $\nu_i \le \xi_i \le \mu_i$ and, hence, the product $\xi_1 \otimes \xi_2$ is in $\Xi_{\nu_1,\nu_2}^{\mu_1,\mu_2}$. Thus an optimal design can be found in the class of product designs and the best product design is constructed as the product of the D–optimal marginal designs. □

In the case of marginally restricted designs, where the first marginal is fixed, $\xi_1 = \mu_1$, and no bounds are imposed on the second marginal ξ_2, we recover the result by Cook and Thibodeau (1980) that the product design $\mu_1 \otimes \xi_2^*$ is D–optimal for the additive model when a D–optimal marginal design ξ_2^* is used.

Obviously, Theorem 2 can be extended directly to additive models with more than two factors. In contrast to the unbounded situation a statement similar to Theorem 2 is not true for Kronecker product type models with complete interactions. This illustrates that there are structural differences between Kronecker product type and additive models.

Example. For multiple linear regression $Y(x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + Z(x_1, x_2)$ on a 20×20 point grid with uniformly bounded marginals, $\xi_i(x_i) \le \mu_i(x_i) = \frac{1}{10}$, $x_i = 1, ..., 20$, the D–optimal design is given by equal weights $\frac{1}{100}$ on each point of the four 5×5 grids located at the corners of the design region as exhibited in Figure 13.2. Note that this design differs from the design obtained by Müller (1998, p. 90) which is optimal under the constraint $\xi \le \mu_1 \otimes \mu_2$.

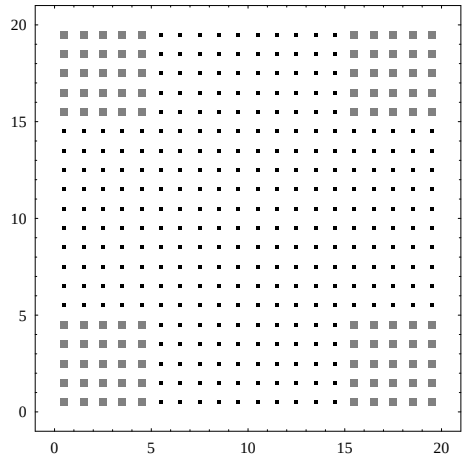


Figure 13.2 *D*–optimal design for multiple linear regression on a 20×20 grid with uniformly bounded marginals

5. APPENDIX: PROOF OF THEOREM 1

We first show that (a) implies (b) by contradiction: Assume that (b) does not hold. Then there are subsets $\mathcal{A}$ and $\mathcal{B}$ of the design region $\mathcal{X}$ such that $a := \xi^*(\mathcal{A}) - \nu(\mathcal{A}) > 0$, $b := \mu(\mathcal{B}) - \xi^*(\mathcal{B}) > 0$ and $\inf_{x\in\mathcal{A}} F_\Phi(\xi^*; x) > \sup_{x\in\mathcal{B}} F_\Phi(\xi^*; x)$. Define a competing design ξ' by $\xi' = \xi^* - \frac{\min(a,b)}{a}(\xi^* - \nu)$ on $\mathcal{A}$, $\xi' = \xi^* + \frac{\min(a,b)}{b}(\mu - \xi^*)$ on $\mathcal{B}$ and $\xi' = \xi^*$ on the complementary set $\mathcal{X} \setminus (\mathcal{A} \cup \mathcal{B})$. Then $\nu \le \xi' \le \mu$ and $F_\Phi(\xi^*, \xi') < 0$. Hence, in view of Theorem A (i) the design ξ^* cannot be optimal in contradiction to (a).

To see that (b) implies (c) we only have to set $\mathcal{X}_1 = \{x;\ F_\Phi(\xi^*; x) > s\}$ and $\mathcal{X}_2 = \{x;\ F_\Phi(\xi^*; x) < s\}$ which have the required properties.

Finally, to show that (c) implies (a) we investigate the directional derivative $F_\Phi(\xi^*, \eta)$ for every competing design η. In view of condition (i) in (c) we have $\eta \ge \xi^*$ on $\mathcal{X}_1$ and $\eta \le \xi^*$ on $\mathcal{X}_2$ for every design $\eta \in \Xi_\nu^\mu$. Hence,

$$\int_{\mathcal{X}_1} F_\Phi(\xi^*; x)\eta(dx) \ge \int_{\mathcal{X}_1} F_\Phi(\xi^*; x)\xi^*(dx) + \inf_{x\in\mathcal{X}_1} F_\Phi(\xi^*; x)(\eta(\mathcal{X}_1) - \xi^*(\mathcal{X}_1))$$

and

$$\int_{\mathcal{X}_2} F_\Phi(\xi^*; x)\eta(dx) \ge \int_{\mathcal{X}_2} F_\Phi(\xi^*; x)\xi^*(dx) - \sup_{x\in\mathcal{X}_2} F_\Phi(\xi^*; x)(\xi^*(\mathcal{X}_2) - \eta(\mathcal{X}_2)).$$

Employing conditions (ii) and (iii) of (c) we, thus, obtain

$$F_\Phi(\xi^*, \eta) = \int_{\mathcal{X}} F_\Phi(\xi^*; x)\eta(dx) \ge \int_{\mathcal{X}} F_\Phi(\xi^*; x)\xi^*(dx) + s(\eta(\mathcal{X}) - \xi^*(\mathcal{X})) = 0$$

which establishes the optimality of ξ^* in Ξ_ν^μ with respect to Φ in view of Theorem A (i). □

Acknowlegdement. The work of the second author was partly supported by the research grant Ku 719/2 of the Deutsche Forschungsgemeinschaft at the Free University, Berlin.

References

Cook, R.D. and Fedorov, V.V. (1995). Constrained optimization of experimental design. *Statistics* **28**, 129–178.

Cook, R.D. and Thibodeau, L.A. (1980). Marginally restricted D-optimal designs. *J. Amer. Statist. Assoc.* **75**, 366–371.

Dette, H. and Sahm, M. (1997). E-optimal designs for the double exponential model. In *MODA 5 – Advances in Model–Oriented Data Analysis and Experimental Design*, Eds A.C. Atkinson, L. Pronzato and H.P. Wynn, pp. 11–20. Heidelberg: Physica-Verlag.

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

Fedorov, V.V. (1989). Optimal design with bounded density: Optimization algorithms of the exchange type. *J. Statist. Plann. Inference* **22**, 1–13.

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

Kiefer, J. (1959). Optimal experimental designs. *J. R. Statist. Soc.* B **21**, 272–304.

Müller, W.G. (1998). *Collecting Spatial Data.* Heidelberg: Physica-Verlag.

Schwabe R. (1996). *Optimum Designs for Multi–Factor Models. Lecture Notes in Statistics* **113**. New York: Springer-Verlag.

Schwabe R. and Wierich W. (1995). D–optimal designs of experiments with non–interacting factors. *J. Statist. Plann. Inference* **44**, 371–384.

Sitter, R.R. and Wu, C.F.J. (1993). Optimal designs for binary response experiments: Fieller–, D– and A–criteria. *Scand. J. Statist.* **20**, 329–341.

Whittle, P. (1973). Some general points in the theory of optimal experimental design. *J. R. Statist. Soc.* B **35**, 916–921.

Wierich, W. (1985), Optimum designs under experimental constraints for a covariate model and an intra–class regression model. *J. Statist. Plann. Inference* **12**, 27-40

Wynn, H.P. (1982). Optimum submeasures with applications to finite population sampling. In *Statistical Decision Theory and Related Topics III. Proc. Third Purdue Symp.* Vol. 2., Eds S.S. Gupta and J.O. Berger, pp. 485–495. New York: Academic Press.

Chapter 14

CONSTRUCTION OF CONSTRAINED OPTIMAL DESIGNS

Ben Torsney and Saumendranath Mandal
Department of Statistics, University of Glasgow
Glasgow G12 8QW, United Kingdom
b.torsney@stats.gla.ac.uk and smandal@stats.gla.ac.uk

Abstract We consider the problem of finding an 'approximate' design maximising a criterion subject to an equality constraint. Initially the Lagrangian is formulated but the Lagrange parameter is removed through a substitution, using linear equation theory, in an approach which transforms the constrained optimisation problem to a problem of maximising two functions of the design weights simultaneously. They have a common maximum of zero which is simultaneously attained at the constrained optimal design weights. This means that established algorithms for finding optimising distributions can be considered. The approach can easily be extended to the case of several constraints, raising the 'prospect' of solving an expanded class of problem.

Keywords: constrained optimal design, multiplicative algorithms, optimizing distributions, directional derivatives, Lagrangian theory, equivalence theorem

Introduction

We consider the problem of computing designs which optimise standard regression design criteria subject to an equality constraint. Examples of such problems include determining designs subject to a given efficiency; see Pukelsheim and Rosenberger (1993). Cook and Fedorov (1995) is an invited discussion paper on constrained optimal design.

Torsney and Alahmadi (1995) considered the case of finding designs subject to zero correlations between the estimates of the underlying parameters. In particular, they considered the case of minimal support designs and transformed the constrained optimal design problem to one of maximizing a criterion with respect to two or three sets of weights or distributions. They employed ex-

tensions of the multiplicative algorithms for finding optimizing distributions developed in Torsney (1983).

This approach does not extend to other equality constraints or to constrained optimal designs subject to zero correlation when the support exceeds minimum size.

It is our purpose to consider the general case and, in fact, to extend the work of Alahmadi (1993) in which the constrained optimization problem is transformed to one of simultaneous maximization of two objective functions with respect to the design weights.

1. LINEAR DESIGN THEORY

For our purposes, consider a general linear model under which we assume independence between observations on a response variable y with constant variance σ^2 and the conditional expectation:

$$E(y|\underline{v}) = \underline{v}^T\underline{\theta} \tag{14.1}$$

where $\underline{v} \in V$, $\underline{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)^T \in \mathbb{R}^k$ are unknown parameters and $V = \{\underline{v} \in \mathbb{R}^k \colon \underline{v} = (f_1(\underline{x}), f_2(\underline{x}), \ldots, f_k(\underline{x}))^T, \underline{x} \in \mathcal{X}\}$ with $\underline{f}(\underline{x}) = (f_1(\underline{x}), f_2(\underline{x}), \ldots, f_k(\underline{x}))^T$, a vector of known regression functions and $\mathcal{X}$ a closed compact design space. That is, V is the image under f of $\mathcal{X}$. So, V is an induced design space.

An approximate design is characterised by a probability measure p defined on $\mathcal{X}$ and hence on V. In practice, we must discretise these spaces. Let $V = \{\underline{v}_1, \underline{v}_2, \ldots \underline{v}_J\}$.

Then p is characterised by a set of weights $p_1, p_2, \ldots, p_J$ satisfying $p_j \geq 0$, $j = 1, 2, \ldots, J$ and $\sum p_j = 1$, weight p_j being assigned to $\underline{v}_j$. We wish to choose $p = (p_1, p_2, \ldots, p_J)$ optimally. If $\underline{\hat{\theta}}$ is the least squares estimator of $\underline{\theta}$, then $cov(\underline{\hat{\theta}}) \propto M^{-1}(p)$, where $M(p)$ is the information matrix:

$$M(p) = \sum_{j=1}^{J} p_j \underline{v_j} \underline{v_j}^T = VPV^T = \sum_{j=1}^{J} p_j \underline{f}(\underline{x_j})\underline{f}^T(\underline{x_j}), \tag{14.2}$$

where V is the $k \times J$ matrix $[\underline{v}_1, \underline{v}_2, \ldots \underline{v}_J]$ and $P = \text{diag}(p_1, p_2, \ldots, p_J)$.

We wish to choose the proportion p_j of observations taken at $\underline{x_j}$ for good estimation of $\underline{\theta}$, by optimizing some criterion. So we have a particular case of the following problem:

Problem (P1):

Maximise $\phi(p)$ over $\mathcal{P} \equiv \{p = (p_1, p_2, \ldots, p_J) : p_j \geq 0, \sum p_j = 1\}$.

The most important design criterion in applications is that of D-optimality, in which the criterion $\phi(p) = \psi\{M(p)\} = \log\{\det(M(p))\}$, is maximised. A

D-optimal design minimises the volume of the conventional ellipsoidal confidence region for the parameters of the linear model. Other choices of maximising criteria are $\psi\{M(p)\} = -\underline{c}^T M^{-1}(p)\underline{c}$ for a given vector $\underline{c}$ (c-optimality; appropriate if there is interest only in $\underline{c}^T\underline{\theta}$) or $\psi\{M(p)\} = -\text{ tr }(AM^{-1}(p)A^T)$ and $\psi\{M(p)\} = -\log\det(AM^{-1}(p)A^T)$ for a given $s \times k$ matrix A, $s < k$ (linear optimality and D_A-optimality respectively; appropriate if there is interest in inference only for $A\underline{\theta}$).

2. OPTIMALITY CONDITIONS

In the above problems we wish to choose p to maximise $\phi(p) = \psi\{M(p)\}$ subject to $p_j \geq 0$, $\sum p_j = 1$. We need to characterise optimality conditions. We do so in terms of point to point directional derivatives.

2.1. DIRECTIONAL DERIVATIVES

Let

$$D(p, q, \varepsilon) = \phi\{(1-\varepsilon)p + \varepsilon q\}, \tag{14.3}$$

$$F_\phi\{p, q\} = \lim_{\varepsilon \downarrow 0} \frac{D(p, q, \varepsilon) - \phi(p)}{\varepsilon} = \frac{dD(p, q, \varepsilon)}{d\varepsilon} \mid_{\varepsilon=0^+} \tag{14.4}$$

and

$$F_\phi^{(2)}\{p, q\} = \frac{d^2 D(p, q, \varepsilon)}{d\varepsilon^2} \mid_{\varepsilon=0^+} . \tag{14.5}$$

Whittle (1973) called $F_\phi\{p, q\}$ the directional derivative of $\phi(.)$ at p in the direction of q. It can exist even if $\phi(.)$ is not differentiable, in which case

$F_\phi(p, q) = (q - p)^T \partial\phi/\partial p$

$$= \sum_{i=1}^{J} (q_i - p_i) d_i, \text{ where } d_i = \partial\phi/\partial p_i \ , \quad i = 1, \ldots, J.$$

Let $F_j = F_\phi(p, e_j)$, where e_j is the j^{th} unit vector in $\mathbb{R}^J$,

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

We call F_j a vertex directional derivative of $\phi(.)$ at p.

2.2. CONDITIONS FOR LOCAL OPTIMALITY

If $\phi(.)$ is differentiable at p^*, then $\phi(p^*)$ is a local maximum of $\phi(.)$ in the feasible region of problem(P1) if

$$F_j^* = F_\phi\{p^*, e_j\} \begin{cases} = 0 & : \text{ if } p_j^* > 0 \\ \leq 0 & : \text{ if } p_j^* = 0 \end{cases} \tag{14.6}$$

and

$$F_j^{*(2)} = F_\phi^{(2)}\{p^*, e_j\} \ \leq 0 \ \ if \ p_j^* > 0. \tag{14.7}$$

If $\phi(.)$ is concave on its feasible region then the first-order stationarity condition (1.6) is both necessary and sufficient for a solution to the problem (P1). This of course is the General Equivalence Theorem in Optimal Design; see Whittle (1973).

3. ALGORITHMS

Problem (P1) has a distinctive set of constraints, namely the variables $p_1, p_2, \ldots, p_J$ must be nonnegative and sum to 1. An iteration which neatly submits to these and has some suitable properties is the multiplicative algorithm:

$$p_j^{(r+1)} = p_j^{(r)} f(d_j^{(r)}) \Big/ \sum_{i=1}^{J} p_i^{(r)} f(d_i^{(r)}), \tag{14.8}$$

where $d_j^{(r)} = \frac{\partial \phi}{\partial p_j} \mid_{p=p^{(r)}}$, while $f(d)$ satisfies the following conditions:
(i) $f(d)$ is positive ;
(ii) $f(d)$ is strictly increasing in d.

This type of iteration was first proposed by Torsney (1977), taking $f(d) = d^{\delta}$, with $\delta > 0$. Subsequent empirical studies include Silvey, Titterington and Torsney (1978), which is a study of the choice of δ when $f(d) = d^{\delta}$, $\delta > 0$; Torsney (1988), which mainly considered $f(d) = e^{\delta d}$ in a variety of applications, for which one criterion $\phi(.)$ could have negative derivatives; Torsney and Alahmadi (1992) and Alahmadi (1993) which explore other choices of $f(.)$ and Mandal and Torsney (2000) which considers systematic choices of $f(.)$. Torsney (1983) establishes a result on monotonicity.

4. CONSTRAINED OPTIMIZING DISTRIBUTIONS

4.1. GENERAL PROBLEM (GP)

The general problem which we will consider is the following:
Maximise a criterion $\phi(p)$ subject to the constraint $g(p) = 0$ as well as $p_j \geq 0$, $\sum p_j = 1$.

The function $g(p)$ could be
$g(p) = \underline{a}^T M^{-1}(p)\underline{a}$ - $\underline{b}^T M^{-1}(p)\underline{b}$, if we want $var(\underline{a}^T\underline{\hat{\theta}}) = var(\underline{b}^T\underline{\hat{\theta}})$,
or $g(p) = \underline{r}^T M^{-1}(p)\underline{s}$, if we want $Cov(\underline{r}^T\underline{\hat{\theta}}, \underline{s}^T\underline{\hat{\theta}}) = 0$,
or $g(p) = \phi(p)/\phi(p_U^*) - e$, if we want an efficiency of e ($0 < e < 1$), where p_U^* is the unconstrained optimiser of $\phi(p)$.

In the last case the criterion to be optimised is $\phi(p)$. In the first case an appropriate criterion might be $\phi(p) = -\text{logdet}(AM^{-1}(p)A^T)$ or $-\text{tr}(AM^{-1}(p)A^T)$, where $A = [\underline{a}, \underline{b}]^T$. The second case is in fact equivalent to the first case with $\underline{r} = \frac{\underline{a}+\underline{b}}{2}$ and $\underline{s} = \frac{\underline{a}-\underline{b}}{2}$.

4.2. FIRST ORDER LAGRANGIAN CONDITIONS TRANSFORMED

First we establish first-order conditions for a constrained local maximum using Lagrangian theory applied only to the constraint $g(p) = 0$.

Let

$$L(\phi, p, \lambda) = \phi(p) + \lambda g(p). \tag{14.9}$$

Then for $i = 1, 2, \ldots, J$

$$d_i^L = \frac{\partial L}{\partial p_i} = \frac{\partial \phi}{\partial p_i} + \lambda \frac{\partial g}{\partial p_i} = d_i^\phi + \lambda d_i^g. \tag{14.10}$$

Also the vertex directional derivatives of L are

$$\begin{aligned} F_i^L &= d_i^L - \sum p_j d_j^L \\ &= d_i^\phi - \sum p_j d_j^\phi + \lambda (d_i^g - \sum p_j d_j^g) \\ &= F_i^\phi + \lambda F_i^g, \end{aligned} \tag{14.11}$$

where $F_i^\phi = d_i^\phi - \sum p_j d_j^\phi$ and $F_i^g = d_i^g - \sum p_j d_j^g$. Equivalently $\underline{F}^L = \underline{F}^\phi + \lambda \underline{F}^g$.

Now we acknowledge the constraint $\sum p_j = 1$, but, for the moment, assume $p_j > 0,\ j = 1, 2, \ldots, J$. Then from the equivalence theorem, or Section 2.2, conditions for a constrained local turning point of L (subject to $\sum p_j = 1$) are

$$\begin{aligned} F_j^L &= 0 \qquad \text{given } p_j > 0,\ j = 1, 2, \ldots, J, \\ \text{i.e.} \quad \underline{F}^L &= \underline{0}, \\ \text{i.e.} \quad \underline{F}^g \lambda &= -\underline{F}^\phi. \end{aligned} \tag{14.12}$$

Now suppose $A = \underline{F}^g$, $\underline{b} = -\underline{F}^\phi$ and $\underline{\lambda} = \lambda$. Then $\underline{\lambda}$ must satisfy

$$A\underline{\lambda} = \underline{b}. \tag{14.13}$$

We now deal with λ using the approach of Alahmadi (1993). The set of solutions to the system of equations (1.13), if solutions exist, is given by:

$$\underline{\lambda} = A^- \underline{b} + (I - A^- A)\underline{z} \qquad \text{for any } \underline{z}, \tag{14.14}$$

where A^- is any generalised inverse of A. These are solutions iff $AA^-\underline{b} = \underline{b}$.

Assuming $A^T A$ is nonsingular one choice of A^- is

$$A^- = (A^T A)^{-1} A^T = [(\underline{F}^g)^T \underline{F}^g]^{-1} (\underline{F}^g)^T. \tag{14.15}$$

Then $A^-\underline{b} = \frac{-(\underline{F}^g)^T \underline{F}^\phi}{(\underline{F}^g)^T \underline{F}^g} = \hat{\lambda}$ and $AA^-\underline{b} = \hat{\lambda}\, A = \hat{\lambda}\, \underline{F}^g$. Thus p^* should be such that

$$\hat{\lambda}\underline{F}^g = -\underline{F}^\phi \quad i.e. \quad \hat{\lambda}\underline{F}^g + \underline{F}^\phi = \underline{0} \tag{14.16}$$

and then substituting by $\hat{\lambda}$ in (1.16), we get

$$-\frac{(\underline{F}^g)^T\underline{F}^\phi}{(\underline{F}^g)^T\underline{F}^g}\underline{F}^g + \underline{F}^\phi = \underline{0}, \tag{14.17}$$

$$\text{or} \quad \underline{h} = [(\underline{F}^g)^T\underline{F}^g]\underline{F}^\phi - [(\underline{F}^g)^T\underline{F}^\phi]\underline{F}^g = \underline{0} \tag{14.18}$$

$$\text{i.e.} \quad \underline{h}^T\underline{h} = 0. \tag{14.19}$$

In (1.19)

$$\begin{aligned} h_i &= [(\underline{F}^g)^T\underline{F}^g]F_i^\phi - [(\underline{F}^g)^T\underline{F}^\phi]F_i^g \\ &= [\sum_{j=1}^{J}(F_j^g)^2]F_i^\phi - [\sum_{j=1}^{J}(F_j^gF_j^\phi)]F_i^g. \end{aligned} \tag{14.20}$$

So, since $\underline{h}^T\underline{h} \geq 0$, p^* should minimise $\underline{h}^T\underline{h}$ or maximise $\{-\underline{h}^T\underline{h}\}$ with a maximum value of zero.

Alternatively

$$\underline{h}^T\underline{h} = [(\underline{F}^g)^T\underline{F}^g]^2[(\underline{F}^\phi)^T\underline{F}^\phi][1-R], \tag{14.21}$$

where

$$R = \frac{[(\underline{F}^g)^T\underline{F}^\phi]^2}{[(\underline{F}^\phi)^T\underline{F}^\phi][(\underline{F}^g)^T\underline{F}^g]} = \frac{\xi_1^2}{\xi_2\,\xi_3}, \tag{14.22}$$

$\xi_1 = (\underline{F}^g)^T\underline{F}^\phi$, $\xi_2 = (\underline{F}^\phi)^T\underline{F}^\phi$ and $\xi_3 = (\underline{F}^g)^T\underline{F}^g$. Clearly $0 \leq R \leq 1$. So p^* should maximise [R-1] yielding a maximum value of zero.

Now consider the possibility of zero p^*, which was not considered by Alahmadi (1993). Then, by the equivalence theorem,

$$F_j^L \begin{cases} = 0 & : \text{ if } p_j^* > 0 \\ \leq 0 & : \text{ if } p_j^* = 0. \end{cases} \tag{14.23}$$

There are two ways in which we can adapt the above approach to this case. One is to argue that

$$\begin{aligned} l_i &= p_ih_i = 0 \quad \text{for } i = 1, 2, \ldots, J, \\ \text{i.e.} \quad \underline{l} &= P\underline{h} = 0, \end{aligned} \tag{14.24}$$

where P = diag$(p_1, p_2, \ldots, p_J)$. So we replace $\underline{h}$ in the argument above by $\underline{l} = P\underline{h}$ to conclude that p^* should maximise

$$Q = -\underline{h}^TP^2\underline{h}, \tag{14.25}$$

yielding a maximum value of zero.

A second option is to argue that

$$\begin{aligned} E_j^L &= p_j F_j^L = 0 \quad \text{for } j = 1, 2, \ldots, J \\ \text{or} \quad \underline{E}^L &= P\underline{F}^L = \underline{0}. \end{aligned} \tag{14.26}$$

So we replace $\underline{F}^L$ in the above argument by $\underline{E}^L$, concluding that we should have at p^*

$$\underline{h} = \underline{h}(\underline{E}) = \underline{0}. \tag{14.27}$$

Hence p^* should maximise $Q = -\underline{h}^T \underline{h}$ yielding a maximum of zero.

So we have dealt with the Lagrange multiplier by 'substitution' in terms of p and transformed attainment of the first order conditions in the Lagrangian to an example of (P1), where $\phi_L(p) = Q$ with several possible choices of Q. Then p^* must maximise any increasing function of Q. We consider some such transformations but choose them and Q such that the maximum of the revised criterion is still zero e.g. $(e^Q - 1)$ or $\{-\log(1 - Q)\}$.

The reason for this is that we must still ensure the constraint $g(p^*) = 0$. We deal with this by transformation to an optimization problem too; namely p^* should maximise $G(p) = -[g(p)]^2$, subject to $p_j \geq 0$, $\sum p_j = 1$. So $G(p^*) = 0$.

Thus p^* should simultaneously maximise $G(p)$ and $Q(p)$, with a common maximum of zero if $Q(p^*) = 0$ subject to $p_j \geq 0$, $\sum p_j = 1$. The sharing of a common optimal value of zero means that we have transformed the original constrained optimization problem to an example of (P1) with various choices of $\phi(p)$ which we here denote by $\phi_L(p)$, namely

$$\begin{aligned} \phi_L(p) = \phi_1(p) &= G(p) + Q(p), \\ \phi_L(p) = \phi_2(p) &= \alpha\, G(p) + (1 - \alpha)\, Q(p) \quad 0 < \alpha < 1, \\ \phi_L(p) = \phi_3(p) &= \min\{G(p), Q(p)\}. \end{aligned}$$

Hence we can consider applying methods for solving (P1).

First we illustrate the above theory by simple examples in the context of quadratic regression on three points $:-1, 0, 2$. Thus the design corresponds to the distribution

$$\begin{array}{lccc} x & : & -1 & 0 & 2 \\ p(x) & : & p_1 & p_2 & p_3 \end{array}$$

and the information matrix is

$$M(p) = \begin{pmatrix} 1 & -p_1 + 2p_3 & p_1 + 4p_3 \\ -p_1 + 2p_3 & p_1 + 4p_3 & -p_1 + 8p_3 \\ p_1 + 4p_3 & -p_1 + 8p_3 & p_1 + 16p_3 \end{pmatrix}.$$

We focus on constraints concerning the element $(M^{-1}(p))_{23}$ of $M^{-1}(p)$. This is given by

$$\begin{aligned} C &= [(8p_3 - p_1) - (2p_3 - p_1)(4p_3 + p_1)]/det(M(p)) \\ &= [(p_2 - 1/2)^2 - 9(p_3 - 1/2)^2 + 2]/det(M(p)). \end{aligned} \quad (14.28)$$

Of course the numerator is quadratic in the weights. We first consider the constraint $C = 0$, i.e. taking $g(p) = C$ so that we want $Cov(\hat{\theta}_2, \hat{\theta}_3)$ = 0. To produce plots we deal with this constraint and the '$\sum p_j$ = 1' constraint by substitution to transform the constrained optimization problem to an unconstrained optimization with respect to one weight, thereby ensuring $G(p)$ = 0.

Of the various possibilities only the following formulae are valid here and there is no restriction on p_2:

$$\begin{aligned} p_1 &= (\frac{1}{2} - p_2) + \frac{1}{3}\sqrt{(p_2 - \frac{1}{2})^2 + 2}, \\ p_3 &= \frac{1}{2} - \frac{1}{3}\sqrt{(p_2 - \frac{1}{2})^2 + 2}. \end{aligned}$$

We now show plots of $\phi(p_2)$, $Q(p_2)$ for $\phi(p)$ = logdet$(M(p))$, $-$tr$(M^{-1}(p))$. See Fig. 14.1 and Fig. 14.2. Clearly both functions are unimodal and both are maximised at p_2^* = 0.4925325 with $Q(p_2^*)$ = 0.

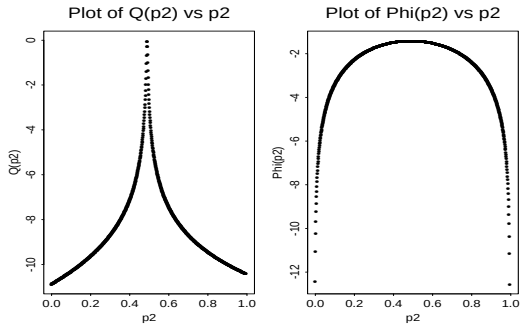


Figure 14.1 Plots of $Q(p_2)$ and $\phi(p_2)$ versus p_2 for D-optimality

We also consider the constraint det$(M(p))$ C = 2, since this leads to linear constraints on the weights. It is simpler to substitute for p_1 and p_2 in terms of p_3. Relevant formulae are:

$$\begin{aligned} &p_1 = -2(2p_3 - 1), \quad p_2 = 3p_3 - 1 \quad \text{for} \quad \tfrac{1}{3} \le p_3 \le \tfrac{1}{2}, \\ &p_1 = 2p_3 - 1, \qquad\quad p_2 = 2 - 3p_3 \quad \text{for} \quad \tfrac{1}{2} \le p_3 \le \tfrac{2}{3}. \end{aligned}$$

Plots of $\phi(p_3)$ and $Q(p_3)$ (Fig. 14.3 and Fig. 14.4) reveal that both functions now have two maximal turning points, at the same values of p_3, one in each

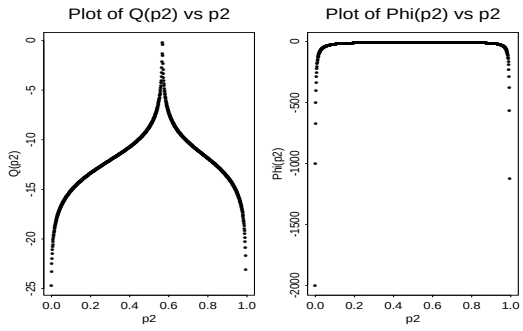


Figure 14.2 Plots of $Q(p_2)$ and $\phi(p_2)$ versus p_2 for A-optimality

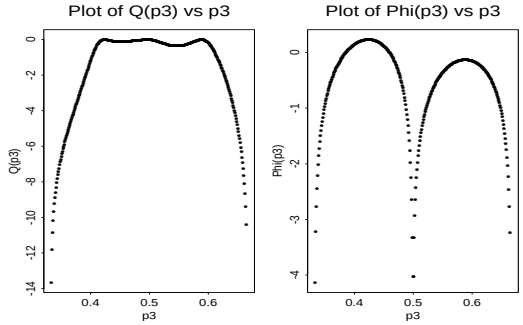


Figure 14.3 Plots of $Q(p_3)$ and $\phi(p_3)$ versus p_3 for D-optimality

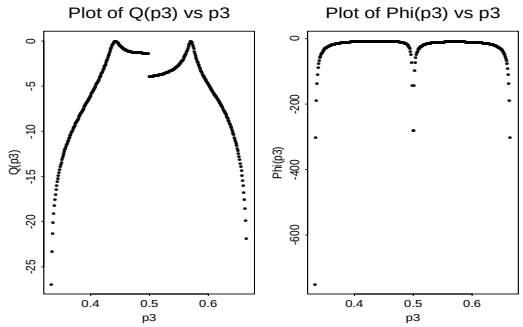


Figure 14.4 Plots of $Q(p_3)$ and $\phi(p_3)$ versus p_3 for A-optimality

of the regions $p_3 < 0.5$ and $p_3 > 0.5$. Further $Q(p_3)$ has a maximum of zero at both turning points, while $\phi(p_3)$ is maximised at the lower value, i.e. p_3 = 0.442635.

5. EMPIRICAL RESULTS

We have transformed the original constrained optimization problem (GP) to examples of (P1). So we consider the class of algorithms outlined in Section 3.

First, for the above quadratic model, we consider deriving the constrained D-optimal design subject to $Cov(\hat{\theta}_2, \hat{\theta}_3) = 0$. At equal initial weights with $Q = -(1-R)^2 = Q_1$, say, $(G, Q) = (-28.4444, -1.0), -67.556 \leq F_i(G) \leq 39.111$ and $-83508 \leq F_i(Q) \leq 144240$. After 200 iterations starting from this design with $f(x) = \exp(\delta x)$, $x = F$ and $\delta = 0.001$, we converge to the design above i.e. $p_2^* = 0.4925325$ and the new values are $(G, Q) = (-2.310^{-8}, -9.41^{-7}), -0.0012280 \leq F_i(G) \leq 0.0415182$ and $-0.0640491 \leq F_i(Q) \leq 0.0.0027849$. These derivatives could be closer to zero. This illustrates that while G and various choices of Q can be approximately zero, directional derivatives can be distinctly non zero.

We further consider two examples considered by Silvey et al (1978), Wu (1978) and Alahmadi (1993). The examples are defined by their design spaces.
Example-1. $\mathcal{V} = \{(1,-1,-1)^T, (1,-1,1)^T, (1,1,-1)^T, (1,2,2)^T\}$.
Example-2. $\mathcal{V} = \{(1,-1,-2)^T, (1,-1,1)^T, (1,1,-1)^T, (1,2,2)^T\}$.
The design criterion is $\phi(p) = -\text{tr}(AM^{-1}(p)A^T)$, where $A = [\underline{a}, \underline{b}]^T$, $\underline{a} = (1,0,0)$ and $\underline{b} = (0,0,1)$ while the constraint is $Cov(\hat{\theta}_1, \hat{\theta}_3) = 0$. Alahmadi (1993) took $\phi_L(p) = \phi_1(p)$. He too found convergence to be slow with $Q = Q_1$, $f(x) = (1+sx)^{s\delta}$, $x = d$, $s = \text{sign}(x)$, $\delta = 0.3$, 0.2 for Examples 1 and 2 respectively. Approximate solutions seemed to be (0.237, 0.270, 0.330, 0.163) and (0.255, 0.355, 0.215, 0.175) for Examples 1 and 2 respectively.

We have endeavoured to improve on this with a view to accommodating zero optimal weights. We took $\phi_L(p) = \phi_3(p)$ since there is a danger of converging to zero values for $\underline{F}^L = \underline{F}^G + \underline{F}^Q$ but not for $\underline{F}^G$, $\underline{F}^Q$. We note only preliminary results which highlight some difficulties.

(a) In Example-1 starting from equal initial weights with $Q = -(1-e^{Q_2})^2$, $Q_2 = \xi_1^2 - \xi_2\,\xi_3$, $f(x) = (1+sx)^{s\delta}$, $s = \text{sign (x) and } x = F$, the pair (G, Q) change from $(-0.177, -0.153)$ to $(-0.040, -0.051)$ after 700 iterations. Yet at this point $-1.393 \leq F_i(G) \leq 2.045, -1.884 \leq F_i(Q) \leq 2.017$.

To explore this issue further, in the remaining illustrations, we take $p^{(0)}$ to be the approximate optimal solution of Alahmadi (1993). We denote this by $p^{(A)}$.

(b) In Example-1 $(G, Q) = (-5.210^{-8}, -0.0298)$ with $-38.7634 \leq F_i(Q) \leq 81.6632$ at $p^{(A)} = (0.237, 0.270, 0.330, 0.163)$ when $Q = -(\log(1-Q_3))^3$, $Q_3 = -\underline{h}^T P^2 \underline{h}$. After 354 iterations with $f(x) = (1+sx)^{s\delta}$, $s = \text{sign } (x)$ and $x = F$, $(G, Q) = (-0.00001, -0.00001)$ with $-0.020 \leq F_i(G) \leq 0.025, -0.491 \leq F_i(Q) \leq 0.355$.

(c) In Example-1 again $(G, Q) = (-0.000087, -0)$ with $F_i(Q) = 0, i = 1,2,3,4$ while $-0.057 \leq F_i(G) \leq 0.070$, at a variation of $p^{(A)}$ when $Q = -(\log(1-Q_2))^2$. After 100 iterations with $f(x) = \exp(\delta x), x = F$ and $\delta =$

$0.001, (G, Q) = (-0.000001, 0)$ while $F_i(Q) = 0$ and $-0.0070 \leq F_i(G) \leq 0.0085$.
(d) In Example-2 for A-optimality $(G, Q) = (-0.000005, 0)$ with $F_i(Q) = 0$ while $-0.0134 \leq F_i(G) \leq 0.0121$ at $p^{(A)}$ when $Q = Q_1$. After 300 iterations with $f(x) = \exp(\delta x), x = F$ and $\delta = 0.00001, G = Q = -1.010^{-8}, F_i(Q) = 0$ and $-0.00017 \leq F_i(G) \leq 0.00015$.

6. DISCUSSION

There is clearly difficulty in attaining the first-order conditions, at least in respect of Q. This is not surprising in view of the plots of Q in Fig. 14.1 and Fig. 14.2. The function seems virtually non-differentiable at the optimum while being convex with steep derivatives on either side. These are hostile conditions for the iterations of Section 3.

Further development is needed. Possibly further transformations of Q to a more concave shape would help e.g.

$$-Q = -\tilde{Q} \begin{cases} = \{-Q\}^t & \text{for} \quad -Q < 1 \\ = \{-Q\}^{1/t} & \text{for} \quad -Q > 1, \end{cases}$$

for some t. Modifications using the clustering approach of Mandal and Torsney (2000) to deal with zero optimal weights is another option.

There is also the issue of multiple local maxima to be considered. This and the extension to the case of several equality constraints, which in principle is straightforward, is the focus of ongoing work.

Acknowledgements

One of the authors, Saumendranath Mandal, acknowledges the financial support of a University of Glasgow Scholarship and a UK CVCP Overseas Research Student (ORS) Award (96017016).

References

Alahmadi, A. M. (1993). *Algorithms for the Construction of Constrained and Unconstrained Optimal Designs*. Ph.D. Thesis, University of Glasgow.

Cook, D. and Fedorov, V. (1995). Constrained optimization of experimental design. *Statistics* **26**, 129–178.

Mandal, S. and Torsney, B. (2000). Algorithms for the construction of optimizing distributions. *Communications in Statistics: Theory and Methods* (to appear).

Mandal, S. and Torsney, B. (2000). Construction of optimal designs using a clustering approach (in preparation).

Pukelsheim, F. and Rosenberger, J. L. (1993). Experimental designs for model discrimination. *J. Am. Statist. Assoc.* **88**, 642–649.

Silvey, S. D., Titterington, D. M. and Torsney, B. (1978). An algorithm for optimal designs on a finite design space. *Communications in Statistics* A **14**, 1379–1389.

Torsney, B. (1977). Contribution to discussion of 'Maximum Likelihood estimation via the EM algorithm' by Dempster et al. *J. Roy. Statist. Soc.* B **39**, 26–27.

Torsney, B. (1983). A moment inequality and monotonicity of an algorithm. *Proc. Internat. Symp. on Semi-Infinite Programming and Appl.* Eds K.O. Kortanek and A.V. Fiacco, pp. 249–260. *Lecture Notes in Econom. and Math. Sys.* **215**. Heidelberg: Springer-Verlag.

Torsney, B. (1988) Computing optimizing distributions with applications in design, estimation and image processing. *Optimal Design and Analysis of Experiments* Eds Y. Dodge, V.V. Fedorov and H.P. Wynn, pp. 361–370. Amsterdam: Elsevier.

Torsney, B. and Alahmadi, A. M, (1992). Further development of algorithms for constructing optimizing distributions. *Model Oriented Data Analysis.* Proc. 2nd IIASA Workshop in St. Kyrik, Bulgaria, Eds V.V. Fedorov, W.G. Müller and I. Vuchkov, pp. 121–129. Heidelberg: Physica-Verlag.

Torsney, B. and Alahmadi, A. M. (1995). Designing for minimally dependent observations. *Statistica Sinica* **5**, 499–514.

Whittle, P. (1973). Some general points in the theory of optimal experimental design. *J. Roy. Statist. Soc.* B **35**, 123–130.

Wu, C. F. J. (1978). Some iterative procedures for generating nonsingular optimal designs. *Communications in Statistics* A **14**, 1399–1412.

PART TWO

A P P L I C A T I O N S

Chapter 15

PHARMACEUTICAL APPLICATIONS OF A MULTI-STAGE GROUP TESTING METHOD

Brian Bond and Valeri Fedorov
Smithkline Beecham Pharmaceuticals
Statistical Sciences Department
Brian_C_Bond@sbphrd.com and Valeri_V_Fedorov@sbphrd.com

Matthew Jones and Anatoly Zhigljavsky
Cardiff University
School of Mathematics
JonesMC1@Cardiff.ac.uk and ZhigljavskyAA@Cardiff.ac.uk

Abstract An important problem in pharmaceutical research is whether individual testing of components should be made, or alternatively, if the components should be tested in groups. It is important that the experiment is economically viable since, for multi-stage procedures, the cost of additional stages must be taken into consideration along with the cost of testing the mixtures of components. Optimum group sizes are calculated for two-stage and three-stage members of Li's family of algorithms and for row-and-column procedures, along with the minimum number of tests required to determine all of the active components. Finally, comparisons are made between the costs of the one-, two- and three-stage procedures using two different cost functions for the cost of testing mixtures of components.

Introduction

Pharmaceutical background

High-throughput (HTP) drug screening is an important and widely used strategy for identifying novel compounds as leads for future work in the pharmaceutical industry. Leads are later subjected to further biological investigation through functional assays on whole cells or tissue. They are also used in initiating chemical exploration for better compounds through techniques such as quantitative structure activity relationships (QSAR).

As company collections of compounds grow, technologies have been and continue to be developed for efficient screening using minimal resources in smaller time scales, whilst maintaining the quality of the resulting data. One method implemented by a number of companies is that of pooling samples. If a pooled sample produces a positive result, the individual components are assayed separately to identify which individual compound or compounds are active.

A number of statistical issues arise from this strategy:

- What is the optimal number of compounds to pool?
- How will the strategy affect false positive and false negative error rates?
- How does this strategy compare to a single-stage screen of assaying all individual compounds?

Additionally there are a number of practical constraints:

- Much of the assay work is performed by robots and hence the final strategy needs to be compatible with robotics;
- Random access to samples of single compounds in the compound library is performed manually and adds a substantial amount of time to a multi-stage process.

New assay technologies such as miniaturisation, single molecule detection and chip-based assays are fast improving throughput and out-pacing the time and resource savings a two-stage procedure makes. Complete data from all compounds is perceived to be a clear advantage of running the single-step process when data-mining of the resulting database is expected to become the norm.

Statistical background

For simplicity, assume that we are only interested in the activity of a large number of different components, potential drug compounds, against certain enzymes relating to certain diseases.

In a particular test, either an individual component or a mixture of them is tested for activity. The result of the test is in the form of a numerical value. If we were to look at the distribution of the activity of all the components we would find a positively-skewed distribution. We are only interested in the components that are on the far right tail of the activity scale. In a typical experiment we would search for the $d = 50$ most active components among a total of $n = 500\,000$ components.

A cut-off point is specified whereby, if a component has an activity reading greater than this point, it is deemed to be a hit. However, if the activity is

less, the component is judged to be inactive. When a component is labelled as being active when in fact it isn't, it is called a *false positive*. In cases where the component is judged to be inactive when it is active, it is said to be a *false negative*. After all of the active components have been identified they are ranked in ascending order of activity, which determines their importance relative to one another.

The problem with *false positives* could be overcome in an obvious way: following the main tests components are individually retested for their activity. The issue about *false negatives* is only related to those components around the cut-off point. We assume here that the errors in the tests are reasonably small, therefore the *false negative* problem does not concern the very active components.

There are two slightly different ways to formalise the problem as a rigorous group testing problem. One way is to consider it as the so-called *hypergeometric group testing problem* where the number d of active components is fixed. This problem is mathematically identical to the problem called the *search for defectives* (see Sobel and Groll, 1958). Alternatively, d can be the upper bound for the number of defectives. Another way, known as the *binomial group testing problem*, is to assume that the probability of finding an active component by one simple trial is $p = d/n$, the activity of different components being independent.

An important problem in pharmaceutical research is whether components should be tested individually or, alternatively, if the components should be tested in groups. We call a group *active* if it contains at least one active component, assuming that in an experiment we are able to detect the activity of a group without error.

The main difference between the present study and the papers on group testing to be mentioned below is the consideration both of costs (penalties) for additional stages and for the number of components in a test group.

Let Λ represent the cost incurred between successive stages and let c_s be the cost of testing a mixture of s components. We shall assume that $c_1 = 1$ (that is the cost of individually testing the components is 1). Let λ represent the normalised cost between successive stages i.e. $\lambda = \frac{\Lambda}{n}$.

Two simple cost functions that can be used are:

(i) $c_s = 1 + \kappa s^{\gamma}$ with $0 \leq \kappa < 1$ and $0 \leq \gamma \leq 1$;

(ii) $c_s = 1 + \kappa \log s$ with $\kappa \geq 0$.

We thus parameterise the costs with an additional two or three parameters, namely λ, κ, and perhaps γ. If $\kappa = 0$, the cost function in (i) is one and the cost of the experiment is exactly the number of tests.

1. MINIMISING THE NUMBER OF TESTS

Different group testing strategies, as well as upper and lower bounds for the length of optimal algorithms, have been extensively studied for both formulations of the problem. In this section we ignore the costs (that is assume $\lambda = 0$ and $c_s = 1$ for all s) and characterise different methods by the number of tests.

We provide references for some of the best-known results, always assuming, which is in agreement with the practical requirements, that the total number of components n is large, the number of active components d is relatively small and, in the binomial group testing model, the probability that a random component is active, p, is small.

The origin of group-testing is credited to R. Dorfman (1943), from whose work future studies stemmed. Sobel and Groll (1959) extensively studied the binomial group testing model. In their main procedure components that are proven to be active or inactive are never used in the subsequent tests. At every stage all other components are required to be separated into at most two sets. One, which is called the *defective set*, contains at least one active component. In the other, the *binomial set*, the components act like independent binomial random variables with probability p of being active. Letting EN denote the expected number of group tests remaining to be performed, Sobel and Groll show that

$$EN \cong -n \log_2 \left(\frac{1}{(1-p)} \right) + np \log_2 \left[\log_2 \frac{1}{(1-p)} \right]^{-1}, \quad n \to \infty . \quad (15.1)$$

Li's s-stage algorithm (1962) was tuned to minimise the number of tests in the worst case when using combinatorial group testing to detect the d active components. At the first stage the n components are divided into g_1 groups of k_1 or $k_1 - 1$ components. All groups are then tested and those that are inactive are removed. In general, at stage i, $2 \leq i \leq s$, items from the active groups of stage $i-1$ are pooled and arbitrarily divided into g_i groups of size k_i or $k_i - 1$ and a test performed on each group. The value of k_s is one; thus every component is identified at stage s. Letting N denote the number of tests required to determine the active components, it has been found that

$$N \leq e\, d(\log n - \log d), \quad \text{where } \; e = 2.7182818.... \quad (15.2)$$

Hwang's (1972) generalised binary splitting algorithm is an extension of the binary splitting procedure, which involves the partitioning of n components into two disjoint groups such that neither has size exceeding $2^{\lceil \log_2 n \rceil - 1}$, where $\lceil x \rceil$ denotes the smallest integer value larger than or equal to x. One group is then tested, the outcome indicating that either the test group or the other group is active. Hwang suggested a way to include d, the number of active components, into applications of binary splitting such that the total number of tests could be reduced.

The algorithm is as follows:

(i) If $n \leq 2d-2$ perform individual testing. If $n \geq 2d-1$, define $\alpha = \lfloor \log_2((n-d+1)/d) \rfloor$, where $\lfloor x \rfloor$ denotes the largest integer value less than or equal to x;

(ii) Test a group of size 2^α. If inactive, the group is identified as bad, set $n = n - 2^\alpha$ and return to Step 1. If the outcome is active use, binary splitting to identify an active component and an unspecified number say x of inactive components. Let $n = n - 1 - x$ and $d = d - 1$. Return to Step 1.

If n is large enough, the number of tests for this algorithm satisfies

$$N \leq d(\log_2 n - \log_2 d + 3). \tag{15.3}$$

General formulae for the expected number of tests to determine active components in multi-stage procedures are also discussed in Patel (1962).

An alternative literature on the hypergeometric group testing problems deals with the probabilistic technique of derivation of existence theorems for one-stage designs. The pioneering work in this area was by Renyi (1965). It has been successfully continued by many authors; some examples are listed in Du and Hwang (1993). For a fixed number of active components d and $n \to \infty$, the best known upper bound was derived in Dyachkov, Rykov and Rashad (1989); see also Du and Hwang (1993), p.68:

$$N \leq dA_d(1+o(1)) \log_2 n,$$

where

$$\frac{1}{A_d} = \max_{0\leq q\leq 1} \max_{0\leq Q\leq 1} \Big\{ -(1-Q)\log_2(1-q^d) + dQ \log_2 \frac{q}{Q} + d(1-Q)\log_2 \frac{1-q}{1-Q} \Big\} \tag{15.4}$$

and $A_d = \frac{2}{d}\log_2 e(1+o(1))$ when $d \to \infty$. Asymptotically, when both n and d are large,

$$N \leq N_*(n,d) \sim \frac{e}{2} d^2 \log n\,, \quad n \to \infty,\ d = d(n) \to \infty,\ d(n)/n \to 0.$$

In the case where d is fixed and the number of components in every test group, say s, is also fixed, the upper bound for the length of the optimum one-stage design is derived in Zhigljavsky (1998): $N \leq N^* = N^*(n,d,s)$ where N^* is the minimum over $k = 1, 2, \ldots$ such that

$$\frac{1}{2}\sum_{i=0}^{d-1} \binom{n}{i\ \ d-i\ \ d-i\ \ n-2d+i} \left(1 - 2 \cdot \frac{\binom{n-t}{s} - \binom{n-2d+i}{s}}{\binom{n}{s}} \right)^k < 1, \tag{15.5}$$

where $\binom{n}{a\ b\ c\ d} = n!/(a!b!c!d!)$ is the multinomial coefficient. When $n \to \infty$, the results in Zhigljavsky (1998) imply that $N(n,d,s) = \lceil N^{(\mathrm{as})}(n,d,s) + o(1) \rceil$, where

$$N^{(\mathrm{as})}(n,d,s) = \frac{(d+1)\log n - \log(d-1)! - \log 2}{-\log(1 - 2\frac{s}{n}(1-\frac{s}{n})^d)}\,. \tag{15.6}$$

Analogous results hold when d is the upper bound for the number of active components. Optimisation of the right-hand side of (15.6) with respect to s, the size of the test groups, gives $s_{opt} = s(n) = n/(d+1)$ and

$$\min_s N^{(\mathrm{as})}(n,d,s) \sim \frac{e}{2}\, d^2 \log n\,.$$

The approximations (upper bounds) for the lengths of different group testing strategies are compared in Table 15.1 for $n = 500\,000$ and $d = 10, 50,$ and 100. (The corresponding values of p are $0.00002,\ 0.0001,$ and 0.0002).

As we see from Table 15.1, the multi-stage strategies are approximately d times better than the best one-stage procedures. However, the situation changes when the cost λ for additional stages is taken into account. We show below, for reasonable values of λ in multi-stage strategies from Li's family of algorithms, that three or more stages become less efficient than the one-stage and two-stage strategies; see Figures 15.3 (a) and 15.4 (a). The same holds for the other sequential algorithms.

Formula (15.6) which gives the upper bound for the length of the optimal one-stage procedure can easily be extended to calculate the costs; (15.6) implies that there exist one-stage procedures with normalised cost

$$C^{(\mathrm{as})}(n,d,s) = \frac{c_s}{n} \cdot \frac{(d+1)\log n - \log(d-1)! - \log 2}{-\log(1 - 2\frac{s}{n}(1-\frac{s}{n})^d)}\,. \tag{15.7}$$

For the cost function $c_s = 1 + \kappa \log s$, optimisation of the right-hand side of (15.7) with respect to s again gives the asymptotically optimum rate $s = n/(d+1)$. In the case of the cost function $c_s = 1 + \kappa s$, $0 < \kappa < 1$, the individual testing procedure ($s = 1$) is asymptotically optimum.

2. TWO-STAGE PROCEDURE

A typical procedure used in the pharmaceutical industry to detect active components is the classical Dorfman procedure (see Dorfman (1943); short description can also be found in Feller (1960), Chapter 9, Exercise 26), which is a particular case of Li's family of algorithms and is described as follows: the motherplate consists of m columns and k rows, giving in total km cells, where each cell contains a different component (assume for simplicity $n = km$). At the first stage, a mixed sample of the m components in each row is taken and

Procedure	$d = 10$ ($p = 0.00002$)	$d = 50$ ($p = 0.0001$)	$d = 100$ ($p = 0.0002$)
Sobel and Groll procedure (15.1)	137	566	1 032
Li's *s*-Stage algorithm (15.2)	128	544	1 006
Hwang algorithm (15.3)	187	815	1 197
One-stage Algorithm (15.6)	1 801	35 702	131 402

Table 15.1 Approximations for expected number of tests in various procedures for $n = 500\,000$

deposited into the daughterplate. The mixtures are then tested for activity. At the second stage, if the mixture is active then it is deemed to be a hit: the m components that constitute the hit are then tested individually to determine their activity.

We shall follow the binomial group testing model and assume that, prior to the experiment, the probability that a component is active is p. In practice, p is very small with a typical value being $p = 0.0001$ (this would correspond to $d = 50$ and $n = 500\,000$). We also assume that the activity of every component is independent of every other component.

We have:

Pr(a component is inactive) = $1 - p$;

Pr(a group of components is inactive) =$(1 - p)^m$.

Thus,

Pr(a group of components is active) = $1 - (1 - p)^m = P_{m,p}$.

Hence, the first stage can be modelled by a sequence of $k = n/m$ Bernoulli trials, with the probability of success being $P_{m,p}$. The number of successes (that is the number of active groups) is k', which is a random variable with a binomial distribution. The normalised cost is then

$$\widetilde{C}(m, p, \lambda) = \frac{1}{m}c_m + \frac{k'm}{n} + \lambda.$$

In practice p is small, therefore, by Taylor's expansion, $P_{m,p} \sim mp$ $(p \to 0)$ and

$$E[\tilde{C}(m, p, \lambda)] \sim \frac{1}{m}c_m + pm + \lambda\,;$$

$$Var[\tilde{C}(m, p, \lambda)] \sim \frac{Var(k')m^2}{n^2} = \frac{mP_{m,p}(1 - P_{m,p})}{n} = \frac{m^2 p}{n}.$$

This implies that when p is small and $n \to \infty$, the expected cost tends to infinity, but the variance of the cost remains bounded.

The optimum value of m for minimising the total cost may be found by numerical optimisation. For the case where the cost function is $c_s = 1 + \kappa s$, including the case $\kappa = 0$,

$$\frac{d\tilde{C}(m,p,\lambda)}{dm} = -\frac{1}{m^2} + p = 0; \quad \text{thus} \quad m^2 = \frac{1}{p} \quad \text{which gives} \quad m_{opt} = \sqrt{\frac{1}{p}}.$$

Figure 15.1 (a) shows the mean number of tests required to detect the active components for the optimum two-stage procedure, Figure 15.1 (b) shows the optimum value of m required to minimise the number of tests.

3. THREE-STAGE PROCEDURE

The three-stage procedure (again a particular case of Li's family of algorithms) has the same first stage as the two-stage procedure. However, with the three-stage procedure, components from the active mixture are analysed in groups of size l rather than individually to detect activity. On the third stage, the groups that were active on the second stage are tested individually for activity.

We again adopt the binomial group testing model. The cost of determining the number of active components for the three-stage procedure can be calculated as

$$C(n,m,l,p,\Lambda) = \frac{n}{m}c_m + \frac{k'm}{l}c_l + lk'' + 2\Lambda, \tag{15.8}$$

where $k' \sim Bin(k, P_{m,p})$ and $k'' \sim Bin(k', P_{l,p'})$.

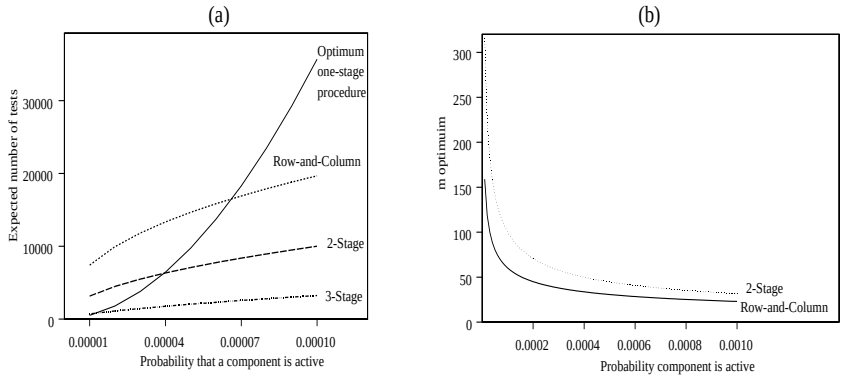


Figure 15.1 (a) the mean number of tests as a function of p for the one-stage, two-stage, three-stage (section 3) and row-and-column (section 4) procedures with optimum parameters; (b) the optimum values of m minimising the mean number of tests for the 2-stage and the row-and-column procedures

The first term in (15.8) counts the number of tests in the first stage which is $k = \frac{n}{m}$. At the second stage we have $n' = k'm$ components and we test these components in groups of l items. This gives $n'/l = k'm/l$ tests at the third stage. As a result of the second stage, we have k'' active groups each of size l, where analogously to the results above, $k'' \sim Bin(k', P_l, p')$, with p' being the posterior probability of an individual component being active. Since p is small, the probability of two or more active components in a group of m items in the first stage is negligible. Thus, we may assume that $p' = \frac{1}{m}$.

Hence, when p is small, the expected normalised cost is

$$E[\widetilde{C}(m, l, p, \lambda)] \sim \frac{1}{m}c_m + \frac{pm}{l}c_l + lp(1 - (1 - \frac{1}{l})^m) + 2\lambda, \quad p \to 0.$$

The optimum values of m and l for minimising the number of tests to find active components can be found by means of numerical optimisation.

Figure 15.1 (a) shows the mean number of tests required to detect the active components for the optimum three-stage procedure, Figure 15.2 (a) shows the optimum values of m and l required to minimise the number of tests.

4. ROW-AND-COLUMN PROCEDURE

For the row-and-column procedure the motherplate consists of $m > 1$ columns and $k > 1$ rows, giving in total km cells, each cell containing a different component. Without loss of generality, we assume that $m \leq k$. The number of motherplates to be tested is $r = \frac{n}{km}$, for simplicity we assume that r is an integer (in a typical experiment r is large). At the first stage, a mixed sample of the m components in each row is taken, along with mixed samples of the k components in each column, both of which are deposited in the daughterplate. The mixtures are tested for activity. At the first stage we thus make $\frac{n}{k} + \frac{n}{m}$ tests in total. The number of active components in each motherplate is ξ, which is a random variable with a binomial distribution, $\xi \sim Bin(km, p)$.

At the second stage we test the components that could be active for their activity (these components are located at the intersections of the active rows and columns). If there is either zero or one active component in the motherplate then no further tests are required at the second stage. However, if the number of active components in the motherboard ξ is larger than 1, then we must at most test all the intersections of the active rows and columns to detect the active components. If the active components are in different columns and rows, at most ξ^2 further tests will be required (if the active components are in the same row or column then the number of tests is smaller, since the number of intersection points to test for the active components will be less). Also, when there are $\xi \geq k$ active components then at most mk (the full motherplate) tests will be required. This implies that the upper bound for the expected number of tests required to determine the number of active components at the second

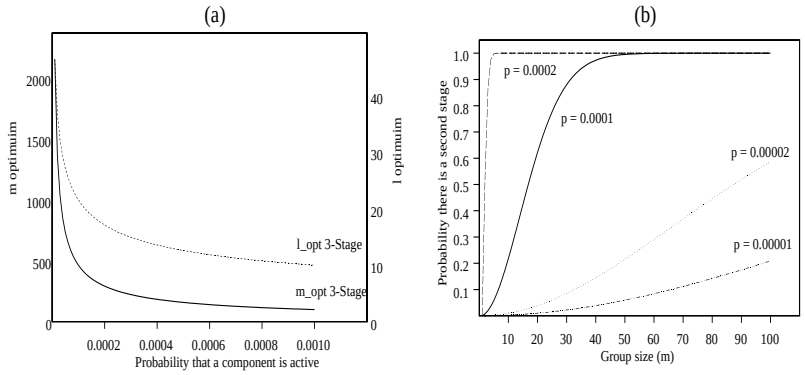


Figure 15.2 (a) the optimum values of m and l minimising the mean number of tests for the 3-stage procedure; (b) the probability that there is a second stage in the row-and-column procedure, as a function of m

stage may be estimated as

$$N(n,m,k,p) \leq \frac{n}{mk}(p_0 \cdot 0 + p_1 \cdot 0 + p_2 \cdot 2^2 + ... + p_m \cdot m^2 + (p_{m+1} + ... + p_{km})) \cdot km). \qquad (15.9)$$

If the value of p is small, the expected number of tests at the second stage is

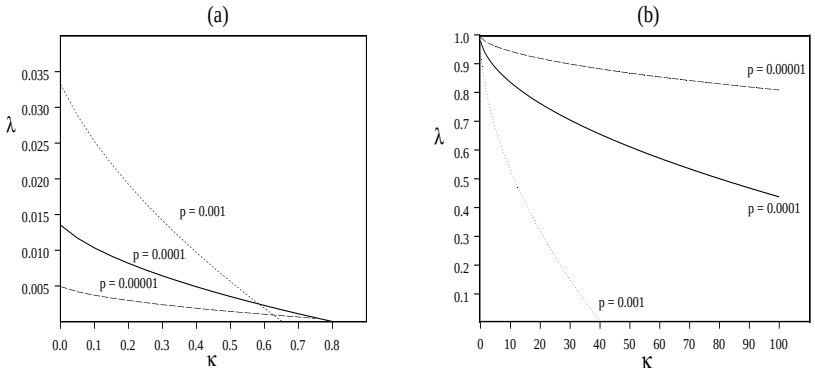


Figure 15.3 (a) the values of λ as a function of κ such that the 2-stage procedure has the same cost as the 3-stage procedure with $c_s = 1 + \kappa s$; (b) the values of λ as a function of κ such that the 1-Stage procedure has the same cost as 2-stage procedure with $c_s = 1 + \kappa \log s$

estimated as

$$E[N(n,m,k,p)] = \frac{n}{mk}\left(4\binom{km}{2}p^2(1-p)^{km-2} + (1-p_0-p_1-p_2)km\right)$$

since $(p_3 3^2 + ... + p_m m^2 + (p_{m+1} + ... + p_{km})km \leq (1 - p_0 - p_1 - p_2)km$. Therefore, the expected normalised cost of the row-and-column procedure may be estimated as

$$E[\tilde{C}(n,m,k,p)] \leq c_m \frac{1}{m} + c_k \frac{1}{k} + \frac{1}{mk}(2km(km-1)p^2(1-p)^{km-2} + (1 - p_0 - p_1 - p_2)km) + \lambda. \quad (15.10)$$

It is easy to estimate the probability that there will be a second stage for the row-and-column procedure. We do not need the second stage if there is never more than one active component in the motherplate. There are $r = n/(mk)$ motherplates and the probability of having two or more active components on each motherplate is

$$Q = 1 - p_0 - p_1 = 1 - (1-p)^{km} - kmp(1-p)^{km-1}.$$

Therefore, the probability that there is a second stage is $1 - (1-Q)^r$. These probabilities, for the optimum case $k = m$, are plotted in Figure 15.2 (b). We see from this plot that, for practical values of k and m, the probability that the row-and-column procedure is a one-stage procedure is large for small values of p. We assume that $k = m$ because, for constant mk, the expected number of tests is always smaller when $m < k$ than when $m = k$.

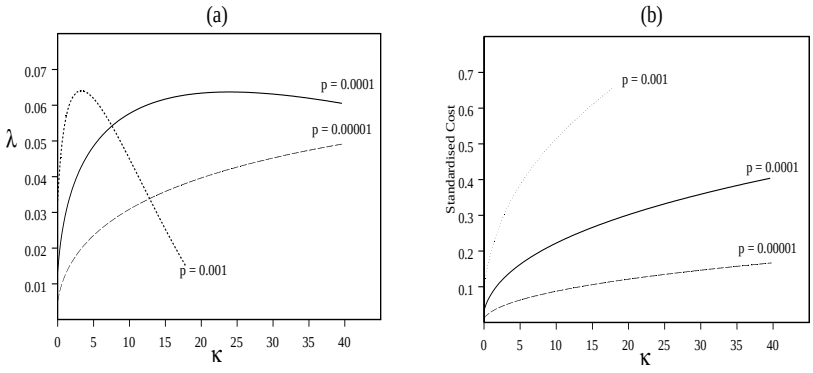


Figure 15.4 (a) the values of λ as a function of κ such that 2-stage procedure has the same cost as 3-stage procedure with $c_s = 1 + \kappa \log s$; (b) the standardised cost as a function of κ such that 2-stage procedure has the same cost as 3-stage procedure with $c_s = 1 + \kappa \log s$

5. CONCLUSIONS

We have shown that if the number of active components, d, is reasonably small (say, $d \leq 10$), then the optimum one-stage procedure can be considerably more cost-effective than the best multi-stage procedures.

By increasing the number of components in the mixtures in the two-stage procedure, we can significantly reduce the number of tests required to detect active components (the number of tests could be reduced by a factor of approximately 4 in the standard case, that is $p = 0.0001$).

The reduction in the number of tests can be even greater if we apply a multi-stage procedure. However, multi-stage procedures with three or more stages, are much less effective if we take into account the costs associated with the number of components in a mixture, and the penalties for extra stages. The two-stage procedure is often a good compromise.

In terms of the number of tests involved, the row-and-column method is typically worse than the two-stage standard with the same parameters. However, the number of components to be tested at the second stage for the row-and-column procedure is much smaller than in the two-stage and three-stage procedures. In Section 4 we found the probability that the row-and-column procedure is a one-stage procedure.

References

Dorfman, R. (1943). The detection of defective members of large populations. *Ann. Math. Statist.* **14**, 436–440.

Du, D. Z. and Hwang, F. K. (1993). *Combinatorial Group Testing*. Singapore: World Scientific.

Dyachkov, A.G., Rykov, V.V and Rashad, A.M. (1989). Superimposed distance codes. *Problems Control Inform. Th.* **18**, 237–250.

Feller, W. (1960). *An Introduction to Probability Theory and its Applications, Vol. 1 (2nd edn)*. New York: Wiley.

Hwang, F.K. (1972). A method for detecting all defective members in a population by group testing. *J. Amer. Statist. Assoc.* **67**, 605–608.

Li, C.H. (1962). A sequential method for screening experimental variables. *J. Amer. Statist. Assoc.* **57**, 455–477.

Patel, M. (1962). Group screening with more than two stages. *Technometrics* **4**, 209–217.

Renyi, A. (1965). On the theory of random search. *Bull. Amer. Math. Soc.* **71**, 809–828.

Sobel, M. and Groll, P.A. (1959). Group testing to eliminate efficiently all defectives in a binomial sample. *Bell System Tech. J.* **38**, 1179–1252.

Zhigljavsky, A.A. (1998). Asymptotic upper bounds for the optimal design length in factor screening experiments. In *MODA 5 - Advances in Model-Oriented Data Analysis and Experimental Design* Eds A.C. Atkinson, L. Pronzato and H.P. Wynn, pp. 85–93. Heidelberg; Physica Verlag.

Chapter 16

BLOCK DESIGNS FOR COMPARISON OF TWO TEST TREATMENTS WITH A CONTROL

Steven M. Bortnick
Battelle Memorial Institute Columbus, Ohio
bortnick@battelle.org

Angela M. Dean and Jason C. Hsu
The Ohio State University Columbus, Ohio
amd@stat.ohio-state.edu and jch@stat.ohio-state.edu

Abstract In an experiment to compare $p = 2$ test treatments with a control, simultaneous confidence bounds (or intervals) for the amount by which each test treatment is better than (or differs from) the control are required. When an experiment is arranged in b blocks of size k, the optimal allocation of a fixed number of experimental units to the individual test treatments and the control within each block need to be determined. The optimality criteria of interest are the minimization of the expected average allowance (EAA) or the minimization of the expected maximum allowance (EMA) of the simultaneous confidence bounds for the $p = 2$ treatment-control mean contrasts. This paper provides bounds on the EAA and EMA which are then used in search algorithms for obtaining optimal or highly efficient experimental design solutions.

Keywords: block design, confidence bounds, multiple comparisons, optimal design, treatment versus control

1. INTRODUCTION

Multiple comparisons of test treatments with a control are common in a wide range of disciplines, such as agriculture, medicine, biopharamceutics, engineering and environmental studies. For example, in environmental pollution monitoring, the toxicity of whole effluents is assessed by comparing increasing concentrations of an effluent with distilled water in terms of growth, reproductive, and survival rates of various species, (cf. U.S. EPA, 1994). Hsu and Berger

(1999) cite several examples in dose-response studies in which the treatment comparisons are with respect to a control. In this paper, we consider treatment versus control comparisons in a block design setting. For example, John (1961) discussed an experiment involving nine detergents, where one detergent was the standard detergent, and only three treatments could be observed under similar experimental conditions. Daly and Cooper (1967) described a study for reducing stuttering which compared two electro-shock treatments with a control treatment, where the subjects formed blocks.

Concern for controlling the experiment-wise error rate in such studies leads to the use of simultaneous inference in the data analysis phase of the experiment, which results in a set of simultaneous confidence intervals or bounds for the treatment versus control contrasts. Optimal designs for this situation are those that minimize either the expected average half-width or the expected maximum half-width of simultaneous confidence intervals for the treatment versus control contrasts. In the case of one-sided upper or lower confidence bounds, the optimal design minimizes the average or maximum amount above or below the contrast estimates that is needed to obtain a specified level of confidence. Spurrier and Nizam (1990) labelled the first criterion as the expected average allowance, where an allowance represents the "plus and/or minus" one attaches to point estimates (see Tukey, 1994). We call the second criterion the expected maximum allowance. We use the abbreviations EAA and EMA to refer to these criteria.

The proposed EAA and EMA criteria should be preferred over the traditional criteria of A- and MV-optimality, since the latter ignore the correlations among the estimated treatment versus control contrasts. Using Dunnett's method of multiple comparisons (Dunnett, 1955), the correlations determine the simultaneous critical value associated with the set of confidence intervals (bounds) and this plays an integral role in determining the size of the allowances for the intervals (bounds).

Spurrier and Nizam (1990) studied the EAA criterion in the completely randomized design setting with $p = 2$ test treatments and a control treatment. They showed that, for a fixed number of replications n_0 of the control treatment, the design most balanced with respect to the number of test treatment replications minimizes both the average standard deviation and Dunnett's simultaneous critical values. Therefore, the EAA-optimal completely randomized design with a fixed number of control replications is unequivocally the design most balanced in the number of test treatment replications. The same is true under EMA-optimality, although Spurrier and Nizam (1990) did not consider this criterion. The situation is more complicated, however, in a block design setting as discussed below.

Let d be a design in the class of block designs $D\{2, b, k\}$ with $p = 2$ test treatments, one control treatment, and b blocks of size k. Let $\lambda_{st}^{(d)} = \sum_h n_{sh} n_{th}$,

where n_{sh} and n_{th} are, repectively, the number of times that treatments s and t occur in block h; ($h = 1, \ldots, b$; $s \neq t = 0, 1, 2$, where treatment 0 is the control treatment). It is straightforward to show that the variance-covariance matrix for the two treatment versus control contrast estimators is

$$\left(\frac{k\,\sigma^2}{\left(\lambda_{01}^{(d)} + \lambda_{12}^{(d)}\right)\left(\lambda_{02}^{(d)} + \lambda_{12}^{(d)}\right) - \left(\lambda_{12}^{(d)}\right)^2} \right) \times \begin{bmatrix} \lambda_{02}^{(d)} + \lambda_{12}^{(d)} & \lambda_{12}^{(d)} \\ \lambda_{12}^{(d)} & \lambda_{01}^{(d)} + \lambda_{12}^{(d)} \end{bmatrix}. \tag{16.1}$$

Let $\eta_i^{(d)}$ represent the standard deviation of the least squares estimator of the treatment versus control contrast $\tau_0^{(d)} - \tau_i^{(d)}$, $i = 1, \ldots, p$ and let $c_j^{(d)}(\alpha)$ be Dunnett's j-sided ($j = 1$ or 2) simultaneous critical value. Using (16.1) and the notation of $r_1 = \lambda_{01} + \lambda_{12}$ and $r_2 = \lambda_{02} + \lambda_{12}$, the average standard deviation of the treatment versus control contrast estimators can be written as

$$\bar{\eta} = \tfrac{\eta_1+\eta_2}{2} = \tfrac{\sqrt{k}\,\sigma}{2}\left[\left(\tfrac{r_1}{r_1 r_2 - \lambda_{12}^2}\right)^{1/2} + \left(\tfrac{r_2}{r_1 r_2 - \lambda_{12}^2}\right)^{1/2}\right], \tag{16.2}$$

and the maximum standard deviation can be written as

$$\breve{\eta} = \tfrac{(\eta_1+\eta_2)+|\eta_1-\eta_2|}{2} = \bar{\eta} + \tfrac{\sqrt{k}\,\sigma}{2}\left|\tfrac{\sqrt{r_1}-\sqrt{r_2}}{\sqrt{r_1 r_2 - \lambda_{12}^2}}\right|. \tag{16.3}$$

Let $\mathbf{n}_0 = [n_{01}, \ldots, n_{0b}]$ denote the allocations of the control treatments to the blocks, and let $D_{(\mathbf{n}_0,\lambda_{12})}\{2, b, k\}$ represent the subset of designs with fixed control layout $\mathbf{n}_0$ and fixed parameter λ_{12}. Theorems 4.1, 4.2 and 4.3 of Bortnick (1999) are stated below in Theorem 1 for future reference. They imply that, among designs within $D_{(\mathbf{n}_0,\lambda_{12})}\{2, b, k\}$, both the average standard deviation $\bar{\eta}$ and the maximum standard deviation $\breve{\eta}$ are minimized by the design(s) *most balanced* with respect to the parameters λ_{01} and λ_{02}, but Dunnett's simultaneous critical values are minimized by the design(s) *least balanced* with respect to these parameters.

Theorem 1. *a) Within the subclass $D_{(\mathbf{n}_0,\lambda_{12})}\{2, b, k\}$, the average standard deviation and the maximum standard deviation of the treatment versus control contrast estimators are both increasing in the difference $|\lambda_{01} - \lambda_{02}|$.*

b) Within the subclass $D_{(\mathbf{n}_0,\lambda_{12})}\{2, b, k\}$, Dunnett's simultaneous critical values $c_j(\alpha)$, $j = 1, 2$, are decreasing in the difference $|\lambda_{01} - \lambda_{02}|$.

Full details of the proof are given in Bortnick (1999). Part (a) is proved by showing that the average standard deviation $\bar{\eta}$ is a Schur-convex function of r_1 and r_2. Set $\mathbf{r}^{(d_a)} = \left[r_1^{(d_a)}, r_2^{(d_a)}\right]^T$ and $\mathbf{r}^{(d_b)} = \left[r_1^{(d_b)}, r_2^{(d_b)}\right]^T$. Design d_b has the smaller average standard deviation if $\mathbf{r}^{(d_a)}$ for design d_a majorizes $\mathbf{r}^{(d_b)}$

for design d_b. Then, since

$$\mathbf{r}^{(d_a)} \succ \mathbf{r}^{(d_b)} \quad \Leftrightarrow \quad \left|\lambda_{01}^{(d_a)} - \lambda_{02}^{(d_a)}\right| \geq \left|\lambda_{01}^{(d_b)} - \lambda_{02}^{(d_b)}\right|,$$

the result for the average standard deviation follows.

Also, from (16.3) with λ_{12} fixed, the maximum standard deviation is given by

$$\check{\eta} \;=\; \bar{\eta} + |h(r_1, r_2)| \;=\; \bar{\eta} + \sqrt{h^2(r_1, r_2)}\,.$$

It can be shown that $h^2(r_1, r_2)$ is decreasing in $r_1 r_2$. Also $r_1 r_2$ is decreasing in $|\lambda_{01} - \lambda_{02}|$ and, combined with the result for the average standard deviation, the result for the maximum standard deviation follows.

Part (b) of the theorem is proved by showing that the correlation ρ between the treatment versus control estimators is a Schur-convex function of r_1 and r_2. From Slepian's inequality and Šidák's inequality, respectively, (see Tong, 1980), $c_1(\alpha)$ and $c_2(\alpha)$ are decreasing in the correlation ρ for any specified α. Therefore, from Table 1 of Marshall and Olkin (1979, p.61), it follows that $c_1(\alpha)$ and $c_2(\alpha)$ are Schur-concave functions of r_1 and r_2. The result then follows by majorization.

Since the optimality criteria of interest are based on products of a standard deviation and a critical value, the optimal designs are not obvious. Block size $k = 2$ is discussed in Section 2, where a computer search strategy and tables of optimal designs are given. The settings $k \geq 3$ are discussed briefly in Section 3

2. THE INCOMPLETE BLOCK SETTING

2.1. BOUNDING THE STANDARD DEVIATIONS

When all of the blocks are of size $k = 2$, non-binary designs necessarily contain blocks with two replications of a single treatment. Since such blocks do not give information on the contrasts of interest, we do not consider non-binary designs. For binary designs, if n_i is the number of replications of treatment i, we have

$$\begin{array}{rclcrcl} \lambda_{12} & = & b - n_0\,, & & \lambda_{01} + \lambda_{02} & = & n_0, \\ \lambda_{01} + \lambda_{12} & = & n_1 & \text{and} & \lambda_{02} + \lambda_{12} & = & n_2. \end{array} \tag{16.4}$$

In a binary design with $k = 2$, the control layout is determined by the number of control replications n_0. Since $\lambda_{12} = (b - n_0)$ is also determined by n_0, we may write $D_{(\mathbf{n}_0, \lambda_{12})}\{2, b, 2\}$ as $D_{(n_0)}\{2, b, 2\}$.

Since $D_{(n_0)}\{2, b, 2\}$ contains only binary designs, lower and upper bounds on n_1 and n_2 are given by $(b - n_0)$ and b. Therefore, from (16.4), the difference $|\lambda_{01} - \lambda_{02}|$ is maximized in $D_{(n_0)}\{2, b, 2\}$ when one of n_1 or n_2 is equal to $(b - n_0)$ and the other is equal to b. Using this fact together with Theorem 1a

and (16.2), an upper bound on $\bar{\eta}$ is given by

$$\bar{\eta} \;\leq\; \tfrac{\sqrt{2}\sigma}{2}\left(\frac{\sqrt{(b-n_0)}+\sqrt{b}}{\sqrt{(b-n_0)n_0}}\right). \tag{16.5}$$

Likewise, Theorem 1b, (16.3) and 16.5 imply, for all designs in $D_{(n_0)}\{2,b,2\}$, an upper bound on $\check{\eta}$ of

$$\check{\eta} \;\leq\; \sqrt{2}\sigma\left(\frac{\sqrt{b}}{\sqrt{(b-n_0)n_0}}\right). \tag{16.6}$$

From (16.4), we have $n_1+n_2=(2b-n_0)$. Therefore, the difference $|\lambda_{01}-\lambda_{02}|$ is minimized in $D_{(n_0)}\{2,b,2\}$ when $n_1=n_2=(2b-n_0)/2$. In this case, $r_1=r_2$ in (16.3) and $\bar{\eta}$ and $\check{\eta}$ are equal. For all designs in $D_{(n_0)}\{2,b,2\}$, a lower bound on $\bar{\eta}$ and $\check{\eta}$ is, therefore, given by

$$\bar{\eta} \text{ or } \check{\eta} \;\geq\; \sigma\left(\frac{\sqrt{2b-n_0}}{\sqrt{(b-\frac{3}{4}n_0)n_0}}\right). \tag{16.7}$$

Since the designs being compared are all binary in the treatments and $k=2$, it follows that n_1 and n_2 must differ by at least one when n_0 is an odd number and bounds slightly sharper than those in (16.7) can be obtained, as follows:

$$\bar{\eta} \;\geq\; \tfrac{\sqrt{2}\sigma}{2}\left(\frac{\sqrt{\left[\frac{2b-n_0}{2}\right]}+\sqrt{\left[\frac{2b-n_0}{2}\right]+1}}{\sqrt{\left[\frac{2b-n_0}{2}\right]\left(\left[\frac{2b-n_0}{2}\right]+1\right)-(b-n_0)^2}}\right) \tag{16.8}$$

and

$$\check{\eta} \;\geq\; \sqrt{2}\sigma\left(\frac{\sqrt{\left[\frac{2b-n_0}{2}\right]+1}}{\sqrt{\left[\frac{2b-n_0}{2}\right]\left(\left[\frac{2b-n_0}{2}\right]+1\right)-(b-n_0)^2}}\right), \tag{16.9}$$

where $[\cdot]$ represents the greatest integer function.

2.2. BOUNDING THE CRITICAL VALUE

Using (16.1), the correlation between the treatment versus control contrast estimators can be written as

$$\rho=\frac{\lambda_{12}}{r_1^{1/2}\;r_2^{1/2}}\,, \tag{16.10}$$

where $r_1=\lambda_{01}+\lambda_{12}$ and $r_2=\lambda_{02}+\lambda_{12}$.

As noted above, the difference $|\lambda_{01}-\lambda_{02}|$ is minimized in $D_{(n_0)}\{2,b,2\}$ when $n_1=n_2=(2b-n_0)/2$. Using this fact together with Theorem 1b, (16.10) and (16.4), an upper bound on ρ is given by

$$\rho \;\leq\; \frac{(b-n_0)}{\sqrt{(b-n_0)b}} \;=\; \sqrt{\tfrac{b-n_0}{b}}\,, \tag{16.11}$$

and a lower bound on ρ is given by

$$\rho \geq \frac{2b-2n_0}{2b-n_0} \quad \text{or} \quad \frac{b-n_0}{\sqrt{\left[\frac{2b-n_0}{2}\right]\left(\left[\frac{2b-n_0}{2}\right]+1\right)}}, \tag{16.12}$$

for n_0 even or odd, respectively.

Since $c_1(\alpha)$ and $c_2(\alpha)$ are decreasing in the correlation ρ, (16.11) can be used in the numerical calculation of a lower bound on $c_j(\alpha)$. Similarly, (16.12) can be used to calculate an upper bound on $c_j(\alpha)$ for n_0 even or odd.

2.3. OBTAINING EAA- AND EMA-OPTIMAL DESIGNS

For a specified $0 < \alpha < 1$ and $j = 1$ or 2, we define $\text{EAA}_j(\alpha) = \bar{\eta} \times c_j(\alpha)$, which is to be minimized under the EAA-optimality criterion, and $\text{EMA}_j(\alpha) = \check{\eta} \times c_j(\alpha)$, which is to be minimized under the EMA-optimality criterion. Since we are interested in the comparison of $\text{EAA}_j(\alpha)$ and $\text{EMA}_j(\alpha)$ among designs within $D\{2, b, 2\}$, all reported values of $\bar{\eta} \times c_j(\alpha)$ and $\check{\eta} \times c_j(\alpha)$ are given, without loss of generality, with $E(S) = 1.0$.

Within each subclass $D_{(n_0)}\{2, b, 2\}$, smaller $\bar{\eta}$ and $\check{\eta}$ are obtained by designs more balanced in the test treatment replications (i.e. which have a smaller $|\lambda_{01} - \lambda_{02}|$); however, smaller $c_j(\alpha)$ are obtained by designs which are less balanced in the test treatment replications. Although it is not obvious which designs within $D_{(n_0)}\{2, b, 2\}$ are most efficient with respect to the EAA-optimality and EMA-optimality criteria, an efficient search still can be made through the use of the standard deviation and correlation bounds provided in the previous subsection.

We first observe that the upper and lower bounds on the correlation between the treatment versus control contrast estimators are fairly close to one another. As an example, consider the subclass $D_{(n_0)}\{2, b, 2\}$ with $b = 20$ and $n_0 = 18$. Using (16.11) and (16.12), the bounds for ρ are $0.182 \leq \rho \leq 0.316$. This example was chosen because an enumeration of $1 \leq n_0 \leq 20$ showed this range to be the widest (i.e. worst case scenario) among all n_0 when $b = 20$. The bounds for ρ give rise to the bounds $2.066 \leq c_1(0.05) \leq 2.079$ for Dunnett's one-sided critical value at $\alpha = 0.05$. This is a very narrow range and so even the most balanced designs within $D_{(18)}\{2, b, 2\}$ produce critical values that are close to the minimum for that subclass.

On the other hand, (16.5), (16.6) and (16.7) give bounds on $\bar{\eta}$ and $\check{\eta}$ of $0.307 \leq \bar{\eta} \leq 0.491$ and $0.307 \leq \check{\eta} \leq 0.745$ for designs in $D_{(18)}\{2, 20, 2\}$. Thus, the critical values appear almost constant when compared with the range in the average and maximum standard deviations and it seems likely that the designs most balanced in the number of test treatment replications will produce $\text{EAA}_j(\alpha)$ and $\text{EMA}_j(\alpha)$ values close to the subclass lower bound. This is indeed the case in $D_{(18)}\{2, 20, 2\}$, since the design with $\lambda_{01} = \lambda_{02}$ yields a

value of 0.638 for both $\text{EAA}_1(0.05)$ and $\text{EMA}_1(0.05)$, which is close to the bound 0.634. The efficiency of this design is $(0.634/0.638)^2 = 0.988$.

In the design space $D\{2, 20, 2\}$, we have shown via an exhaustive set of calculations that the most balanced design always minimizes both $\text{EAA}_j(\alpha)$ and $\text{EMA}_j(\alpha)$ at each n_0. Roughly speaking, $\text{EAA}_j(\alpha)$ and $\text{EMA}_j(\alpha)$ act like convex functions over their domain, although each is made up of the product of a convex function and a concave function.

The $\text{EAA}_j(\alpha)$ and $\text{EMA}_j(\alpha)$ lower bounds can be used to eliminate the majority of subclasses $D_{(n_0)}\{2, b, 2\}$ corresponding to poor choices for n_0 within $D\{2, b, 2\}$. Consider two competing subclasses, say $D_{(n_{0a})}\{2, b, 2\}$ and $D_{(n_{0b})}\{2, b, 2\}$, corresponding to n_{0a} and n_{0b} replications of the control, respectively. Every design in $D_{(n_{0a})}\{2, b, 2\}$ immediately can be eliminated from consideration for minimizing $\text{EAA}_j(\alpha)$ if there exists some design within $D_{(n_{0b})}\{2, b, 2\}$, say d_b, such that d_b produces a value for $c_j(\alpha) \times \bar{\eta}$ at or below the $\text{EAA}_j(\alpha)$ lower bound for $D_{(n_{0a})}\{2, b, 2\}$. This same logic holds true under the EMA-optimality criterion. As a result, any design with n_{0a} replications of the control need not be considered further in the search for an optimal design.

A logical choice for d_b is the design within subclass $D_{(n_0 b)}\{2, b, 2\}$ that is most balanced with respect to the number of test treatment replications (i.e. the design that minimizes $|\lambda_{01} - \lambda_{02}|$). This gives the following algorithm for obtaining the EAA-optimal design:

(i) Select α and j, and calculate the lower bound on $\text{EAA}_j(\alpha)$ in each subclass $D_{(n_0)}\{2, b, 2\}$ ($0 \leq n_0 \leq b$);

(ii) Calculate the $\text{EAA}_j(\alpha)$ for the most-balanced design in each subclass $D_{(n_0)}\{2, b, 2\}$ ($0 \leq n_0 \leq b$);

(iii) Let d_b be the design yielding the minimum $\text{EAA}_j(\alpha)$ calculated in step 2;

(iv) Any subclass $D_{(n_0)}\{2, b, 2\}$ ($1 \leq n_0 \leq b$) with a lower bound (calculated in step 1) greater than or equal to the value of $\text{EAA}_j(\alpha)$ for d_b can be eliminated from consideration;

(v) Search for an optimal design among the reduced set of subclasses still in contention.

The above algorithm applies to a search for EMA-optimal designs by replacing references to EAA by EMA.

As an example, for the design space $D\{2, 20, 2\}$, the top half of Figure 16.1 shows $\text{EAA}_1(0.05)$ for the most balanced designs and the $\text{EAA}_1(0.05)$ lower bound at various n_0. The horizontal dashed line is drawn at the minimum $\text{EAA}_1(0.05)$ among the set of most-balanced designs. This line indicates that only designs with 15, 16 or 17 replications of the control need be considered any

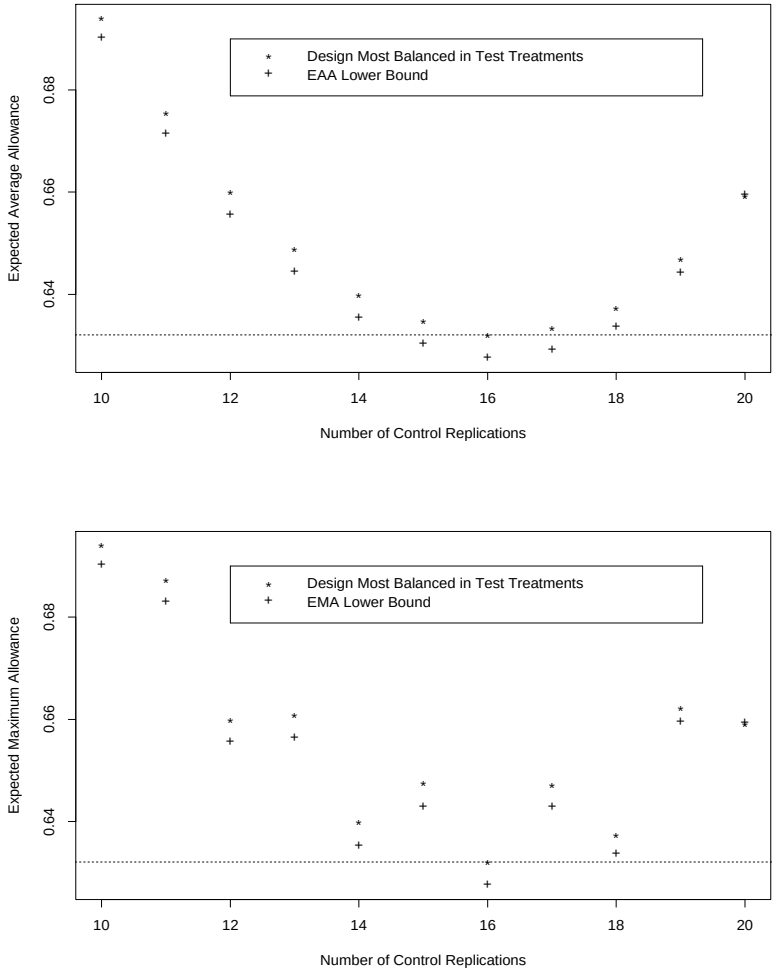


Figure 16.1 Use of the lower bounds on $\text{EAA}_1(0.05)$ and $\text{EMA}_1(0.05)$ to eliminate designs with an inefficient number of control replications

further in the search for an EAA-optimal solution. The results in the bottom half of Figure 16.1 for $\text{EMA}_1(0.05)$ indicate that only designs with 16 replications of the control are in contention to minimize $\text{EMA}_1(0.05)$. As it turns out, the most-balanced design corresponding to $n_0 = 16$ minimizes $\text{EMA}_1(0.05)$ within $D\{2, 20, 2\}$.

Tables 16.1 and 16.2 contain the designs that minimize the EAA and EMA, respectively, for one-sided and two-sided confidence bounds when $p = 2$.

Optimal designs for both criteria are given for $3 \leq b \leq 50$ and for α = 0.01, 0.05 and 0.10. The designs were obtained using the above algorithm. To appreciate the size of the search and the efficiency of the algorithm, observe that with block size equal to b there are $(b+5)!/(b!5!)$ possible design configurations to consider (see Bortnick, 1999). However, the search algorithm involves only one calculation for each of the $b-1$ subclasses $D_{(n_0)}\{2, b, 2\}$, followed by an elementary search for the optimal design within the best subclass. With standard statistical software running on a Windows-based computer, identifying an optimal design via complete enumeration requires approximately 20 hours with 50 blocks, 7 hours with 40 blocks, 2 hours with 30 blocks and 20 minutes with 20 blocks, whereas the search algorithm requires only seconds to complete. More substantial time savings are realized for larger block sizes (i.e., $k > 2$).

In order to construct a specific design from Table 16.1 or Table 16.2, note that (i) every optimal solution in each table is binary in all treatments and (ii) every solution corresponds to the design most balanced with respect to the number of test treatment replications at the given value of n_0. For example, consider the one-sided $\alpha = 0.10$ case when $b = 10$. Under this scenario, the same design minimizes both the EAA and the EMA, namely $n_0 = 8$. There are a total of 12 remaining plots for allocation to the $p = 2$ test treatments. Since the optimal solution corresponds to a binary design most balanced with respect to the number of test treatment replications; that is $n_1 = n_2 = 6$, the design is as follows (where columns are blocks):

0	0	0	0	0	0	0	0	1	1
1	1	1	1	2	2	2	2	2	2

For those cases when the optimal solution corresponds to an odd number for n_0, the design with $n_1 = n_2 - 1$ is equivalent to the design with $n_1 = n_2 + 1$.

Interestingly, every EMA-optimal design in Table 16.2 has an even number of control replications. Since these designs are also most balanced with respect to the number of test treatment replications, each listed EMA-optimal design is an R-type or S-type BTIB design (as defined by Hedeyat et al., 1988). The same is true for EAA-optimal designs in Table 16.1 that have an even number of control replications n_0 and, in fact, in these cases, the EMA- and EAA-optimal designs are the same. When n_0 is odd, for the EAA-optimal binary designs, $\lambda_{01} \neq \lambda_{02}$ and no BTIB designs exist.

2.4. A COMPARISON WITH A- AND MV-OPTIMAL DESIGNS

To this date, the majority of research regarding the multiple comparisons with a control problem in the incomplete block setting has focused on the criterion of A-optimality. From a multiple comparisons standpoint, the goal

Number of Blocks	$(1-\alpha)\times 100\%$											
	One-Sided						Two-Sided					
	90%		95%		99%		90%		95%		99%	
	n_0	EMA	n_0	EMA	n_0	EMA	n_0	EMA	n_0	EMA	n_0	EMA
3	2	3.834	2	7.764	2	38.978	2	7.062	2	14.182	2	70.997
4	3	1.819	3	2.721	3	6.340	3	2.645	3	3.850	3	8.799
5	4	1.338	4	1.838	4	3.416	4	1.811	4	2.401	4	4.324
6	4	1.133	5	1.502	5	2.521	5	1.487	5	1.890	5	3.055
7	6	0.997	6	1.289	6	2.045	6	1.279	6	1.589	6	2.418
8	6	0.898	6	1.151	6	1.772	6	1.148	6	1.407	6	2.066
9	7	0.831	7	1.054	8	1.582	7	1.051	8	1.275	8	1.821
10	8	0.773	8	0.974	8	1.435	8	0.971	8	1.169	8	1.642
11	8	0.730	9	0.914	9	1.329	9	0.912	9	1.092	9	1.512
12	10	0.692	10	0.862	10	1.240	10	0.860	10	1.024	10	1.403
13	10	0.658	10	0.819	11	1.170	10	0.818	11	0.972	11	1.319
14	11	0.630	11	0.783	12	1.108	12	0.781	12	0.924	12	1.246
15	12	0.605	12	0.749	12	1.056	12	0.748	12	0.884	12	1.185
16	12	0.584	13	0.721	13	1.012	13	0.720	13	0.849	13	1.132
17	14	0.564	14	0.695	14	0.971	14	0.694	14	0.816	14	1.084
18	14	0.545	14	0.673	15	0.936	15	0.672	15	0.789	15	1.044
19	15	0.529	16	0.652	16	0.904	16	0.651	16	0.763	16	1.006
20	16	0.514	16	0.632	16	0.876	16	0.631	16	0.740	17	0.974
21	16	0.501	17	0.615	18	0.849	17	0.614	17	0.719	18	0.943
22	18	0.488	18	0.599	18	0.825	18	0.598	18	0.699	18	0.916
23	18	0.476	19	0.584	19	0.803	19	0.584	19	0.682	19	0.891
24	19	0.466	20	0.570	20	0.783	20	0.570	20	0.665	20	0.867
25	20	0.455	20	0.557	21	0.764	20	0.557	21	0.650	21	0.846
26	20	0.446	21	0.545	22	0.746	21	0.545	22	0.636	22	0.826
27	22	0.437	22	0.534	22	0.730	22	0.534	22	0.622	22	0.807
28	22	0.428	23	0.524	23	0.715	23	0.523	23	0.609	24	0.790
29	23	0.420	24	0.514	24	0.700	24	0.513	24	0.597	24	0.773
30	24	0.413	24	0.504	25	0.687	24	0.504	25	0.586	25	0.758
31	24	0.406	25	0.495	26	0.674	26	0.495	26	0.576	26	0.744
32	26	0.399	26	0.487	27	0.662	26	0.486	26	0.566	27	0.730
33	26	0.392	27	0.479	28	0.650	27	0.478	27	0.556	28	0.717
34	27	0.386	28	0.471	28	0.639	28	0.471	28	0.547	28	0.705
35	28	0.380	28	0.464	29	0.629	29	0.464	29	0.538	29	0.693
36	28	0.375	29	0.457	30	0.619	30	0.457	30	0.530	30	0.682
37	30	0.370	30	0.450	31	0.610	30	0.450	30	0.522	31	0.672
38	30	0.364	31	0.444	32	0.601	31	0.444	32	0.515	32	0.661
39	31	0.359	32	0.438	32	0.592	32	0.437	32	0.507	33	0.652
40	32	0.355	32	0.432	34	0.584	33	0.432	33	0.501	34	0.642
41	32	0.350	34	0.426	34	0.576	34	0.426	34	0.494	34	0.634
42	33	0.346	34	0.421	35	0.568	34	0.420	35	0.488	35	0.625
43	34	0.342	35	0.416	36	0.561	35	0.415	36	0.481	36	0.617
44	35	0.338	36	0.410	37	0.554	36	0.410	36	0.476	37	0.609
45	36	0.334	37	0.406	38	0.547	37	0.405	37	0.470	38	0.601
46	36	0.330	38	0.401	38	0.541	38	0.401	38	0.464	38	0.594
47	37	0.326	38	0.396	39	0.534	38	0.396	39	0.459	40	0.587
48	38	0.323	39	0.392	40	0.528	39	0.392	40	0.454	40	0.580
49	39	0.319	40	0.388	41	0.522	40	0.388	40	0.449	41	0.574
50	40	0.316	41	0.384	42	0.517	41	0.384	42	0.444	42	0.567

Table 16.1 EAA-optimal designs

Number of Blocks	$(1-\alpha)\times 100\%$											
	One-Sided						Two-Sided					
	90%		95%		99%		90%		95%		99%	
	n_0	EMA	n_0	EMA	n_0	EMA	n_0	EMA	n_0	EMA	n_0	EMA
3	2	3.834	2	7.764	2	38.978	2	7.062	2	14.182	2	70.997
4	2	1.885	2	2.839	4	6.620	4	2.706	4	3.935	4	8.990
5	4	1.338	4	1.838	4	3.416	4	1.811	4	2.401	4	4.324
6	4	1.133	4	1.507	4	2.545	4	1.500	4	1.911	4	3.096
7	6	0.997	6	1.289	6	2.045	6	1.279	6	1.589	6	2.418
8	6	0.898	6	1.151	6	1.772	6	1.148	6	1.407	6	2.066
9	8	0.838	8	1.059	8	1.582	8	1.053	8	1.275	8	1.821
10	8	0.773	8	0.974	8	1.435	8	0.971	8	1.169	8	1.642
11	8	0.730	8	0.918	10	1.337	8	0.917	10	1.098	10	1.519
12	10	0.692	10	0.862	10	1.240	10	0.860	10	1.024	10	1.403
13	10	0.658	10	0.819	10	1.172	10	0.818	10	0.973	10	1.322
14	12	0.633	12	0.783	12	1.108	12	0.781	12	0.924	12	1.246
15	12	0.605	12	0.749	12	1.056	12	0.748	12	0.884	12	1.185
16	12	0.584	12	0.723	14	1.013	14	0.722	14	0.850	14	1.133
17	14	0.564	14	0.695	14	0.971	14	0.694	14	0.816	14	1.084
18	14	0.545	14	0.673	14	0.938	14	0.672	14	0.790	16	1.046
19	16	0.531	16	0.652	16	0.904	16	0.651	16	0.763	16	1.006
20	16	0.514	16	0.632	16	0.876	16	0.631	16	0.740	16	0.974
21	16	0.501	18	0.616	18	0.849	18	0.615	18	0.719	18	0.943
22	18	0.488	18	0.599	18	0.825	18	0.598	18	0.699	18	0.916
23	18	0.476	18	0.584	20	0.804	18	0.584	20	0.683	20	0.891
24	18	0.466	20	0.570	20	0.783	20	0.570	20	0.665	20	0.867
25	20	0.455	20	0.557	20	0.765	20	0.557	20	0.650	20	0.847
26	20	0.446	22	0.546	22	0.746	22	0.545	22	0.636	22	0.826
27	22	0.437	22	0.534	22	0.730	22	0.534	22	0.622	22	0.807
28	22	0.428	22	0.524	24	0.715	22	0.524	24	0.610	24	0.790
29	22	0.421	24	0.514	24	0.700	24	0.513	24	0.597	24	0.773
30	24	0.413	24	0.504	24	0.687	24	0.504	24	0.587	26	0.759
31	24	0.406	26	0.495	26	0.674	26	0.495	26	0.576	26	0.744
32	26	0.399	26	0.487	26	0.662	26	0.486	26	0.566	26	0.730
33	26	0.392	26	0.479	28	0.650	28	0.479	28	0.556	28	0.717
34	26	0.387	28	0.471	28	0.639	28	0.471	28	0.547	28	0.705
35	28	0.380	28	0.464	30	0.629	28	0.464	28	0.539	30	0.693
36	28	0.375	30	0.457	30	0.619	30	0.457	30	0.530	30	0.682
37	30	0.370	30	0.450	30	0.610	30	0.450	30	0.522	30	0.672
38	30	0.364	30	0.444	32	0.601	32	0.444	32	0.515	32	0.661
39	30	0.360	32	0.438	32	0.592	32	0.437	32	0.507	32	0.652
40	32	0.355	32	0.432	34	0.584	32	0.432	34	0.501	34	0.642
41	32	0.350	34	0.426	34	0.576	34	0.426	34	0.494	34	0.634
42	34	0.346	34	0.421	36	0.568	34	0.420	34	0.488	36	0.625
43	34	0.342	36	0.416	36	0.561	36	0.415	36	0.481	36	0.617
44	34	0.338	36	0.410	36	0.554	36	0.410	36	0.476	36	0.609
45	36	0.334	36	0.406	38	0.547	36	0.406	38	0.470	38	0.601
46	36	0.330	38	0.401	38	0.541	38	0.401	38	0.464	38	0.594
47	38	0.326	38	0.396	40	0.534	38	0.396	38	0.459	40	0.587
48	38	0.323	40	0.392	40	0.528	40	0.392	40	0.454	40	0.580
49	38	0.319	40	0.388	40	0.522	40	0.388	40	0.449	42	0.574
50	40	0.316	40	0.384	42	0.517	40	0.384	42	0.444	42	0.567

Table 16.2 EMA-optimal designs

of minimizing the EAA appears more appropriate than simply minimizing the average variance. To compare these two approaches, we can find the A-optimal design d_1 at each value of $3 \leq b \leq 50$ and compare its Expected Average Allowance to that of the EAA-optimal design d_2. An analogous approach is taken to compare MV-optimal designs with EMA-optimal designs.

Table 16.3 demonstrates the efficiency of the A-optimal solution at each value of b in comparison to the one-sided $\alpha = 0.05$ EAA-optimal solution. The table gives both the average variance and the EAA for each design, along with the EAA efficiency. Table 16.3 gives similar results for the MV-optimal designs compared with those minimizing the EMA, including the EMA efficiency of the former. Asterisks in the n_0 columns of the table indicate those block sizes for which the two optimal solutions differ. As an example, for $b = 11$, the MV- and EMA-optimal designs from Table 16.3 differ as follows (columns are blocks):

MV-optimal design:

0	0	0	0	0	0	0	0	0	0	1
1	1	1	1	1	2	2	2	2	2	2

EMA-optimal design:

0	0	0	0	0	0	0	0	1	1	1
1	1	1	1	2	2	2	2	2	2	2

When the solutions differ, it is always in the same direction; that is, the EAA- or EMA-optimal solution requires fewer replications of the control. Thus, from the point of view of multiple comparisons, the A- and MV-optimal solutions often lead to designs with too many replications of the control. On the other hand, the efficiencies in Table 16.3 indicate that the EAA and EMA values for these less-than-perfect solutions do tend to be close to those of the true EAA-optimal and EMA-optimal designs. Similar results hold at different values of α (i.e., $\alpha = 0.01$ and $\alpha = 0.10$), as well as for the case of two-sided confidence intervals.

3. DESIGNS WITH LARGER BLOCK SIZES

For $k = 3$, a complete block design exists and it is shown in Section 4.5 of Bortnick (1999) that these are, indeed, EAA-optimal and EMA-optimal in the class $D\{p, b, k\}$ with $p = 2$, $k = 3$, as well as being A- and MV-optimal.

For $k > 3$, the problem becomes more complex. However, nearly optimal solutions still can be identified in an efficient and straightforward manner since, for any given control layout $\mathbf{n}_0$, lower bounds on the $\mathrm{EAA}_j(\alpha)$ and $\mathrm{EMA}_j(\alpha)$ can still be obtained. Then, by considering all possible control layouts, an overall lower bound on the $EAA_j(\alpha)$ and $EMA_j(\alpha)$ can be obtained for designs in $D\{2, b, k\}$, $k > 3$. Lemma 4.9 of Bortnick (1999) greatly reduces the total number of control layouts that need to be considered in the search for the overall bound. Once the overall bound is identified, a design is constructed that

Number of Blocks	A-optimal Design				MV-optimal Design			
	n_0	Average Variance	$EAA_1(0.05)$	EAA-Efficiency	n_0	Maximum Variance	$EMA_1(0.05)$	EMA-Efficiency
3	2	0.667	7.764	1.0000	2	0.667	7.764	1.0000
4	3	0.500	2.721	1.0000	∗4	0.500	2.851	0.9913
5	4	0.375	1.838	1.0000	4	0.375	1.838	1.0000
6	5	0.318	1.502	1.0000	4	0.333	1.507	1.0000
7	6	0.267	1.289	1.0000	6	0.267	1.289	1.0000
8	∗7	0.237	1.162	0.9825	6	0.238	1.151	1.0000
9	∗8	0.208	1.059	0.9918	8	0.208	1.059	1.0000
10	8	0.188	0.974	1.0000	8	0.188	0.974	1.0000
11	9	0.171	0.914	1.0000	∗10	0.171	0.923	0.9883
12	10	0.156	0.862	1.0000	10	0.156	0.862	1.0000
13	∗11	0.144	0.821	0.9965	10	0.145	0.819	1.0000
14	∗12	0.133	0.783	0.9992	12	0.133	0.783	1.0000
15	∗13	0.125	0.752	0.9905	12	0.125	0.749	1.0000
16	∗14	0.117	0.723	0.9933	∗14	0.117	0.723	0.9983
17	14	0.110	0.695	1.0000	14	0.110	0.695	1.0000
18	∗15	0.104	0.673	0.9995	∗16	0.104	0.676	0.9891
19	16	0.098	0.652	1.0000	16	0.098	0.652	1.0000
20	∗17	0.093	0.633	0.9958	16	0.094	0.632	1.0000
21	∗18	0.089	0.616	0.9971	18	0.089	0.616	1.0000
22	18	0.085	0.599	1.0000	18	0.085	0.599	1.0000
23	∗20	0.081	0.586	0.9946	∗20	0.081	0.586	0.9951
24	20	0.078	0.570	1.0000	20	0.078	0.570	1.0000
25	∗21	0.075	0.558	0.9980	∗22	0.075	0.560	0.9908
26	∗22	0.072	0.546	0.9988	22	0.072	0.546	1.0000
27	∗23	0.069	0.535	0.9961	22	0.069	0.534	1.0000
28	∗24	0.067	0.524	0.9969	∗24	0.067	0.524	0.9978
29	24	0.064	0.514	1.0000	24	0.064	0.514	1.0000
30	∗26	0.062	0.506	0.9946	∗26	0.062	0.506	0.9946
31	∗26	0.060	0.495	0.9996	26	0.060	0.495	1.0000
32	∗27	0.058	0.487	0.9977	∗28	0.058	0.489	0.9921
33	∗28	0.057	0.479	0.9982	∗28	0.057	0.479	0.9992
34	∗29	0.055	0.472	0.9965	28	0.055	0.471	1.0000
35	∗30	0.053	0.465	0.9967	∗30	0.053	0.465	0.9967
36	∗30	0.052	0.457	1.0000	30	0.052	0.457	1.0000
37	∗31	0.050	0.450	0.9986	∗32	0.050	0.451	0.9948
38	∗32	0.049	0.444	0.9989	∗32	0.049	0.444	0.9999
39	∗33	0.048	0.438	0.9977	32	0.048	0.438	1.0000
40	∗34	0.047	0.432	0.9980	∗34	0.047	0.432	0.9980
41	34	0.046	0.426	1.0000	34	0.046	0.426	1.0000
42	∗36	0.044	0.421	0.9964	∗36	0.044	0.421	0.9964
43	∗36	0.043	0.416	0.9994	36	0.043	0.416	1.0000
44	∗37	0.042	0.411	0.9984	∗38	0.042	0.411	0.9951
45	∗38	0.041	0.406	0.9987	∗38	0.041	0.406	0.9988
46	∗39	0.041	0.401	0.9978	38	0.041	0.401	1.0000
47	∗40	0.040	0.397	0.9975	∗40	0.040	0.397	0.9975
48	∗40	0.039	0.392	0.9997	40	0.039	0.392	1.0000
49	∗42	0.038	0.389	0.9964	∗42	0.038	0.389	0.9964
50	∗42	0.037	0.384	0.9991	∗42	0.037	0.384	0.9994

Table 16.3 $EAA_1(0.05)$- and $EMA_1(0.05)$-efficiency of A- and MV-optimal designs

is balanced in the test treatments. Such designs have $\mathrm{EAA}_j(\alpha)$ and $\mathrm{EMA}_j(\alpha)$ values close to their respective upper bounds. Details can be found in Bortnick (1999).

In a search for an efficient $\mathrm{EMA}_2(0.05)$ design solution in the design space $D\{2, 6, 5\}$, the following design was obtained. The bound on the EMA is 1.060, the actual EMA is 1.065, so the efficiency is at least 0.9896. This design solution is a Balanced Treatment Block design with parameters $\lambda_{01} = \lambda_{02} = 18$ and $\lambda_{12} = 12$.

0	0	0	0	0	0
0	0	0	0	0	0
1	1	1	1	1	1
1	1	1	2	2	2
2	2	2	2	2	2

References

Bortnick, S.B. (1999). *Optimal Block Designs for Simultaneous Comparison of Test Treatments with a Control.* PhD Dissertation, Ohio State University.

Daly, D.A., Cooper, E.B, (1967). Rate of stuttering adaptation under two electro-shock conditions. *Behaviour, Research and Therapy* **5**, 49–54.

Dunnett, C.W., (1955). A multiple comparison procedure for comparing several treatments with a control. *JASA* **50**, 1096–1121.

Hedayat, A., Jacroux, M., and Majumdar, D. (1988). Optimal designs for comparing test treatments with controls. *Statistical Science* **3**, 462–476.

Hsu, J. C., Berger, R. L. (1999). Stepwise confidence intervals without multiplicity adjustment for dose response and toxicity studies. *JASA* **94**, 468–482.

John, P. W. M. (1961). An application of a balanced incomplete design, *Technometrics* **3**, 51–54.

Marshall, A. and Olkin, I. (1979). *Inequalities: Theory of Majorization and Its Applications*. N.Y: Academic Press.

Spurrier, J.D., Nizam, A. (1990). Sample size allocation for simultaneous inference in comparison with control experiments. *JASA* **85**, 181–186.

Tong, Y.L. (1980). *Probability Inequalities in Multivariate Distributions*. N.Y.: Academic Press.

Tukey, J.W. (1994). The problem of multiple comparisons. *The Collected Works of John W. Tukey*. N.Y., London: Chapman & Hall.

US Environmental Protection Agency (1994) *Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms, 3rd edn*. Cincinnati, Ohio: Office of Research and Development, US Environmental Protection Agency.

Chapter 17

OPTIMAL SAMPLING DESIGN WITH RANDOM SIZE CLUSTERS FOR A MIXED MODEL WITH MEASUREMENT ERRORS

A. Giovagnoli
Department of Statistical Sciences, University of Bologna, Via Belle Arti, 41
40126 Bologna, Italy
giovagno@stat.unibo.it

L. Martino
National Institute of Statistics, via A. Ravà, 150
00172 Roma, Italy
Martino@istat.it

Abstract Our investigation concerns sampling in epidemiological studies, in the presence of both strata and clusters; the problem is to choose the number of clusters to sample in each stratum given that the size of the clusters in general is a random variable. The issue of unplanned randomness in the design seems to have been scarcely addressed in the survey sampling literature. We were motivated by a sample survey - carried out in 1990–1995 by the Italian National Institute of Nutrition (INN-CA) - on the food habits of the Italian population, divided into four geographical areas: the household came in both as random factor which influenced the individual response and - due to the varying number of its members - as a random component of the design which affected the sample size. In this paper we assume various mixed models under different hypothesis on measurement errors (typically correlated) in the response and for each of them find the optimal designs under several optimality criteria, namely the determinant, the trace, the maximum eigenvalue of the unconditional Fisher information of the fixed effect parameters. In all the models we deal with in the present paper, the optimal design depends on just one unknown parameter τ, a given function of the variance components and correlation coefficients. The dependence of the design on τ is investigated through some simulations. The solutions given for the special cases motivated by the INN-CA study should be applicable to a wider variety of situations.

Keywords: correlated errors, epidemiological studies, measurement errors, mixed models, random designs, sample surveys.

Introduction

Traditionally, the theories of Sampling Design and of Experimental Design have not developed side by side, nor have they in general borrowed much methodology from each other, with few noticeable exceptions e.g. in some early papers by Wynn (1977a, 1977b). Some of the reasons for such a lack of interchange among them - albeit fairly superficial ones - are put forward in Fienberg and Tanur (1987, 1996), but fundamentally the two theories are different because they address different inferential problems.

Nutritional studies and, more generally epidemiological ones which investigate the effects of various factors on health, fall in the in-between area between sample surveys and experiments: on one hand, there is no control by the investigators on the levels of the factors in question and the design problem only concerns the sampling strategy, while on the other, because of the presence of measurement errors, error-in-variables etc, and because of the interest in prediction, the need for a statistical model that describes the response is now widely recognized in the literature. Well-established tools from the theory of optimal experimental designs could usefully be employed within this context.

Our investigation will concern sampling in the presence of both strata (namely fixed block effects, in the design-of-experiment terminology) and clusters (namely random whole-plot effects); the problem is to choose the number of clusters to sample in each stratum given that the size of the clusters in general is a random variable. For instance, in 1990–1995, the Italian National Institute of Nutrition carried out a sample survey (INN-CA) - see Turrini (1996) - on the food habits of the Italian population, divided into four geographical areas, by interviewing all the members of selected households: the household came in both as a random factor which influenced the individual response and - due to the varying number of its members - as a random component of the design which affected the sample size. A classical way of dealing with this problem consists in choosing a sample that reproduces the distribution in the population of the random quantities. The issue of unplanned randomness in the design - as opposed to randomized designs, in which the random element is there intentionally for inferential purposes - has been scarcely addressed in the survey sampling literature (see for instance Rao, 1979 and Bellhouse, 1984).

In this paper we assume a mixed model with measurement errors (typically correlated) in the response and find the optimal designs under several optimality criteria, namely the determinant, the trace and the maximum eigenvalue of the unconditional Fisher information of the fixed effect parameters. Such designs are mostly only locally optimal, as they depend on unknown variance components and correlation coefficients; however, in dietary investigations,

the presence of validation studies often allows independent estimates of some of these unknown parameters, which makes the optimal design problem more tractable. It may be of interest to point out that, in all the models we deal with in the present paper, the optimal design depends on just one unknown parameter τ, a given function of the variance components and correlation coefficients. Thus we only need to make a "best guess" of τ. The dependence of the design on τ is investigated in this paper through some simulations. The solutions given for the special cases motivated by the INN-CA study should be applicable to a wider variety of situations.

1. DESCRIPTION OF THE MODEL

Assume we observe a continous response, affected by one or more block (i.e. qualitative) factors, on all the units of the whole plots (clusters) nested within the blocks. In addition we assume that the observations are subject to measurement errors, and thus we need to replicate them in order to make estimates more precise and to derive an estimate of the measurement error variances (for a complete review of measurement error models see Fuller, 1987, and Carroll et al., 1995). In Martino (1999a and 1999b) this model was used to describe the daily total fat intake - measured as a proportion of the total energy intake - in the INN-CA nutritional study of the Italian population. The nutrient intake was measured by interviewing all the members (the units) of selected households (the clusters) on their food habits. Even though a "golden method" - such as the weighted food record (see Bingham, 1987) was used, the responses were known to be affected by errors, both at the household and at the individual level. The block effect was due to the geographical area (the stratum), and there were also covariates, namely sex and age.

For simplicity, in this paper we deal with a model with one block factor and without covariates. The model can be expressed as

$$Y_{kjlr} = \beta_k + \alpha_{kj} + \varepsilon_{kjl} + \eta_{kjr} + \nu_{kjlr}, \tag{17.1}$$

where

$k = 1, 2, ..., K$, $j = 1, 2, ..., n_k$, $l = 1, 2, ..., m_{kj}$, $r = 1, 2, ..., R$;

β_k is a fixed effect;

α_{kj} is the random component due to cluster j within stratum k;

η_{kjr} is the measurement error of replicate r in cluster j within stratum k;

ε_{kjl} is the random error of the individual observation, namely variation of the individual unit with respect to the mean response;

ν_{kjlr} is the measurement error in replicate r on unit l in cluster j within stratum k.

Let $N = R\sum_{k=1}^{K}\sum_{j=1}^{n_k} m_{kj} = R\sum_{k=1}^{K} N_k$ be the total number of observations, with $n = \sum_{k=1}^{K} n_k$ = the total number of clusters and $N_k =$

$\sum_{j=1}^{n_k} m_{kj}$ = the total number of observations in stratum k. From now on let I_h denote the h-dimensional identity matrix, J_h the h-dimensional matrix of all ones, $\underline{1}_h$ the h-dimensional vector of all ones and $\underline{1}_h^t$ its transpose. The expected value of the response can be expressed as $E(\underline{Y}) = X\underline{\beta}$, where

$$X = \begin{pmatrix} \underline{1}_{RN_1} & 0 & .. & 0 \\ 0 & \underline{1}_{RN_2} & .. & 0 \\ ... & ... & .. & 0 \\ 0 & 0 & .. & \underline{1}_{RN_K} \end{pmatrix}.$$

The ML estimates of the parameter vector β are obtained through an iterative procedure. Hartley and Rao (1967) and Searle (1970) show that the asymptotic variance-covariance matrix of these estimates is the Fisher information of the unknown β's, given by

$$\Im_C = X^t \Sigma^{-1} X, \tag{17.2}$$

whence

$$V\left(\widehat{\beta}\right) \approx \left(X^t \Sigma^{-1} X\right)^{-1}.$$

Observe that (17.2) clearly depends on the design through the matrix X, hence we speak of the *conditional* information matrix $\Im_C$. When observations in different clusters are independent, this matrix is the sum of information matrices corresponding to the clusters,

$$\Im_C = \sum_{k=1}^{K} \sum_{j=1}^{n_k} (X_{kj})^t \, \Sigma_{kj}^{-1} X_{kj}, \tag{17.3}$$

where X_{kj} denotes the design matrix relative to cluster j of stratum k, and Σ_{kj} is the covariance of observations in cluster j of stratum k.

1.1. UNCORRELATED MEASUREMENT ERRORS

The random quantities α_{kj} are assumed i.i.d., normally distributed, with zero mean and variance σ_α^2. Similarly for ε_{kjl}, for η_{kjr} and for ν_{kjlr}, with variances σ_ε^2, σ_η^2 and σ_ν^2 respectively (the assumption of constant variances appears reasonable in a nutritional context). The α's, ε's, η's and ν's are also assumed to be jointly independent. Thus observations relative to different clusters are independent. Clearly, in the absence of replications, the contributions to variance coming from η_{kjr} and from ν_{kjlr} cannot be distinguished from those coming from α_{kj} and ε_{kjl} respectively, unless one has further information from other sources. The variance-covariance matrix of the observations is given by

$$\begin{aligned} \Sigma = Cov(\underline{Y}) &= diag \begin{pmatrix} ..., \left(J_{m_{kj}} \otimes J_R\right) \sigma_\alpha^2 + \left(I_{m_{kj}} \otimes J_R\right) \sigma_\varepsilon^2 + \\ \left(J_{m_{kj}} \otimes I_R\right) \sigma_\eta^2 + \left(I_{m_{kj}} \otimes I_R\right) \sigma_\nu^2, ... \end{pmatrix} \\ &= diag\left(...., I_{m_{kj}} \otimes \left(\sigma_\nu^2 I_R + \sigma_\varepsilon^2 J_R\right) + J_{m_{kj}} \otimes \left(\sigma_\eta^2 I_R + \sigma_\alpha^2 J_R\right), ...\right). \end{aligned} \tag{17.4}$$

Letting $A = \sigma_\nu^2 I_R + \sigma_\varepsilon^2 J_R$ and $B = \sigma_\eta^2 I_R + \sigma_\alpha^2 J_R$, (17.4) is written as

$$\Sigma = diag\left(...., I_{m_{kj}} \otimes A + J_{m_{kj}} \otimes B, ...\right). \tag{17.5}$$

Applying (17.5) to (17.3), after some algebra the conditional information can be written as a block diagonal matrix

$$\Im_C = diag\left(....., \sum_{j=1}^{n_k} \underline{1}_{Rm_{kj}}^t \left[I_{m_{kj}} \otimes A + J_{m_{kj}} \otimes B\right]^{-1} \underline{1}_{Rm_{kj}}, ...\right),$$

whence

$$\Im_C = R \cdot diag\left(....., \sum_{j=1}^{n_k} \frac{m_{kj}}{\sigma_\nu^2 + R\sigma_\varepsilon^2 + m_{kj}\left(\sigma_\eta^2 + R\sigma_\alpha^2\right)}, ...\right). \tag{17.6}$$

Letting

$$\tau = \frac{\frac{\sigma_\eta^2}{R} + \sigma_\alpha^2}{\frac{\sigma_\nu^2}{R} + \sigma_\varepsilon^2} \tag{17.7}$$

one gets

$$\Im_C = \frac{1}{\frac{\sigma_\nu^2}{R} + \sigma_\varepsilon^2} diag\left(....., \sum_{j=1}^{n_k} \frac{m_{kj}}{1 + m_{kj}\tau}, ...\right). \tag{17.8}$$

1.2. CORRELATED MEASUREMENT ERRORS

Often the hypothesis of independence of measurement errors does not appear to be justified, and it seems preferable to assume some form of dependence among different measurements. Let the assumptions on α_{kj} , ε_{kjl} , η_{kjr}, ν_{kjlr} be the same as before, except that now we suppose that, for given (k, j, l), the errors ν_{kjlr} are correlated for different r's and, for given (k, j), the errors η_{kjr} are correlated for different r's. Let $Cov\left(\underline{\nu}_{kjl}\right) = \sigma_\nu^2 \Omega_\nu$ and $Cov\left(\underline{\eta}_{kj}\right) = \sigma_\eta^2 \Omega_\eta$ with Ω_ν and Ω_η being R-dimensional matrices. The variance-covariance matrix of the observations is now

$$\Sigma = Cov\left(\underline{Y}\right) = diag\left(\begin{array}{c} ...., \left(J_{m_{kj}} \otimes J_R\right)\sigma_\alpha^2 + \left(I_{m_{kj}} \otimes J_R\right)\sigma_\varepsilon^2 \\ + \left(J_{m_{kj}} \otimes\ \Omega_\eta\right)\sigma_\eta^2 + \left(I_{m_{kj}} \otimes \Omega_\nu\right)\sigma_\nu^2, ... \end{array}\right) \tag{17.9}$$

$$= diag\left(...., I_{m_{kj}} \otimes \left(\sigma_\nu^2 \Omega_\nu + \sigma_\varepsilon^2 J_R\right) + J_{m_{kj}} \otimes \left(\sigma_\eta^2 \Omega_\eta + \sigma_\alpha^2 J_R\right), ...\right).$$

Letting $\tilde{A} = \sigma_\nu^2 \Omega_\nu + \sigma_\varepsilon^2 J_R$ and $\tilde{B} = \sigma_\eta^2 \Omega_\eta + \sigma_\alpha^2 J_R$, (17.9) can be rewritten as

$$\Sigma = diag\left(...., I_{m_{kj}} \otimes \tilde{A} + J_{m_{kj}} \otimes \tilde{B}, ...\right), \tag{17.10}$$

so that

$$\Im_C = diag\left(....., \sum_{j=1}^{n_k} m_{kj} \underline{1}_R^t \left[\tilde{A} + m_{kj}\tilde{B}\right]^{-1} \underline{1}_R, ...\right). \tag{17.11}$$

We shall work out the details of just two special cases.

Equicorrelated errors. Let $corr(\nu_{kjlr}, \nu_{kjls}) = \rho_\nu$ and $corr(\eta_{kjr}, \eta_{kjs}) = \rho_\eta$ for all $r \neq s$. Then $\Omega_\nu = (1 - \rho_\nu) I_R + \rho_\nu J_R$ and $\Omega_\eta = (1 - \rho_\eta) I_R + \rho_\eta J_R$ whence

$$\tilde{A} = \sigma_\nu^2 \Omega_\nu + \sigma_\varepsilon^2 J_R = \tilde{\sigma}_\nu^2 I_R + \tilde{\sigma}_\varepsilon^2 J_R$$

and

$$\tilde{B} = \sigma_\eta^2 \Omega_\eta + \sigma_\alpha^2 J_R = \tilde{\sigma}_\eta^2 I_R + \tilde{\sigma}_\alpha^2 J_R,$$

where

$\tilde{\sigma}_\nu^2 = \sigma_\nu^2(1 - \rho_\nu)$, $\tilde{\sigma}_\varepsilon^2 = \rho_\nu \sigma_\nu^2 + \sigma_\varepsilon^2$, $\tilde{\sigma}_\eta^2 = \sigma_\eta^2(1 - \rho_\eta)$, $\tilde{\sigma}_\alpha^2 = \rho_\eta \sigma_\eta^2 + \sigma_\alpha^2$.

Thus, if we let

$$\tilde{\tau} = \left(\frac{\tilde{\sigma}_\eta^2}{R} + \tilde{\sigma}_\alpha^2\right) / \left(\frac{\tilde{\sigma}_\nu^2}{R} + \tilde{\sigma}_\varepsilon^2\right) \tag{17.12}$$

the expression for the conditional Fisher information parallels that of (17.8) for uncorrelated errors

$$\Im_C = \frac{1}{\frac{\tilde{\sigma}_\nu^2}{R} + \tilde{\sigma}_\varepsilon^2} diag\left(....., \sum_{j=1}^{n_k} \frac{m_{kj}}{1 + m_{kj}\tilde{\tau}}, ...\right). \tag{17.13}$$

First-order autocorrelated errors. We shall now assume that there are no measurement errors related to the clusters so that $\sigma_\eta^2 = 0$. Assume furthermore that at the unit level the measurement errors follow a first-order autoregressive process. This assumption is fairly realistic in nutritional studies, when observations are repeated over a period of time, as suggested by several authors (see for instance Wang et al., 1996). Under these assumptions, $corr(\nu_{kjlr}, \nu_{kjls}) = \rho$ for every $r \neq s$ with $|\rho| < 1$. Writing

$$\Omega_\nu = \begin{bmatrix} 1 & \rho & ... & \rho^{R-1} \\ \rho & 1 & ... & \rho^{R-2} \\ ... & ... & ... & ... \\ \rho^{R-1} & \rho^{R-2} & ... & 1 \end{bmatrix}$$

and $\tilde{B} = \sigma_\alpha^2 J_R$ in (17.11) we get:

$$\underline{1}_R^t \left(\sigma_\nu^2 \Omega_\nu + (\sigma_\varepsilon^2 + m\,\sigma_\alpha^2) J_R\right)^{-1} \underline{1}_R = \frac{\tilde{R}}{\sigma_\nu^2 + (\sigma_\varepsilon^2 + m\,\sigma_\alpha^2)\tilde{R}},$$

where $\tilde{R} = \underline{1}^t \Omega_\nu^{-1} \underline{1} = \dfrac{R - R\rho - 2\rho}{1+\rho}$, hence

$$\begin{aligned}
\Im_C &= diag\left(....., \sum_{j=1}^{n_k} m_{kj} \underline{1}_R^t \left[\tilde{A} + m_{kj}\tilde{B}\right]^{-1} \underline{1}_R, ...\right) \\
&= diag\left(....., \sum_{j=1}^{n_k} m_{kj} \frac{\tilde{R}}{\sigma_\nu^2 + (\sigma_\varepsilon^2 + m_{kj}\sigma_\alpha^2)\tilde{R}}, ...\right) \\
&= \frac{1}{\frac{\sigma_\nu^2}{\tilde{R}} + \sigma_\varepsilon^2}\, diag\left(....., \sum_{j=1}^{n_k} \frac{m_{kj}}{1 + m_{kj}\widetilde{\tilde{\tau}}}, ...\right),
\end{aligned}$$

with

$$\widetilde{\tilde{\tau}} = \frac{\sigma_\alpha^2}{\frac{\sigma_\nu^2}{\tilde{R}} + \sigma_\varepsilon^2}. \tag{17.14}$$

2. THE DESIGN

We want to choose optimal designs for the models of Sections 1.1 and 1.2. In general when speaking of an optimal design it is understood that the total sample size is kept fixed, since it is easy to see that information increases when the number of observations, and thus the cost, increases. However in this paper, as already pointed out, the size of the clusters is a random variable, which makes the sample size random. It seems reasonable therefore to fix a quantity related to the total sampling cost, very likely $n = \sum_{k=1}^{K} n_k$, the total number of sampled clusters, or, more generally, if the sampling costs c_k are different from stratum to stratum, the quantity $c = \sum_{k=1}^{K} c_k n_k$. It may be of interest to note that in the classical theory of survey sampling, if the sizes of the clusters are different they are usually assumed to be known a priori and the selection probabilities of the primary sampling units are taken to be proportional to their size; there is no accounting for the case when the size is random - see Cochran (1977).

Several criteria are used in the experimental design literature (see for instance Silvey, 1980) for the choice of the optimal design. They are usually functions of the variance-covariance matrix of the fixed effects estimators or of the Fisher

information matrix relative to the β's. We cannot obviously base the choice of the design on the conditional information, even when it is the one that will be used in the inference. We suggest choosing the design that optimizes the unconditional Fisher information, which will be denoted by $\Im$. In all the cases of the previous Section, the matrix $\Im_C$ is of the form

$$\Im_C \propto diag\left(....., \sum_{j=1}^{n_k} \frac{m_{kj}}{1+m_{kj}\tau}, ...\right),$$

with a suitable interpretation of τ each time. Thus

$$\begin{aligned}\Im = E(\Im_C) &\propto diag\left(....., E\left(\sum_{j=1}^{n_k} \frac{m_{kj}}{1+m_{kj}\tau}\right), ...\right) \\ &= diag\left(....., n_k \cdot E\left(\frac{M_k}{1+M_k\tau}\right), ...\right) \\ &= diag\left(....., n_k \cdot \Psi_k\left(M_k, \tau\right), ...\right),\end{aligned}$$

where M_k denotes the random variable "number of units in cluster k" $(k = 1, ..., K)$ and we define

$$\Psi_k\left(M_k, \tau\right) = E\left(\frac{M_k}{1+M_k\tau}\right). \tag{17.15}$$

A more accurate procedure would be to choose a design which minimizes some function of the unconditional variance-covariance matrix of the fixed effects estimators, i.e. $W = E(V\left(\widehat{\beta}\right))$. This is **not** the inverse of the unconditional Fisher information. However, the computation of $\Im$ is simpler and $\Im^{-1}$ is a good large sample approximation to W.

2.1. THE CRITERION OF THE DETERMINANT

If we want to minimize the generalized asymptotic variance of the ML estimators of the vector β, the target function to be maximized is the determinant of the Fisher information, i.e.

$$\det \Im = \frac{1}{(\frac{\sigma_\nu^2}{R} + \sigma_\varepsilon^2)^K} \prod_{k=1}^{K} n_k \prod_{k=1}^{K} \Psi_k\left(\tau\right), \tag{17.16}$$

where we have written $\Psi_k\left(\tau\right) = \Psi_k\left(M_k, \tau\right)$ for short. Subject to a fixed total cost c, the function (17.16) is a maximum when the $c_k n_k$'s are equal for all k, independently of the unknown σ's, i.e. when$n_i = c/(Kc_i)$ for all $i = 1, 2, \ldots, K$. In this case the size of the clusters does not influence the design.

2.2. THE EIGENVALUE CRITERION (MINIMAX)

If we want to minimize the maximum asymptotic variance of the estimators, the target function is the smallest eigenvalue of matrix $\Im$, namely $min_k\, n_k \Psi_k(\tau)$. The constrained problem

$$\begin{cases} Max \min_k n_k \Psi_k(\tau) \\ \sum_k^K c_k n_k = c \end{cases}$$

has the solution

$$n_i = \frac{1}{\Psi_i(\tau)} \cdot \frac{c}{\sum_{k=1}^K \frac{c_k}{\Psi_k(\tau)}} \qquad \text{for } i = 1, 2, ..., K. \tag{17.17a}$$

2.3. THE TRACE CRITERION

If we want to minimize the sum of the asymptotic variances of the estimators, namely of the entries $\frac{1}{n_k \Psi_k(\tau)}$ on the main diagonal of $\Im^{-1}$, we need to solve the constrained problem

$$\begin{cases} \min \left(\sum_{k=1}^K \frac{1}{n_k \Psi_k(\tau)} \right) \\ \sum_k^K c_k n_k = c \end{cases},$$

which yields

$$n_i = \frac{1}{\sqrt{c_i \Psi_i(\tau)}} \cdot \frac{c}{\sum_{k=1}^K \sqrt{\frac{c_k}{\Psi_k(\tau)}}} \qquad \text{for} \quad i = 1, 2, ..., K. \tag{17.18}$$

3. SOME REMARKS ON τ AND M_K

The parameter τ defined in (17.7) quantifies the relative weight of the variance due to clusters (that of the mean measurement error in the cluster plus the cluster variance component) and the variance of the units (that of the mean measurement error of each unit plus the variability among different units). The meaning of (17.12) and (17.14) is more complicated but similar. The value of τ is in general unknown: however, there may sometimes be partial information on it. For instance, several authors (e.g. in Blumer et al., 1991) emphasize that the weight of measurement errors, especially those at the unit level, in explaining variance is often significantly greater than that of the other variance components.

Since the function $f(x) = \frac{x}{1+\tau x}$ is increasing in x for all $\tau > 0$, it can be shown (see e.g. Ross, 1996, Proposition 9.1.2) that if there is a stochastic

dominance relation $\geq_{st}$ between the random variables M_k and M_h, i.e. if

$$\Pr(M_k > u) \geq \Pr(M_h > u) \qquad \forall\, u,$$

then the order relation carries over to the expected values

$$E(f(M_k)) \geq E(f(M_h)).$$

Thus, for $\tau > 0$, $\Psi_k(\tau, M_k)$ is increasing with respect to stochastic increases of M_k It is also decreasing with respect to τ. Therefore, by the results of (17.17a) and (17.18), the optimal number of clusters to sample within each stratum depends on the frequency distribution of the cluster size in the population and on the relative weight of the variance components relative to the cluster and to the unit. As an example, a simulation has been carried out to show the dependence of the optimal design on τ, under equal costs c_k. Our motivation was the INN-CA study, relative to four Italian geographical areas (North-West, North-East, Centre and South); thus we let $K = 4$ in model (17.1); the clusters are the households and the distribution of the number M_k of household members in the geographical areas has been estimated from the Indagine Multiscopo ISTAT (ISTAT, 1998) - see Table 17.1 and Martino (1999a).

	Cumulative distribution function			
M_k	F_1	F_2	F_3	F_4
1	0.2452	0.1967	0.2385	0.1801
2	0.5290	0.4787	0.5264	0.4023
3	0.7812	0.7502	0.7468	0.6133
4	0.9580	0.9399	0.9424	0.8733
5	0.9945	0.9887	0.9877	0.9712
6	0.9986	0.9977	0.9981	0.9945
7	0.9999	0.9998	0.9995	0.9979
$8-10$	1.0000	1.0000	1.0000	1.0000
Mean	2.49	2.65	2.56	2.97

Table 17.1 Cumulative distribution functions of household sizes in the four regions

Observe that there is an almost perfect stochastic ordering of M_k in some of the different geographical areas, namely M_4 is greater than all the other M_k's and M_1 is the smallest; M_2 and M_3 are not ordered, being approximately equivalent, but the mean household size in area 2 is larger than the mean size in area 3.

The design results with respect to varying τ are shown in Table 17.2, from which we observe that the effect of the household size on the optimal sampling

	Eigenvalue criterion				Trace criterion			
τ	n_1/N	n_2/N	n_3/N	n_4/N	n_1/N	n_2/N	n_3/N	n_4/N
0.01	0.266	0.250	0.259	0.224	0.258	0.250	0.255	0.237
0.05	0.264	0.250	0.258	0.227	0.257	0.250	0.254	0.239
0.1	0.262	0.250	0.258	0.230	0.256	0.250	0.254	0.240
0.5	0.257	0.250	0.255	0.239	0.253	0.250	0.252	0.244
0.8	0.255	0.248	0.255	0.242	0.252	0.249	0.252	0.246
0.9	0.255	0.248	0.255	0.242	0.253	0.249	0.253	0.246
0.99	0.256	0.249	0.253	0.242	0.253	0.250	0.251	0.246
1	0.254	0.250	0.254	0.243	0.252	0.250	0.252	0.246
1.01	0.256	0.249	0.253	0.242	0.253	0.250	0.251	0.246
1.05	0.257	0.249	0.253	0.242	0.253	0.249	0.251	0.246
1.1	0.256	0.248	0.252	0.244	0.253	0.249	0.251	0.247
1.3	0.253	0.249	0.253	0.245	0.252	0.249	0.252	0.247
1.5	0.255	0.250	0.250	0.245	0.252	0.250	0.250	0.248
2	0.253	0.247	0.253	0.247	0.252	0.248	0.252	0.248
3	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250
10	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250

Table 17.2 Optimal sample size by stratum

numbers becomes greater with increasing difference in the frequency distributions among strata. This effect, which leads to sampling more households in the areas with fewer household members (like Area 1 in our example) is more evident when the value of τ is small. For instance, if $\tau = 0.5$, the optimal sampling proportions according the the minimax criterion are 0.257, 0.255, 0.250 and 0.239 respectively for Areas 1, 3, 2 and 4, i.e. 8% more in 1 than in 4, whereas for $\tau \geq 3$ the difference in the optimal sampling proportions is negligible from a practical point of view.

It may be of interest to compare these results with those from the theory of stratified sampling with proportional allocation (the stratum variances do not play any role in this case, because they are all the same); since the latter optimizes the sampling only with respect to the size of the strata, without taking into account the size of the primary sampling units, the sampling proportions in the strata are substantially different from the ones in Table 17.2 for any τ, since they are 0.199, 0.145, 0.184 and 0.472 if the relative frequencies of family size per area are estimated on the basis of the Indagine Multiscopo (1998). This result establishes an important tool for comparison, since stratified sampling was actually used in designing the INN-CA study.

4. LINKS WITH THE EXISTING DESIGN LITERATURE

Ever since Fisher and Neyman, random and mixed models have been the focus of much attention and heated statistical debate (see for instance Samuels et al., 1991, and the related discussion). In general the literature which refers to such models (under more complicated block structures than the one presented here) emphasizes the analysis of designed experiments, rather than planning issues, but there are some exceptions, e.g. Kachlicka and Mejza (1990). In the absence of measurement errors, $\sigma_\eta^2 = \sigma_\nu^2 = 0$ and our model (17.1) becomes

$$Y_{kjl} = \beta_k + \alpha_{kj} + \varepsilon_{kjl}, \tag{17.19}$$

where β is fixed, α randomly nested in β and ε randomly nested in α with $k = 1, 2, ..., K$, $j = 1, 2, ..., n_k$, $l = 1, 2, ..., m_{kj}$, and $\tau = \sigma_\alpha^2/\sigma_\varepsilon^2$.

This model is formally identical to Eq.(4.2) in Giovagnoli and Sebastiani (1989) and, obviously, the expression for the Fisher information relative to the fixed effects is the same. But the optimal solution of Giovagnoli and Sebastiani (1989) for a given N, consisting of a balanced design with all m_{kj} equal, clearly does not make sense in the present context, given the different nature of the values m_{kj}. This remark applies to papers by other authors too, who show, like Roebruck (1987), that for a split-plot experiment a balanced or nearly balanced design is to be preferred.

Clearly mixed models and correlated errors are very closely related topics; one can refer to the vast literature on experimental design for the latter: Kunert is a prolific author here. Measurement errors are a special case, since they are usually correlated: an authorative review of the relevant designs is given by Afsarinejad (1990). Their importance in survey sampling has been emphasized for some time now by several authors; some literature of great relevance is collected in Blumer et al. (1991). Fuller (1995) mentions implications for the design of surveys.

Lastly, we would like to draw attention to the robustness aspects of the analysis, as dealt with by Guseo (1997), that could be paralleled by robustness studies of the design, which so far have not been looked into.

We end by stressing that the optimal plans of the present article do not take into account the need to estimate variance components. In particular, the number of replications R has been kept constant. It would be of interest to explore the interrelation between the optimal number of clusters to be sampled and the optimal number of replicates of individual units. It has been suggested (Kaaks et al., 1995; Stram et al., 1995) that the latter should be kept low with respect to the former.

Acknowledgements. This research was carried out within the Italian National MURST Cofin98 Research Programme "Methods for Planning Experiments and their Applications". The authors are grateful to Prof. Carroll for providing useful suggestions about the model and to Dr Camporese and Dr D'Amicis for providing the Multiscopo and the INN-CA data.

References

Afsarinejad, K. (1990). Repeated measurements designs: a review. *Comm Statist. A - Theory Methods* **19**, 3985–4028.

Bellhouse, D.R. (1984). A review of optimal designs in survey sampling. *Canad. J. Statistics* **12**, 53–65.

Bingham, S.A. (1987). The dietary assessment of individuals: methods, accuracy, new techniques and recommendations. *Nutr. Abstr. Rev.* A **57**, 705–42.

Blumer, P.P., Groves, R.M., Lyberg, L.E., Mathiowetz, N.A. and Sudman, S. (Eds) (1991). *Measurement Errors in Surveys* . Wiley: New York.

Carroll, R.J., Ruppert, D. and Stefanski, L.A. (1995). *Measurement Error in Nonlinear Models*. London: Chapman and Hall.

Cochran,W.G., (1977). *Sampling Techniques*. Wiley: New York.

Fienberg, S.E. and Tanur, J.M. (1987). Experimental and sampling structures: parallels diverging and meeting. *Int. Statist.Rev.* **53**, 75–96.

Fienberg, S.E. and Tanur, J.M. (1996). Reconsidering the fundamental contributions of Fisher and Neyman on experimentation and sampling. *Int. Statist.Rev.* **64**, 237–254.

Fuller, W.A., (1987). *Measurement Error Models*. New York: Wiley.

Fuller, W.A. (1995). Estimation in the presence of measurement error. *Int. Statist. Rev.* **63**, 121–147.

Giovagnoli, A. and Sebastiani, P. (1989). Experimental designs for mean and variance estimation in variance components models. *Comp Statist. Data Analysis* **8**, 21–28. Correction in *Comp Statist. Data Analysis* (1990) **10**, 93.

Guseo, R. (1997). Split-plot design: extension and robustness aspects. *Statist. Appl.* **9**, 61–79.

Hartley, H.O. and Rao, J.N.K. (1967). Maximum likelihood estimation for the mixed analysis of variance models. *Biometrika* **54**, 93–108.

ISTAT, (1998). *Indagine Statistica Multiscopo sulle Famiglie: Aspetti della Vita Quotidiana*. Anno 1997. Roma: ISTAT.

Kachlicka, D. and Mejza, S. (1990). Incomplete spit-plot experiments: whole plot treatments in a row-column design. *Comp. Statist. Data Analysis* **9**, 273–281.

Kaaks, R., Riboli, E. and van Staveren, W. (1995). Sample size requirements for calibration studies of dietary intake measurements in prospective cohort investigations. *Am. J. Epidemiol.* **142**, 557–565.

Martino, L. (1999a). *Problema della Stima e Aspetti del Disegno di Campionamento in Modelli a Effetti Misti con Errori di Misura*. Ph.D. Thesis, Dipartimento di Scienze Statistiche, Università di Bologna.

Martino, L. (1999b). A multilevel mixed linear model with measurement errors. In *Proceedings of the XIV International Workshop on Statistical Modelling, Graz, 19-23 July 1999*, pp. 580–83. Heidelberg: Physica-Verlag.

Rao, J.N.K. (1979). Optimization in the design of sample surveys. In: *Optimizing Methods in Statistics: Proceedings of an International Conference* Ed. J.S. Rustagi, pp. 419–434. New York: Academic Press.

Roebruck, P. (1987). Optimal designs for 3-factorial split-plot models with random block effects. *, J Statist. Planning Inference* **17**, 293–305.

Ross, S. (1989). *Stochastic Processes*. New York: Wiley.

Samuels, M.R., Casella, G. and McCabe, G.P. (1991). Interpreting blocks and random factors (with discussion). *J. Amer. Statist. Assoc.* **86**, 798–808.

Stram, D.O., Longnecker, M.P. and Shames, L. (1995). Cost-efficient design of a diet validation study. *Am. J. Epidemiol.***142**, 353–362.

Searle, S.R. (1970). Large sample variances of maximum likelihood estimators of variance components. *Biometrics* **26**, 505–524.

Silvey, S.D. (1980). *Optimal Designs*. London: Chapman and Hall.

Turrini, A. (1996). Survey form applied in a national food consumption study. In *Proceedings of the III Meeting dell'Azione Concertata AIR-CAT "Methodologies in Product Development and Market Research"*. Barcelona, 3–5 October 1996.

Wang, N., Carroll, R.J. and Liang, K.Y. (1996). Quasilikelihood and variance functions in measurement error models with replicates. *Biometrics* **52**, 401–411.

Wynn, H.P. (1977a). Minimax purposive survey sampling design. *J. Amer. Statist. Assoc.* **72**, 655–657.

Wynn, H.P. (1977b). Optimum designs for finite population sampling. In *Statistical Decision Theory and Related Topics II* Eds S.S. Gupta and J. Berger, pp. 471–478. New York: Academic Press.

Chapter 18

OPTIMIZING A UNIMODAL RESPONSE FUNCTION FOR BINARY VARIABLES

Janis Hardwick
Purdue University, West Lafayette, IN 47907
jphard@umich.edu

Quentin F. Stout
University of Michigan, Ann Arbor, MI 48109
qstout@umich.edu

Abstract Several allocation rules are examined for the problem of optimizing a response function for a set of Bernoulli populations, where the population means are assumed to have a strict unimodal structure. This problem arises in dose response settings in clinical trials. The designs are evaluated both on their efficiency in identifying a good population at the end of the experiment, and in their efficiency in sampling from good populations during the trial. A new design, that adapts multi-arm bandit strategies to this unimodal structure, is shown to be superior to the designs previously proposed. The bandit design utilizes approximate Gittin's indices and shape constrained regression.

Keywords: adaptive, clinical trial, dose-response, experimental design, multi-arm bandit, nonparametric, Polya urn, random walk, sequential sampling, stochastic approximation, unimodal regression, up and down

1. INTRODUCTION

Consider a problem in which there are k linearly ordered populations or "arms". Associated with Arm i is a binary random variable, Y_i, the outcome of which is governed by an unknown distribution function $F(i)$. The expected return from Arm i, $p_i = \mathbf{E}(Y_i)$, is referred to as its *value*. Of interest here is the problem of identifying the arm with the highest value, given a strict unimodal structure on the values $p_1, \ldots, p_k$. As unimodality is location invariant, the arms may be located at integers $\{1, \ldots, k\}$ without any loss of generality. There are n

opportunities to sample and allocation rules or sampling algorithms are assessed according to two measures — a sampling error and a decision error.

We evaluate four nonparametric sampling designs — three from the literature and a new one proposed here. Since a common application motivated development of several of these designs, we briefly review the application here. See Durham et al. (1998) and Hardwick et al. (2000) for details.

1.1. AN APPLICATION

Classical dose response problems focus on locating specified quantiles of the relationship between drug dose and probability of toxic response to drug therapy. A common framework for these problems is to model a patient's response at dose s with a Bernoulli random variable with success probability (i.e., non-toxic outcome) $1 - Q(s)$, where $Q(s)$ is a continuous nondecreasing function.

We focus on a variation of this problem in which there are only k doses. We wish to maximize the probability that a patient being treated exhibits not only a *non-toxic* response but is also *cured* at a given dose. Let $R(i)$ be a nondecreasing response curve that models the probability that dose i is effective, and take $F(i) = R(i)(1 - Q(i))$ to be the product curve for efficacy and nontoxicity, $i = 1, \ldots, k$. Note that in many common parametric dose response settings, the curve F is unimodal.

The goal is to develop a sampling design that identifies the optimal dose, $i^* = \arg\max_i F(i)$. The problem may be formulated with R conditioned on Q, and here we take the special case of R and Q independent. Further, while in some scenarios we are provided with the individual outcomes from R and Q, in this case, we assume that only the outcome of F is observed (see Hardwick et al., 2000).

In clinical trials there are typically two populations of patients to consider — the n trial subjects and the unknown number of patients in the future who will be affected by the terminal decision of the experiment. We consider a good design to be one that has a relatively high probability that both

(i) trial subjects will be cured, and
(ii) the optimal dose is selected as best at the trial's termination.

Since these two criteria are opposing, a single optimal design for this problem doesn't exist. Instead, we seek designs that lie on or close to an optimal trade-off curve representing performance along these two measures. Note that it is not known how to optimize either measure in this setting.

In this paper, we compare a new shape constrained multiarm bandit sampling rule with three other designs that have been proposed for virtually the same

problem. In Sections 2.1 and 2.2, respectively, we describe urn and up and down designs as proposed in Durham et al. (1998). In Section 2.3, we describe a stochastic approximation type of design delineated in Herkenrath (1983). The bandit design, adapted to the unimodal structure, is outlined in Section 2.4. In Section 3 we describe evaluation criteria and experimental results. In Section 4 we discuss asymptotic behaviour, and in Section 5 we close with a short discussion.

2. SAMPLING DESIGNS

We begin with some notation. Recall that i^* denotes the best arm and let p^* be the value of Arm i^*. Arm i_m is sampled at stage m (the m^{th} observation). Let $\mathcal{I}_{im} = 1$ if $i = i_m$ and 0 otherwise. Then $n_{im} = \sum_{j=1}^{m} \mathcal{I}_{ij}$ is the number of observations sampled from Arm i by stage m for $i = 1, \ldots, k$; $m = 1, \ldots, n$. Thus $\sum_{i=1}^{k} n_{im} = m$.

Let Y_{im} represent the outcome of Arm i at trial m. For convenience we take Y_{im} to be 0 unless $i = i_m$, in which case, Y_{im} has a Bernoulli outcome with success rate p_i. Then, $r_m = \sum_{i=1}^{k} Y_{im}$ is the return or "reward" received at stage m. Two designs considered here take observations in pairs, (Y_1, Y_2) as opposed to individually. To keep notation consistent, we set $Y_1(ij) = Y_{im}$ and $Y_2(ij) = Y_{i(m+1)}$ for $j = 1, \ldots, n/2$, $m = 1, \ldots, n-1$.

The empirical mean $\hat{p}_{im}$ of p_i after stage m is given by $\sum_{j=1}^{m} Y_{ij}/n_{im}$. For Bayesian designs we assume that each p_i follows an independent beta distribution with parameters a_i, b_i. Thus the posterior mean $\bar{p}_{im}$ of p_i after stage m is given by $(a_i + \sum_{j=1}^{m} Y_{ij})/(a_i + b_i + n_{im})$.

To specify a design, one needs:

– A *sampling rule* that determines which arm to sample at each stage. This may involve special rules to handle startup, boundaries, ties, and observations near the end of the experiment.

– A *terminal decision rule*, to determine the arm declared best at the end of the experiment.

2.1. RANDOMIZED POLYA URN DESIGN

A *randomized Polya urn for selecting optima* is proposed in Durham et al. (1998). In this design, arms are sampled according to a draw from the urn. For each i in $\{1, \ldots, k\}$, the urn initially contains $\alpha_i > 0$ balls labeled i. (In the present case, $\alpha_i = 1$). At stage m, $m = 1, \ldots, n$:

(i) A ball is drawn at random from the urn (and replaced), and an observation is taken from the arm corresponding to the ball's label;

(ii) If the response is a success, then another ball with the same label is added to the urn;

(iii) If the response is a failure, then no new balls are added.

A stopping rule is associated with this urn process and the authors' urn contains some additional information that pertains only to the stopping time. Here, in order to compare the urn design with the others discussed, we assume a fixed sample size. The terminal decision rule is to select the arm with the highest number of balls in the urn as best.

2.2. UP AND DOWN DESIGN

Random walks or up and down designs are well known for the problem of locating quantiles of a non-decreasing dose response function (see Dixon (1965) and Flournoy et al. (1995)). For the case in which the response curve is strictly unimodal, Durham et al. (1998) propose an *up and down rule for targeting the optimum*, defined as follows.

At each stage j, $j = 1, \ldots, \frac{n}{2}$, a pair of arms is sampled. Let $M(j)$ represent the midpoint between the two arms sampled at stage j. Observations $Y_1(ij)$ and $Y_2(ij)$, are taken at $M(j) - c$ and $M(j) + c$, respectively, where $c = \frac{b}{2}$ and b is an odd positive integer. The midpoint for the next two observations at stage $j + 1$ is given by $M(j + 1) = M(j) + V_j$, where

$$V_j = \begin{cases} 1 & \text{if } Y_1(ij) = 0 \text{ and } Y_2(ij) = 1 \\ 0 & \text{if } Y_1(ij) = 0 \text{ and } Y_2(ij) = 0 \\ & \text{or } Y_1(ij) = 1 \text{ and } Y_2(ij) = 1 \\ -1 & \text{if } Y_1(ij) = 1 \text{ and } Y_2(ij) = 0. \end{cases}$$

If $M(j + 1)$ would fall outside the range, then $M(j + 1) = M(j)$. In practice (and herein), b is typically 1, in which case the start-up rule for the process selects $M(1) = 1.5$, the midpoint of the leftmost two arms.

Durham et al. (1998) do not specify a terminal decision rule, but here we assume that the decision is to select the arm having the largest sample mean, i.e., $\arg\max_i \hat{p}_{in}$. Note that with a Markov chain such as this, one would ordinarily select the site most visited as best. For this design, however, such a decision process does not always converge to the optimal arm in the limit so the empirical mean is used instead.

2.3. STOCHASTIC APPROXIMATION DESIGN

Another approach for this problem is to use a Keifer-Wolfowitz type of stochastic approximation rule for locating a local maximum (Kiefer and Wolfowitz, 1952; Sacks, 1958). Typically a design of this sort samples at points

$X(j) - c_j$ and $X(j) + c_j$, $c_j > 0$, where $\lim c_j \to 0$. The "decision point" $X(j+1)$ is given by

$$X(j+1) = X(j) + \frac{a_j}{c_j} V_j,$$

where the sequence $\{a_j > 0\}$ is such that $\Sigma a_j = \infty$ and $\Sigma a_j^2 c_j^{-2} < \infty$.

One difficulty with the above rule is that the sampling points $X(j) \pm c_j$ do not lie on a lattice. Herkenrath (1983) proposed a modified procedure adapted both to the discrete case and to Bernoulli observations. With this procedure, observations are only taken at arms representing the left and right neighbours of $X(j)$. The method is as follows. Let $0 < d < 1/2$ and $q(x) = x - x^l$ where $x^l = \lfloor x \rfloor$ (though k^l is set to $k-1$), and let $x^r = x^l + 1$.

According to Herkenrath (1983), the process begins with $X(1) = 1$. At stage j, allocation of 2 observations is randomized between left and right neighbours of $X(j)$ according to their relative distances. The sampling is guided by the position of $x = X(j)$ as specified in the first column of the following table. The succeeding three columns of the table display the probabilities of sampling pairs of observations as indicated at the top of the columns.

If $\downarrow$ then sample $\rightarrow$	x^l, x^l	x^l, x^r	x^r, x^r
$x < x^l + d$	$1 - q(x^l + d)$	$q(x^l + d)$	0
$x^l + d \leq x < 1/2(x^l + x^r)$	$1 - q(x)$	$q(x)$	0
$x = 1/2(x^l + x^r)$	0	1	0
$1/2(x^l + x^r) < x \leq x^r - d$	0	$1 - q(x)$	$q(x)$
$x^r - d < x$	0	$1 - q(x^r - d)$	$q(x^r - d)$

Table 18.1 Stochastic Approximation Design: probabilities of sampling left and right neigbours of $X(j)$

Once sampling has taken place, the $j + 1^{st}$ decision point is given by

$$X(j+1) = Proj_{[1,k]}\{X(j) + a_j S(Y_1(ij), Y_2(ij), X(j))\},$$

where $a_j \to 0$ and S is a time invariant function that depends on $X(j)$ and the outcomes of the observations at stage j (see Herkenrath, 1983, for details).

For this design, the terminal decision is to select the arm closest to $X(n+1)$.

2.4. UNIMODAL BANDIT DESIGN

In general k-arm bandit problems, the k populations have unknown reward structures, *arm pulling* or sampling takes place sequentially, and decisions are

made with the goal of optimizing a discounted sum of all returns, $\Sigma_1^n \beta_m r_m$ for discount sequence $\{\beta_m \geq 0\}$ (see Berry and Fristedt (1985) for details). This formulation covers a broad class of problems in learning theory, and optimal strategies balance the impulse to earn immediate rewards against the need to gather information for future decisions.

It is natural to model initial information about the arms of a bandit using a prior distribution on the "values", p_i, $i = 1, \ldots, k$, of the arms. As sampling takes place, posterior distributions reflect the additional information acquired. Optimal strategies for Bayesian bandits with finite "horizon" or sample size can be determined via dynamic programming (Bellman, 1956). However, computing such solutions is a daunting task. For the simple case involving independent Bernoulli arms, the dynamic programming equations require computational space and time that grow as $n^{2k}/(2k-1)!$ for a problem of horizon n. (See Hardwick et al., 1999, for the largest problems yet solved).

As the bandit model becomes more complex (e.g., the number of outcomes per arm increases or there is structure on the arms), the problem quickly outgrows available computer resources. In the present situation, we face k *dependent* arms, and thus can obtain optimal solutions in only trivial cases. One option in handling these more complex problems, however, is to exploit known characteristics of optimal solutions to the simpler bandit problems.

In particular, there is a bandit model that, appropriately parameterized, approximates the independent finite horizon model with discount sequence $\beta_j = 1,\ j = 1, \ldots, n$. This model, the *geometric* bandit in which $\beta_j = \beta^{j-1}$ for $0 < \beta < 1$ and $j = 1, 2, \ldots$, offers a surprising solution. Note first that the infinite horizon in this model makes dynamic programming impractical. However, with independent arms, optimal solutions for the geometric bandit may be defined in terms of index rules. Thus, at each stage of the experiment and for each arm, there exists an index depending only on the arm, such that sampling from the arm with the highest index yields the optimal solution, (Gittins and Jones, 1974). Known as Gittin's Indices, these quantities incorporate available "knowledge" about an arm along with the arm's value. The existence of an index rule greatly simplifies the solution to the geometric bandit problem because it reduces the complexity of the k-arm bandit from being exponential in k to being linear in k. However, despite the elegant solution presented in Gittins and Jones (1974), the indices themselves are extremely complicated. Except in very simple models, they cannot be calculated exactly. Still, the idea of using an index rule with traits similar to the Gittin's index has great appeal. Because the objective function of the finite horizon problem is to maximize return, it is reasonable to ask how well the bandit solution accommodates statistical goals of gathering information for inferential purposes.

Here, we use a lower bound for the Gittin's index proposed in Hardwick (1995). If the beta prior parameters for an arm are A and B, then a lower bound

for the Gittins index for this arm is given by $\Lambda^* = \sup\{\Lambda_r : r = 1, 2, \ldots\}$, where

$$\Lambda_r = \frac{\frac{\Gamma(A+1)}{\Gamma(A+B+1)} - B\Sigma_1^r \beta^i \frac{\Gamma(A+i)}{\Gamma(A+B+i+1)}}{\frac{\Gamma(A)}{\Gamma(A+B)} - B\Sigma_1^r \beta^i \frac{\Gamma(A+i-1)}{\Gamma(A+B+i)}}.$$

Generally speaking, it is not difficult to compute Λ^* since Λ_r is a unimodal function of r.

Interestingly, geometric discounting can be viewed as an ethically equitable mechanism for balancing the well being of current and future patients. The parameter β represents the relative weight of the outcome of each subject as compared to the weight of all future patients. In the finite horizon setting, optimal decisions will be virtually identical if $n \sim 1/(1-\beta)$ when $n \to \infty$ and $\beta \to 1$. This relationship is quite accurate when n is as low as, say, 30.

In the present problem, we do not focus on these interpretations, but instead wish to choose β to optimize our joint goals of minimizing loss and making good decisions. It is an open question how best to do this. For simplicity in our simulation experiments, a fixed value of $\beta = 0.95$ was used for all sample sizes. Naturally, somewhat better results would be obtained if β changed somehow with n. Note that it may even make sense to adjust β during a given experiment. Since exploration is best early in the experiment, higher values may be more useful then, with lower values towards the end helping to diminish experimental losses.

The *unimodal bandit* sampling algorithm used to generate the results in the next section is constructed as follows. Let the prior distributions of the p_i be beta with parameters (a_i, b_i) (we use $a_i = b_i = 1$). At stage $m+1$,

(i) Calculate the posterior means, $\bar{p}_{im}$, $i = 1, \ldots, k$;

(ii) Using least squares, fit the best unimodal curve to the posterior means using weights $n_{im} + a_i + b_i$ (Stout and Hardwick, 2000);

(iii) Adjust the posterior distributions so that their means lie on the curve from (ii) by adding the smallest (fractional) number of successes, u_{im}, or failures, v_{im}, needed to bring the posterior means into alignment. For Arm i the *adjusted posterior parameters* are

$$A_{im} = a_i + \sum_{j=1}^{m} Y_{ij} + u_{ij} \text{ and } B_{im} = b_i + n_{im} - \sum_{j=1}^{m} Y_{ij} + v_{ij};$$

(iv) For each arm i, calculate $\Lambda^*_{i,m}$, based on the adjusted posterior distributions in (iii), and let $j = \arg\max_i \Lambda^*_{i,m}$. If there are ties for the maximum, then pick j at random from the tied arms;

(v) Determine where to take the next observation:

(a) If observation m was a success, or was not on arm j, then sample next from arm j;

(b) *Exploration rule:* If observation m was a failure and was on arm j, then pick j' uniformly from $\{j-1, j+1\}$ for $j' \in \{1, \ldots, k\}$. If the p-value for the one sided test of equality of j and j' is at least $1/(n-m+1)$ then sample from j', otherwise sample again from j.

At the end of the experiment, fit a unimodal curve to the weighted posterior means and select the arm corresponding to the mode.

Note the *exploration rule* in Step v. This is used to avoid premature convergence to a suboptimal arm. While the allocation index also serves this purpose, this may not force as much exploration as needed. The exploration rule also lends robustness to prior misspecification.

3. EVALUATION CRITERIA AND RESULTS

As discussed, we seek designs that behave well along two performance measures – an experimental or sampling error to assess losses during the experiment and a decision error to assess future losses based on the terminal decision.

Given any decision rule and sampling design there exists a probability measure on the arms which reflects the chance, $\pi_n(i)$, that Arm i is selected as best at the end of an experiment of size n. One could take $\pi_n(i^*)$, the *probability of correct selection*, as a measure of decision efficiency. However, in clinical trials, a major goal is to minimize harm to patients. In such cases, selecting an arm with a success rate close to p^* is much preferred to selecting those with lower rates. This leads to the following definition:

$$\text{Decision Efficiency}: \quad \mathcal{D}_n = (\Sigma_{i=1}^{k} \pi_n(i) p_i)/p^*.$$

The sampling error is the normalized expected loss incurred when sampling from arms other than i^*. Noting that $\mathbf{E}[r_m]$ is the expected return at stage m, we define

$$\text{Sampling Efficiency}: \quad \mathcal{S}_n = (\Sigma_{m=1}^{n} \mathbf{E}[r_m])/(n \cdot p^*),$$

which is closely related to the *expected successes lost*, $\Sigma_{m=1}^{n} p^* - \mathbf{E}[r_m]$.

An initial simulation study was conducted to examine the relative behaviour of the designs in Section 2. To assess the impact of variation in the unimodal functions themselves, we selected four curves with diverse characteristics. These are shown in Figure 18.1. For each curve and each design, we consider sample sizes 25, 50, 100 and 200. At each configuration 5,000 iterations of the experiments were run and used to estimate the decision efficiency

and sampling efficiency, shown in Figure 18.2. Figure 18.2 also includes results for equal allocation, in which there are $\lfloor k/n \rfloor$ rounds of sampling from each arm and the remaining observations are assigned at random and without replacement to $\{1, \ldots, k\}$.

Generally speaking, the bandit design performed the best along both measures, and no design was uniformly worst. For the urn, bandit and equal allocation rules, the efficiencies move from the southwest towards the northeast as the sample size increases, which is the behaviour that one would expect. However, for the up and down design, this movement is reversed in Curves 2 and 3, and the stochastic approximation design shows no consistent behaviour.

These anomalies point to the impact of design details that are often ignored. Start-up rules, for example, can have a grave impact on design performance. For the urn method, there is a random start, while for the bandit design, the start is dictated by the prior distribution, which in the present case yields a random start. For the up and down and stochastic approximation designs, however, sampling begins with the leftmost arm, as designated by the developers of the designs. If a curve has its mode to the left, these designs tend initially to stay near the mode. For example, with Curve 3, which is strictly decreasing, the up and down design samples *only* from Arms 1 and 2 nearly 9% of the time when $n = 25$. This is primarily due to observing repeated failures on these arms yet not moving away. Note, however, that as the sample sizes increase, this tendency is overcome and the design's performance stabilizes.

We reran the experiments with the ordering on the arms reversed, i.e., so that curves that originally were decreasing were now increasing and so forth. This switch caused the efficiencies of the up and down and stochastic approximation designs to plummet for Curves 2 and 3, whereas they improved for Curve 4.

To draw more direct comparisons among the rules, we also reran the experiments using a random start-up for all designs. The results of these runs appear in Figure 18.3. For the up and down design, the unbiased start-up reordered the efficiencies upward and to the right with the sample size. Further, the results show an overall improvement over those obtained using the leftmost start-up. Note, however, that the preferred start-up rule may depend on the application. For the dose response example discussed in the Introduction, Durham et al. (1998) determine that it is preferable to have a patient succumb to disease than it is to be exposed to toxic levels of drug therapy, and thus it was deemed desirable to begin with low rather than arbitrary drug doses.

Figure 18.3 shows that, even when utilizing a random start-up, the stochastic approximation design still behaved erratically. In Curve 1, for example, the value for $n = 25$ was wildly high on the E_n measure – so much so that it has been omitted from the figure. A less extreme but similar behaviour can be observed in Curve 3. For neither of these curves did stochastic approximation show the desired trend of efficiency increasing with the sample size. We believe that this

unusual behaviour may be the result of the choice of the S function mentioned in Section 2.3. Herkenrath (1983) notes that the system of equations defining S are underdetermined and thus adds conditions to guarantee unique solutions. The specific form of S chosen by the author appears to be questionable, and it is likely that a different solution would improve performance and produce less erratic behaviour.

While the bandit rule performs the best with regard to decision making, note the up and down and equal allocation rules have similar $\mathcal{D}_n$ when n is large. As expected, however, the bandit rule outmatches all rules on the $\mathcal{S}_n$ measure. Among the adaptive rules examined, the urn model seems to perform least well on our measures. However, the main strength of this model is that it is randomized and can better protect patients from selection bias.

4. CONVERGENCE

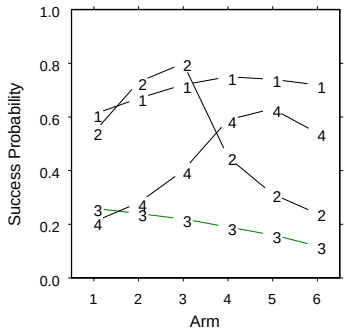


Figure 18.1 Unimodal curves used for evaluation.

While in practice these designs are intended for use with small to moderate sample sizes, it is also important that they exhibit good asymptotic efficiency, i.e., that $\mathcal{D}_n$ and $\mathcal{S}_n$ converge almost surely (*a.s.*) to 1 as $n \to \infty$. If a design is asymptotically efficient with respect to both kinds of error, then the design is asymptotically (first order) optimal. In this section, we assume that efficiency is "asymptotic" and thus drop the adjective.

To obtain sampling efficiency, it is necessary and sufficient that the rate at which arm i^* is sampled goes to 1 as $n \to \infty$, while decision efficiency requires that the probability of selecting arm i^* as best goes to 1 as $n \to \infty$. Note that for a design to insure that it has not prematurely converged in its selection of a good arm, with the unimodal assumption it is necessary that arms $i^* \pm 1$ be sampled infinitely often. Without such an assumption, all arms would have to be sampled infinitely often (*i.o.*).

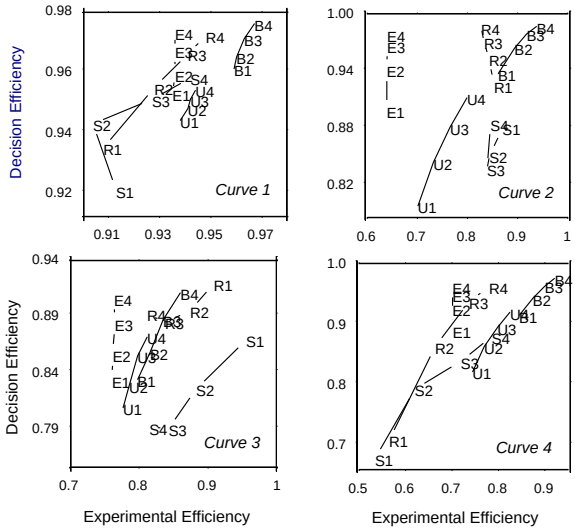


Figure 18.2 Efficiencies, for n = (1) 25, (2) 50, (3) 100, (4) 200. B = Bandit, E = Equal, R = Up-Down, S = Stochastic Approximation, U = Urn. Start-up as proposed with methods

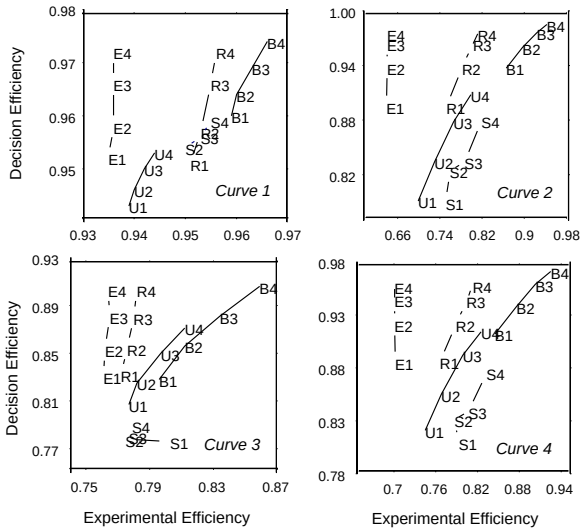


Figure 18.3 Efficiencies, for n = (1) 25, (2) 50, (3) 100, (4) 200. B = Bandit, E = Equal, R = Up-Down, S = Stochastic Approximation, U = Urn. Random start-up for all methods

Equal Allocation: It is straightforward to show that this design is decision efficient, but not sampling efficient.

Up and Down Design: The decision rule for this design selects the arm corresponding to $\max_i \hat{p}_{in}$ as best. Since all arms are sampled *i.o.*, $\hat{p}_{in} \to p_i$ *a.s.* for each i. Thus, choosing the arm with the highest sample mean guarantees *a.s.* selection of Arm i^* in the limit, i.e., the up and down rule is decision efficient. However, because the asymptotic rate at which suboptimal arms are sampled is nonzero, the design is not sampling efficient. Note that unimodality is not required for decision efficiency.

Urn Design: Let pr_i be the asymptotic proportion of balls of type i in the urn as $n \to \infty$. In Theorem 2 of Section 4.2, Durham et al. (1998) show that $pr_{i^*} = 1$ and $pr_i = 0, i \neq i^*$ *a.s.* Since the decision rule chooses the arm with the maximum proportion of balls at stage n as best, the rule *a.s.* makes the correct selection as $n \to \infty$. Since $pr_{i^*} = 1$, the design is also sampling efficient. Note that unimodality is not required for the urn design to be both decision and sampling efficient.

Stochastic Approximation: Herkenrath (1983) shows that when the sequence $\{a_n\} \to 0$, $\sum_1^\infty a_n = \infty$ and $\sum_1^\infty a_n^2 < \infty$, then $X(n)$ converges to i^* *a.s.* Thus, in the limit, the decision to select the arm closest to $X(n)$ as best is *a.s.* the optimal decision, and hence this design is decision efficient. Unfortunately, with the probability of sampling adjacent pairs being fixed, this design cannot be sampling efficient.

Unimodal Bandit: The exploration rule, part vb, ensures that the arm thought to be best and its neighbours are sampled *i.o.*, to avoid premature convergence. Under the unimodal assumption, it can be shown that this insures that asymptotically the bandit *a.s.* selects i^* as the best arm, and thus is decision efficient (see Hardwick and stout, 2000). Further, since the rate of sampling other arms goes to zero, the design is also sampling efficient. Note that if β is bounded away from 1, then the arms not adjacent to i^* are *a.s.* sampled only finitely often. Also note that sampling efficiency can be lost if β goes to 1 too rapidly as n increases.

The conditions needed for first-order sampling efficiency are not difficult to achieve. However, to obtain second-order efficiency, more delicate control of the rates at which suboptimal arms are sampled is needed. For the unimodal bandit design, the discount parameter β and the exploration rule together dictate this rate, as ordinarily would the sequence $\{a_n\}$ for the stochastic approximation design. (As mentioned, it is the deterministic sampling of adjacent pairs in the Herkenrath (1983) design that precludes sampling efficiency.) One must be careful, however, in changing the convergence rates since they affect *both* efficiency types. In fact, it is precisely these rates that control the trade-off

between the competing goals of gaining high reward ($\mathcal{S}_\infty$) and drawing good inferences ($\mathcal{D}_\infty$). Forcing the sampling of suboptimal arms to go to zero too fast reduces the rate at which we arrive at an optimal decision.

In summary, for the unimodal problem, a design needs to sample the best arm and its neighbours *i.o.* to be decision efficient. However, sampling too much from suboptimal arms negatively impacts a design's experimental regret. Thus a good rule will be one that carries out enough exploration to determine the best arm and then samples mostly from it.

5. DISCUSSION

While asymptotic analytical results can give some guidelines, they don't appear to be able to determine which designs are best on useful sample sizes, and hence computer experiments are needed. The experiments reported here, and those in Hardwick and Stout (2000), show that while all of the designs considered are asymptotically decision efficient, for fixed sample sizes the unimodal bandit appears to do slightly better than the others, at least on the curves and sample sizes examined. It appears to achieve both a good sampling efficiency and good decision efficiency for a wide range of situations.

However, there is significant work needed to tune these basic approaches to produce better designs for this setting. We are conducting experiments to evaluate designs on large numbers of unimodal functions, and to evaluate a number of variations. For example, there are many alternatives possible for the exploration rule in the unimodal bandit, and it is unclear which is best. It is also not obvious how β should go to 1 as n goes to ∞, and whether β should be reduced as an experiment progresses, to optimize performance. Further, there is significant interaction between the choice of β and the exploration rule. This is a problem for future study. There are also many variations possible for random walks, urn designs, etc., and in initial experiments we have been able to achieve some improvement.

Returning to the motivating example of dose-response problems, we note that there is also the important case in which the competing failure modes, Q and R in the Introduction, are observed. As expected, one can do better by using designs which exploit the structure of the failure modes (see Hardwick et al., 2000).

References

Bellman, R. (1956). A problem in the sequential design of experiments. *Sankhya* A **16**, 221–229.

Berry, D.A. and Fristedt, B. (1985). *Bandit Problems: Sequential Allocation of Experiments*. London: Chapman and Hall.

Dixon, W. (1965). The up and down method for small samples. *J. Amer. Statist. Assoc.* **60**, 967–978.

Durham, S., Flournoy, N. and Li, W. (1998). A sequential design for maximizing the probability of a favourable response. *Can. J. Statist.* **26**, 479–495.

Flournoy, N., Durham, S., and Rosenberger, W. (1995). Toxicity in sequential dose-response experiments. *Sequential Anal.* **14**, 217–227.

Gittins, J.C. and Jones, D.M. (1974). A dynamic allocation index for the sequential design of experiments. In *Progress in Statistics* Eds J. Gani et al., pp. 241–266. Amsterdam: North Holland.

Hardwick, J. (1995). A modified bandit as an approach to ethical allocation in clinical trials. In *Adaptive Designs: Institute Math. Stat. Lecture Notes* **25**, Eds B. Rosenberger & N. Flournoy, pp. 223–237. Hayward CA: Institute of Mathematical Statistics.

Hardwick, J., Oehmke, R. and Stout, Q.F. (1999). A program for sequential allocation of three Bernoulli populations. *Comput. Stat. and Data Analysis* **31**, 397–416.

Hardwick, J., Meyer, M. and Stout, Q.F. (2000). Adaptive designs for dose response problems with competing failure modes. Submitted.

Hardwick, J. and Stout, Q.F. (2000). Bandit designs for optimizing a unimodal response function. In preparation.

Herkenrath, U. (1983). The N-armed bandit with unimodal structure. *Metrika* **30**, 195–210.

Keifer, J. and Wolfowitz, J. (1952). Stochastic estimation of the maximum of a regression function. *Ann. Math. Statist.* **25**, 529–532.

Sacks, J. (1958). Asymptotic distribution of stochastic approximation procedure. *Ann. Math. Statist.* **29**, 373–405.

Stout, Q.F. and Hardwick, J. (2000). Optimal algorithms for univariate unimodal regression. *Computing Science and Statistics* **32**. To appear.

Discussion

Nancy Flournoy

Consider a set of ordered Bernoulli populations for which $P\{\text{success}\}$ is strictly unimodal. The dual-purpose goal is to select the most successful Bernoulli population while allocating more subjects to the better treatment arms. Hardwick and Stout introduce a new design that adapts the multiarm bandit approach to the unimodal success probability structure, and they compare it to several designs in the literature. Let me start by saying that their *Unimodal Bandit Design* looks terrific! It is a great addition to the emergent literature on this problem.

The designs compared to the Unimodal Bandit Design are (1) a version of Stochastic Approximation restricted to operate on a lattice (Herkenrath, 1983), (2) the Randomized Polya Urn and (3) the Optimizing Up-and-Down Design; the latter two were proposed by Durham, Flournoy and Li (1988). Let me add another to the list for future comparisons: the Birth and Death Urn by Ivanova, Rosenberger, Durham and Flournoy (2000). Procedures for obtaining exact distribution functions for statistics of interest using the Birth and Death Urn are analogous to those given by Ivanova and Flournoy (2000) for ternary outcomes.

In general, the Unimodal Bandit Design performed as well or better than the other designs in every setting examined. In Figure 18.1 design parameters are set at values described in the literature. This includes starting the Randomized Polya Urn and the Optimizing Up-and-Down at the lowest dose, which may be dictated by experimental requirements, but certainly slows down convergence, considerably so for response functions that are skewed to the left (curves 1 and 4). Thus, I focus on the comparisons in Figure 18.3 where initial treatments are randomized. This may not be acceptable in some practical applications, but it makes more equitable comparisons.

I must say that I was surprised to see how well all the designs performed! This is very exciting. Additional criteria may determine the choice of a design. For example, the Randomized Polya Urn design converges slowly, and hence, tends to be less efficient than the other designs, but it offers the most randomization. So if protection against selection bias or the confounding effects of unknown covariates is a primary concern, this design may still be preferred over the others.

There is much fine tuning left to be done, which bodes well for all the designs. For example, alternative logical decision rules exist, and need to be compared. For instance, the decision rule studied for the Randomized Polya Urn is to select the treatment arm having the maximum number of balls in the urn at stage n. Alternatively, one could select the treatment with the highest proportion of successes.

Finally, it is important to recognize a related goal that is an important alternative to the one studied by Hardwick and Stout. Instead of seeking to select the best treatment arm from an ordered set, say, $\{1, ..., K\}$ of treatments, one could seek to infer the optimal treatment on the continuum $[1, K]$ from trials on the ordered set. With this objective, the interval between treatments is important and, for some applications, may not need to be fixed at unity. Also, different decision rules are needed. A reasonable rule is to obtain a smoothed isotonic estimate of a unimodal response function from the observed proportion of successes, and then to select the treatment with the maximum isotonic estimator on the interval. This decision rule has been proposed, and is being studied, by Kpamegan (2000) for the Optimizing Up-and-Down Design. Indeed, when the estimator of the optimal treatment belongs to an interval, designs whose

allocations do not converge to a single treatment (such as the Optimizing Up-and-Down Design, which converges to a distribution) may perform better than those that do converge to a point.

In conclusion, great progress has been made on the development of sequential treatment allocation procedures, operating on a lattice, that will reliably reveal the optimal treatment from a unimodal set and tend to treat more subjects at the better treatments. Hardwick and Stout have made a significant contribution to this effort with their Optimizing Bandit Design and with their analytical techniques. We now anticipate a rapid maturing of statistical approaches to this important practical problem.

Additional References

Ivanova, A. and Flournoy, N. (2000). A birth and death urn for ternary outcomes: stochastic processes applied to urn models. In *A Volume in Honor of Theophilos Cacoulous* Eds C.A. Charalambides, M.V. Koutras and N. Balakrishnan. Florida: CRC Press/Chapman and Hall. (In press).

Ivanova, A., Rosenberger, W.F., Durham, S.D. and Flournoy, N. (2000). A birth and death urn for randomized clinical trials: asymptotic methods. *Sankhya*. (To appear).

Kpamegan, E.E. (2000). *An Optimizing Up-and-Down Design*. Dissertation in progress, American University.

Chapter 19

AN OPTIMIZING UP-AND-DOWN DESIGN

Euloge E. Kpamegan and Nancy Flournoy
Department of Mathematics and Statistics
American University, Washington DC 20016-8050 U.S.A.
Ckeet9@aol.com and Flournoy@american.edu

Abstract Assume that the probability of success is unimodal as a function of dose. Take the response to be binary and the possible treatment space to be a lattice. The Optimizing Up-and-Down Design allocates treatments to pairs of subjects in a way that causes the treatment distribution to cluster around the treatment with maximum success probability. This procedure is constructed to use accruing information to limit the number of patients that are exposed to doses with high probabilities of failure. The Optimizing Up-and-Down Design is motivated by Kiefer and Wolfowitz's stochastic approximation procedure. In this paper, we compare its performance to stochastic approximation. As an estimator of the best dose, simulation studies demonstrate that the mode of the empirical treatment distribution using the Optimizing Up-and-Down Design converges faster than does the usual estimator using stochastic approximation.

Keywords: constant gain stochastic algorithm, Markov chains, phase I/II clinical trial, random walk designs, stationary treatment distribution, stochastic approximation

Introduction

Up-and-down rules are procedures that specify the treatment for the next trial to be one level higher, one level lower, or the same as the treatment selected for the current trial. Our procedure differs significantly from other up-and-down designs in the literature in that it is to be used when the probability of success $\alpha(x)$ is a unimodal function of dose rather than an increasing one. We call the dose having maximum probability of success the *optimal dose* and denote it by μ. Because the proposed designs cluster treatments around the optimal dose given a unimodal response function, they are called Optimizing Up-and-Down Designs. In Section 1 we define the Optimizing Up-and-Down Designs as proposed by Kpamegan and Flournoy (2000). In Section 2 we give the stationary treatment distribution. In Section 3 we give theoretical results

characterizing the stationary treatment distribution. In particular, we describe how the empirical mode converges 'near' to the optimal dose. Finally, we compare the performance of the up-and-down with stochastic approximation in Section 4

1. THE OPTIMIZING UP-AND-DOWN DESIGN

Consider a finite set of possible dosages $\Omega = \{x_1, \ldots, x_K\}$. Suppose the interval between possible dosages is a positive constant Δ, i.e., $x_j - x_{j-1} = \Delta$ for $j = 2, \ldots, K$. Let $X(n)$ be the midpoint of the dose interval for the nth pair of subjects. Then the Optimizing Up-and-Down Design for selecting the dose with maximum success probability is defined by the following algorithm, with adjustments at the treatment boundaries.

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

$$X(n+1) = X(n) + \Delta V(\mathbf{n}), \tag{19.1}$$

where

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

For each pair of subjects, we refer to the one treated at the lower (higher) dose as subject 1 (2) of the nth pair. Note from the recursive equation (19.1) that if $X(n) = x_k$, then the possible treatment levels for the individuals in the $(n+1)th$ pair are

$$X_1(n+1) \in \begin{cases} x_{k-1} - \frac{\Delta}{2} \\ x_k - \frac{\Delta}{2} \\ x_{k+1} - \frac{\Delta}{2} \end{cases} ; \quad X_2(n+1) \in \begin{cases} x_{k-1} + \frac{\Delta}{2} \\ x_k + \frac{\Delta}{2} \\ x_{k+1} + \frac{\Delta}{2} \end{cases} . \tag{19.2}$$

We refer to $\{X_1(n)\}$ and $\{X_2(n)\}$, respectively, as the left sequence and right sequence of treatments. On the boundaries, when the dose cannot go down or up, it stays constant.

Each sequence is a nonhomogeneneous random walk that converges to a stationary distribution with appropriate boundary conditions. This stationary distribution is described in terms of the transition probabilities between treatments in the next section.

2. THE STATIONARY TREATMENT DISTRIBUTION

Define $\overline{\alpha}(x) = 1 - \alpha(x) = P(\text{failure}|x)$. Let p_k, q_k, and r_k denote, respectively, the probability that the treatment dosage for the left subject will move down from level $x_k - \frac{\Delta}{2}$ to $x_{k-1} - \frac{\Delta}{2}$, up from level $x_k - \frac{\Delta}{2}$ to $x_{k+1} - \frac{\Delta}{2}$, and stay at level $x_k - \frac{\Delta}{2}$, with $p_k + q_k + r_k = 1$. Assume that subject 1 of the nth pair has just been allocated treatment $x_k - \frac{\Delta}{2}$. Then for the left sequence, the treatment transition probabilities for $k = 2, \ldots, K-1$ that result are

$$\begin{aligned} p_k &= P(X_1(n+1) = x_{k-1} - \frac{\Delta}{2} | X_1(n) = x_k - \frac{\Delta}{2}) \\ &= P(Y_1(n) = 1 | X_1(n) = x_k - \frac{\Delta}{2}) \times P(Y_2(n) = 0 | X_2(n) = x_k + \frac{\Delta}{2}) \\ &= \alpha(x_k - \frac{\Delta}{2})\overline{\alpha}(x_k + \frac{\Delta}{2}), \end{aligned} \tag{19.3}$$

$$\begin{aligned} q_k &= P\{X_1(n+1) = x_{k+1} - \frac{\Delta}{2} | X_1(n) = x_k - \frac{\Delta}{2}\} \\ &= P(Y_1(n) = 0 | X_1(n) = x_k - \frac{\Delta}{2}) \times P(Y_2(n) = 1 | X_2(n) = x_k + \frac{\Delta}{2}) \\ &= \overline{\alpha}(x_k - \frac{\Delta}{2})\alpha(x_k + \frac{\Delta}{2}), \end{aligned}$$

$$\begin{aligned} r_k &= P\{X_1(n+1) = x_k - \frac{\Delta}{2} | X_1(n) = x_k - \frac{\Delta}{2}\} \\ &= P(Y_1(n) = 1 | X_1(n) = x_k - \frac{\Delta}{2}) \times P(Y_2(n) = 1 | X_2(n) = x_k + \frac{\Delta}{2}) \\ &+ P(Y_1(n) = 0 | X_1(n) = x_k - \frac{\Delta}{2}) \times P(Y_2(n) = 0 | X_2(n) = x_k + \frac{\Delta}{2}) \\ &= \alpha(x_k - \frac{\Delta}{2})\alpha(x_k + \frac{\Delta}{2}) + \overline{\alpha}(x_k - \frac{\Delta}{2})\overline{\alpha}(x_k + \frac{\Delta}{2}). \end{aligned}$$

At the boundaries $p_1 = q_K = 0$, so that $r_1 = 1 - q_1 = 1 - \overline{\alpha}(x_1 - \frac{\Delta}{2}) \times \alpha(x_1 + \frac{\Delta}{2})$ and $r_K = 1 - p_K = 1 - \alpha(x_K - \frac{\Delta}{2}) \times \overline{\alpha}(x_K + \frac{\Delta}{2})$.

The transition probabilities (p'_k, q'_k, r'_k) for the right sequence are equivalent, i.e.,

$$\begin{cases} p'_k = P\left\{X_2(n+1) = x_{k-1} + \frac{\Delta}{2} | X_2(n) = x_k + \frac{\Delta}{2}\right\} = p_k, \\ q'_k = P\left\{X_2(n+1) = x_{k+1} + \frac{\Delta}{2} | X_2(n) = x_k + \frac{\Delta}{2}\right\} = q_k, \\ r'_k = P\left\{X_2(n+1) = x_k + \frac{\Delta}{2} | X_2(n) = x_k + \frac{\Delta}{2}\right\} = r_k. \end{cases}$$

As long as $r_k > 0$ and $r'_k > 0$ for some $k, k' = 1, \ldots, K$, the allocation of treatments for each sequence describes an irreducible recurrent Markov chain. Thus

left and right stationary distributions $\pi = (\pi_1, \ldots, \pi_K)$ and $\pi' = (\pi'_1, \ldots, \pi'_K)$ exist (c.f. Harris, 1952), where $\pi_k = \lim_{n\to\infty} P\left\{X_1(n) = x_k - \frac{\Delta}{2}\right\}$ and $\pi'_k = \lim_{n\to\infty} P\left\{X_2(n) = x_k + \frac{\Delta}{2}\right\}$ for the left and right sequence of subjects, respectively.

Kpamegan and Flournoy (2000) evaluate the balance equations

$$\pi_k = \pi_{k-1}q_{k-1} + \pi_k r_k + \pi_{k+1}p_{k+1}, k = 1, \ldots, K,$$

as indicated by Karlin and Taylor (1975, pp. 86–87) to obtain the stationary treatment probabilities for the left sequence of subjects:

$$\pi_k = \pi_1 \prod_{j=2}^{k} \frac{\overline{\alpha}(x_{j-1} - \frac{\Delta}{2})}{\overline{\alpha}(x_j + \frac{\Delta}{2})}, j = 2, \ldots, k, \tag{19.4}$$

where

$$\pi_1^{-1} = 1 + \sum_{k=2}^{K} \prod_{j=2}^{k} \frac{\overline{\alpha}(x_{j-1} - \frac{\Delta}{2})}{\overline{\alpha}(x_j + \frac{\Delta}{2})}. \tag{19.5}$$

Because $q_k = q'_k$, $r_k = r'_k$, and $p_k = p'_k$, the right sequence has an identically shaped stationary treatment distribution that is shifted Δ to the right of the one for the left sequence. The stationary treatment distributions of the left and right sequences are defined on the lattices of points

$$\Omega_L = \left\{x_1 - \frac{\Delta}{2}, x_2 - \frac{\Delta}{2}, \ldots, x_K - \frac{\Delta}{2}\right\}$$

and

$$\Omega_R = \left\{x_1 + \frac{\Delta}{2}, x_2 + \frac{\Delta}{2}, \ldots, x_K + \frac{\Delta}{2}\right\},$$

respectively. Randomly assigning subjects within each pair to the left and right sequences, we have

$$\begin{aligned} & P\{\text{subject is treated at dose } x\} \\ = \; & \frac{1}{2}P\{\text{subject is treated at dose } x | \text{ left sequence}\} \\ & + \frac{1}{2}P\{\text{subject is treated at dose } x | \text{ right sequence}\}, \end{aligned}$$

and it follows that the combined stationary treatment distribution is a mixture of the left and right treatment distributions with mixing parameter $1/2$. The combined treatment space is

$$\Omega_C = \left\{x_1 - \frac{\Delta}{2}, x_2 - \frac{\Delta}{2}, \ldots, x_K - \frac{\Delta}{2}, x_K + \frac{\Delta}{2}\right\}.$$

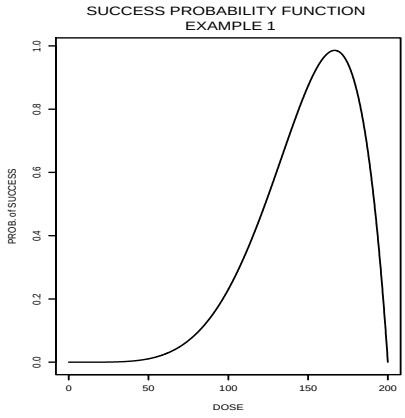


Figure 19.1 Example 1: success probability function

Note that Ω_L and Ω_R each have K treatments levels, whereas Ω_C has $K+1$ treatments levels. Examples of success probability functions are shown in Figures 19.1 and 19.3, and the stationary distributions that they generate are shown in Figures 19.2 and 19.4, respectively.

3. CONVERGENCE OF THE EMPIRICAL TREATMENT MODE OF THE STATIONARY TREATMENT DISTRIBUTION

In this section, we first define a modal set and unimodality on a lattice.
The set $M(\pi)$ of points of absolute maximum for π is called the *modal set.* The minimum point of $M(\pi)$ will be called the mode of π.
(Keilson and Gerber, 1971). A distribution $\pi = \{\pi_k\}$ with all support on the lattice of integers is said to be unimodal if there exists at least one integer M such that

$$\begin{cases} \pi_k \geq \pi_{k-1}, \text{ for all } k \leq M \\ \pi_{k+1} \leq \pi_k, \text{ for all } k \geq M. \end{cases}$$

Note if π is unimodal, then for all $m \in M(\pi)$, $x < m \leq y$ implies $\pi(x) \leq \pi(m) \geq \pi(y)$.

Let $\mathbb{k}$ be the largest integer such that $q_{\mathbb{k}-1} > p_{\mathbb{k}}$. Kpamegan and Flournoy (2000) have shown that the combined stationary treatment distribution is unimodal with (1) mode at $x_{\mathbb{k}} - \Delta/2$ if $\pi_{\mathbb{k}-1} > \pi_{\mathbb{k}+1}$ and at $x_{\mathbb{k}} + \Delta/2$ if $\pi_{\mathbb{k}-1} < \pi_{\mathbb{k}+1}$; (2) the modal set contains $\{x_{\mathbb{k}} - \Delta/2\,, x_{\mathbb{k}} + \Delta/2\}$ if $\pi_{\mathbb{k}-1} = \pi_{\mathbb{k}+1}$. They also showed that, in general, $-\frac{3}{2}\Delta \leq \mu - x_{\mathbb{k}} < \frac{\Delta}{2}$. Tighter bounds are obtained when $\alpha(x)$ is symmetric about its mode μ, in a neighborhood of μ. Then μ is within $\Delta/2$ of the asymptotic treatment $x_{\mathbb{k}}$, that is, $-\frac{\Delta}{2} < \mu - x_{\mathbb{k}} < \frac{\Delta}{2}$.

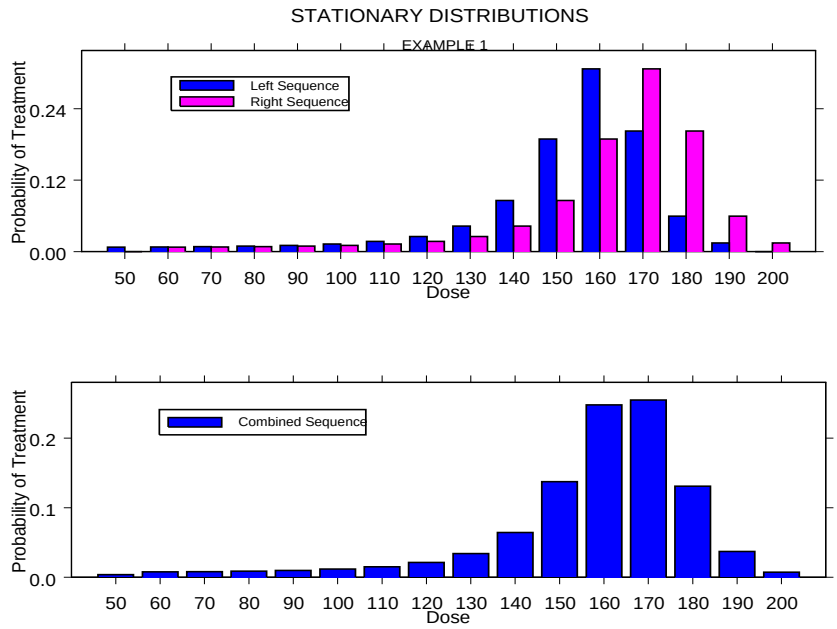


Figure 19.2 Example 1: stationary treatment distributions

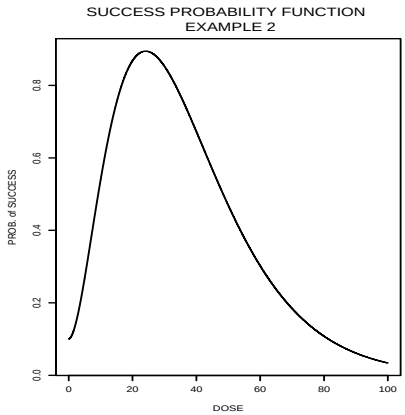


Figure 19.3 Example 2: success probability function

Indeed the mode of the combined treatment sequence is either $x_{\mathbb{k}} - \frac{\Delta}{2}$ or $x_{\mathbb{k}} + \frac{\Delta}{2}$ if unique, and if not unique, the model set includes both these treatments.

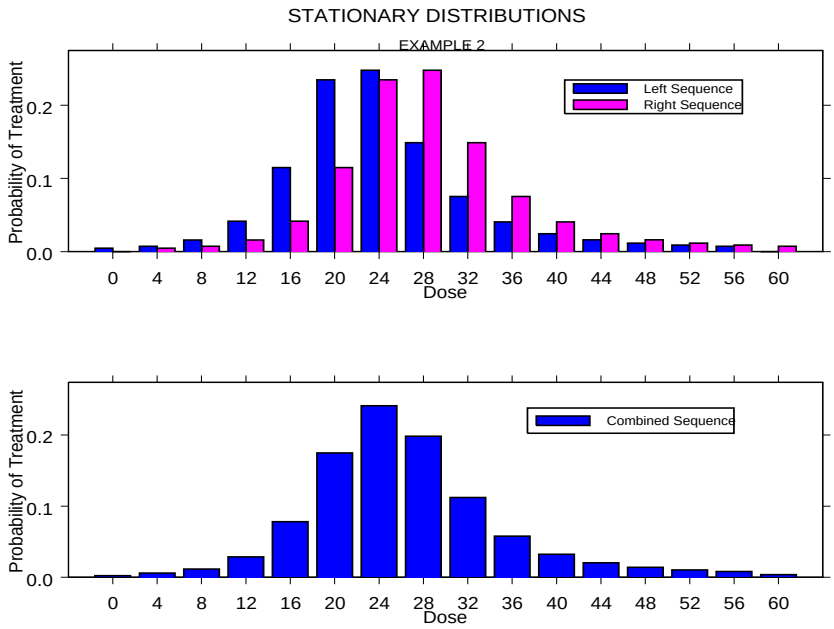


Figure 19.4 Example 2: stationary treatment distributions

In a couple of examples we examine the performance of the empirical mode of the treatment distribution as an estimator of the optimal dose. If the empirical treatment distribution has more than one element, the median of this modal set is taken to be the empirical mode. Figures 19.1 and 19.3 display examples of success probability functions. Figure 19.2 shows the stationary treatment distributions for the left and right sequences that are generated by the success probability function shown in Figure 19.1. The stationary treatment distribution for the left and right sequences combined also is given in Figure 19.2. Alternatively, Figure 19.4 shows the stationary treatment distributions when the success probability function is as shown in Figure 19.3.

4. COMPARISON OF THE UP-AND-DOWN AND STOCHASTIC APPROXIMATION METHODS

Kiefer and Wolfowitz (1952) defined a procedure for sequential estimation of the maximum of a regression function which has a maximum at the unknown point μ. This procedure has been widely researched. See, for example, Blum (1954), Chung (1954), Derman (1956) and Sacks (1957). Our motivation is its

application to binary responses as given by Durham, Flournoy and Li (1998). For this application, the regression function is $\alpha(x) = P(\text{success}|x)$and μ is the optimal dose. Let $\alpha(x)$ be strictly increasing for $x < \mu$ and strictly decreasing for $x > \mu$. Define infinite sequences of positive numbers $\{a_n\}$ and $\{c_n\}$ such that $c_n \to 0, \sum a_n = \infty, \sum a_n c_n < \infty, \sum a_n^2 c_n^{-2} < \infty$. Treat subjects in pairs at dose levels determined recursively by a_n and c_n.

If the nth pair of subjects has been treated at $X(n) - c_n$ and $X(n) + c_n$, the midpoint of the $(n+1)th$ pair is

$$X(n+1) = X(n) + \frac{a_n}{c_n}(Y_2(n) - Y_1(n)), \tag{19.6}$$

where $Y_1(n)$ and $Y_2(n)$ are Bernoulli random variables with success probabilities $\alpha(y|X(n)+c_n)$ and $\alpha(y|X(n)-c_n)$, respectively. Note that, if $a_n/c_n = \Delta$, we have the Optimizing Up-and-Down Design. Thus the Optimizing Up-and-Down Design is similar to stochastic approximation, but the treatment space is restricted to a discrete lattice and the changes in treatment levels from trial to trial are limited. The stochastic approximation estimator of the dose with maximum probability of response is defined recursively by $\widehat{\mu(n)} = X(n)$. Note also that while the sequence of treatments generated in stochastic approximation converges to a point, the sequence generated by the up-and-down procedure converges to a stationary distribution.

For estimating μ, we compare the mean squared error (MSE) of the mode of the empirical treatment distribution from the Optimizing Up-and-Down design with $X(n)$ from stochastic approximation. For stochastic approximation, we use $a_n = a/n$ and $c_n = c/n^{1/4}$, where a and c are positive constants. These sequences provide optimal convergence rates (c.f. Wasan, 1969). Also, we start the stochastic approximation procedure with $a_1/c_1 = \Delta$. For this comparison, we used the success probability function in Example 1. Note that, in this example, $\mu = 166.7$ and the probability of success is approximately zero for doses less than 50 and greater than 200. Each MSE was computed using 2,000 estimates of the mode of the success probability function. The MSEs are presented in side-by-side boxplots in Figures 19.5, 19.6 and 19.7 for comparison. The results suggest that the up-and-down procedure produces estimates with smaller MSE than does stochastic approximation.

4.1. SAMPLE SIZE EFFECT

Figure 19.5 displays boxplots of the MSEs computed with fixed step size 10, fixed initial dose 125 and sample sizes $20, 30, 50$ and 100. They clearly show that the up-and-down procedure performs better than stochastic approximation for all sample sizes and, the larger the sample size, the better it is.

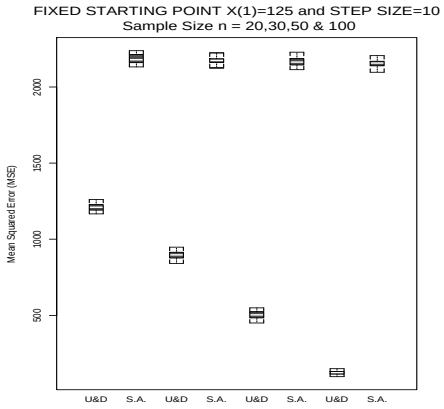


Figure 19.5 Sample size effect for Optimizing Up-and-Down and Stochastic Approximation Designs

4.2. STARTING POINT EFFECT

Figure 19.6 displays side-by-side boxplots of the MSEs for various initial treatments. The MSEs are large for the smallest starting point. This point is near the tail of the success probability function, where the curve is almost flat. This demonstrates that, if initial observations are made near the tail of the curve of the success probability function (see Figure 19.1), convergence to the mode will be very slow. The reason is that in the tail of the success probability curve, the difference between the responses will be close to zero. For initial values close to the mode, the MSEs are similar for the two procedures. In general, however, the Optimizing Up-and-Down procedure generates estimates with smaller MSEs.

4.3. STEP SIZE EFFECT

Starting at dose 120, the side-by-side boxplots in Figure 19.7 show that the Optimizing Up-and-Down procedure has more variability than stochastic approximation. For small step sizes $2, 4$ and 6, there is not much difference between median mean square errors for the two procedures. As the step size Δ increases to 8 and 10, the Optimizing Up-and-Down procedure performs better than stochastic approximation. Recall that all experiments started at the same dose. If the procedures start in the tail of the success probability function and the step size is very small, both converge extremely slowly. Therefore, if one observes a series of trials without success, some intervention should be taken. For example, Δ could be increased.

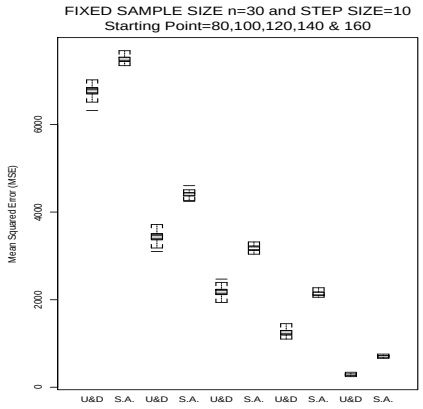


Figure 19.6 Starting point effect for Optimizing Up-and-Down and Stochastic Approximation Designs

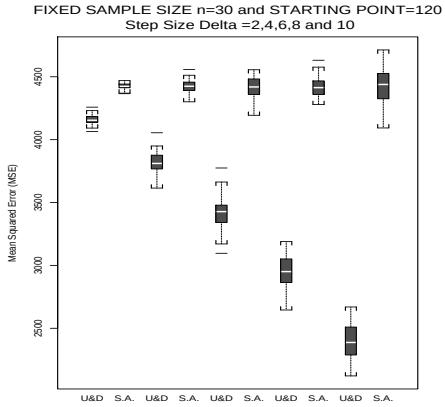


Figure 19.7 Step size effect for Optimizing Up-and-Down and Stochastic Approximation Designs

5. CONCLUDING COMMENTS

We have presented an Optimizing Up-and-Down Design that causes the distribution of treatments to cluster around the dose with maximum success probability. Theoretical results characterizing the asymptotic behaviour of this design were given. A brief analysis of the small sample behaviour of the treatment mode, as an estimator of the optimal dose, was based on simulation. Note, however, that because the design produces a non-homogenous random walk, exact moments can be calculated for the number of treatments and successes at each

dose, as was done for other Biased Coin Up-and-Down Designs by Durham, Flournoy and Haghighi (1995) and Flournoy, Durham and Rosenberger (1995).

There are several practical adjustments to the basic procedure that are recommended. First, consider the problem of starting the experiment far in the tails of the success probability function, in which case the basic procedure gets stuck there intolerably long. Suppose toxic failures occur if the dose is too high and the disease persists if the dose is too low. Then if one intentionally starts the experiment at a very low dose, for example to minimize toxicity while disregarding efficacy, we recommend beginning with a dose escalation scheme. In particular, increase the dose until the first success (no toxicity and no disease failure) is observed and then begin the Optimizing Up-and-Down Design.

Hardwick and Stout (1998) suggest a way to improve convergence of the empirical mode estimator, which will also help move quickly away from doses with low probability of success. Based on simulations, they suggest that when the response to treatment for both subjects in a pair is either success or failure, the next dose be increased or decreased according to the flip of an unbiased coin.

When one can distinguish between failure types, such as toxicity and persistent disease, it is desirable to modify the basic Optimizing Up-and-Down Design to take failure type into account. This is not so easy to do. An attractive idea proposed by Flournoy (1985) was to increase the dose if a toxic response was observed and decrease the dose if disease persisted. This idea was studied via simulations by Gooley *et al.* (1994). Unfortunately, Theorem 1 in Durham and Flournoy (1994), applied to this situation, proves that the stationary treatment distribution resulting from this procedure does not center treatments around the dose with maximum success probability. Indeed, the treatment distribution and the dose with maximum success probability can be quite distant from each other.

Thus a challenge is to find a modification of the basic design that takes failure type in account and yet still clusters treatments around the dose with maximum success probability. Yet convergence of the estimator is so fast, outperforming stochastic approximation, that the Optimizing Up-and-Down Design should be considered (via simulation for particular expected scenarios) for human experiments. For animal experiments, the benefits of this design over current Phase I/II strategies are clear: fast convergence of the estimator will yield the best dose with few animals. Larger group sizes can easily be accommodated.

References

Chung, K. L. (1954). On a stochastic approximation method. *Ann. Math. Statist.* **25**, 463–483.

Derman, C. (1956). Stochastic approximation. *Ann. Math. Statist.* **27**, 879–886.

Derman, C. (1957). Non-parametric up and down experimentation. *Ann. Math. Statist.* **28**, 795–799.

Dixon, W. J. and Mood, A. M. (1948). A method of obtaining and analyzing sensitivity data. *JASA* **43**, 109–126.

Dixon, W. J. (1965). The up-and-down method for small samples. *JASA***60**, 967–978.

Durham, S. D. and Flournoy, N. (1994). Random walks for quantiles estimation. In *Statistical Decision Theory and Related Topics V* Eds J. Berger and S. Gupta, pp. 467–476. New York: Springer-Verlag.

Durham, S. D. and Flournoy, N. (1995). Up-and-down designs I: stationary treatment distributions. In *Adaptive Designs*. IMS lecture Notes- Monograph Series **25**, pp. 139–157. Hayward CA: IMS.

Durham, S. D. and Flournoy N. and Montazer-Haghighi, A. (1995). Up-and-down designs II: exact treatment moments. In *Adaptive Designs*. IMS lecture Notes- Monograph Series **25**, pp. 158–178. Hayward CA: IMS.

Durham, S. D., Flournoy, N. and Li, W. (1998). A sequential design for maximizing the probability of a favorable response. *Canadian J. Statist.* **26**, 479–495.

Flournoy, N., Durham, S. D. and Rosenberger, W. F. (1995). Toxicity in sequential dose-response experiments. *Sequential Analysis* **14**, 217–228.

Giovagnoli, A. and Pintacuda, N. (1998). Properties of frequency distributions induced by general "up-and-down" methods for estimating quantiles. *JSPI* **25**, 51–63.

Gooley, T. A., Martin, P.J., Lloyd, D.F., Pettinger, M. (1994). Simulation as a design tool for Phase I/II clinical trials: An example from bone marrow transplantation. *Controlled Clinical Trials* **15,** 450-460.

Hardwick, J. P. and Stout, Q. (1998). *Personal communication.*

Harris, T. (1952). First passage and recurrence distributions. *Transactions of the American Mathematical Society* **73** , 471–486 .

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

Keilson, J. and Gerber, H. (1971). Some results for discrete unimodality. *JASA* **66**, 386–389.

Kiefer, J. and Wolfowitz, J. (1952). Stochastic approximation of the maximum of a regression function. *Ann. Math. Statist.* **25** , 529–532.

Kpamegan, E. E. and Flournoy, N. (2000). Up-and-down designs for selecting the dose with maximum success probability. Submitted.

Sacks, J. (1958). Asymptotic distribution of stochastic approximation procedures. *Ann. Math. Statist.* **29**, 373–405.

Tsutakawa, R. K. (1967). Random walk design in bio-assay. *JASA* **62**, 842–856.

Wasan, M. T. (1969). *Stochastic Approximation*. Cambridge: University Press.

Wetherwill, G. B. (1963). Sequential estimation of quantal response curves. *Journal of the Royal Statistical Society* B **25**, 1–48.

Discussion

Janis Hardwick

It may be of interest to consider additional performance measures. One goal of this procedure is to "limit the number of patients that are exposed to doses with high probabilities of failure". The authors note that exact moments can be calculated for the number of times a dose is sampled and how often it is successful. This work should be encouraged since such values would offer greater insights into how well the design is responding to the ethical challenge. Certainly the MSE of the modal dose estimator can be useful, but it also ignores important features of the response curve. In Example 1, the curve is left skewed. The mode is at 167 and the success rate at this value is about 0.98. We have seen that the MSE varies depending on the design parameters, but the figures suggest that a reasonable MSE for the up-down design is about 1200. This gives a standard error of approximately 35. If the estimator is 1 SE above the mode (i.e., above 200), then the success rate has gone from 0.98 to 0, whereas if it is 1 SE below the mode the rate is roughly 0.35. Thus, for this curve, it is far better to err to the left than to the right. Note also that if the curve were flat, the MSE could be quite large, but the procedure may still be highly successful since most doses would produce the same results. The reverse is true, of course, if the curve is peaked. One measure that avoids these problem is the *expected successes lost*, E[SL] = $\Sigma_{j=1}^{n/2}[2\alpha(\mu) - \mathbf{E}(Y_{1j} + Y_{2j})]$, where $\alpha(\mu)$ is the success rate at the mode, μ. With this measure or some function thereof, one compares how well a design performs against one of its main goals which is to treat all subjects at the dose with the highest success rate.

Since the step size parameter impacts the results rather significantly, a possible future direction for this work is to examine the same rule but allow for step size changes during the experiment. One way to do this is to specify the time at which a step size shift will take place. Another approach is to make the shift adaptively, perhaps after certain data patterns have occurred.

Other approaches have been taken to Flournoy's problem. Some of these take greater advantage of the unimodal shape constraint on the response curve. In the paper by Hardwick and Stout in this volume, the problem is addressed using a least squares curve fit and various sampling algorithms. The best of these, a multi-arm bandit strategy, which is designed to minimize patient failures improves upon the optimizing up and down design on measures relating to the terminal decision as well as to E[SL]. A disadvantage of the bandit/curve fitting method is that the theory is complex. For example, while convergence

is established, *rates* of convergence are unknown. Furthermore, the exact moments of operating characteristics such as proportion sampled per dose are not available. Finally, the sampling algorithm is far less intuitive and a computer is required to generate it.

Kpamagen and Flournoy mention the problem in which the two failure modes are distinguishable. Hardwick, Meyer and Stout (2000) have recently submitted a paper on this topic. The paper examines parametric and nonparametric curve fitting techniques for estimating the individual toxicity and efficacy response curves.

One feature impacting all of the designs mentioned relates to the shape of the "true" underlying response functions. For flat curves, the sampling design is far less important than it is for skewed and/or peaked functions. Useful research projects would include attempts:

- to narrow down the class of response functions for a specific problem – perhaps using prior elicitation methods;
- to define a prior on the space of unimodal functions and conduct sampling with respect to this space.

Finally, it would interesting to examine bivariate versions of the present problem. For example there may an added factor, such as level of disease or age, controlling the unimodal response function generated from the toxicity and efficacy curves. In this case one would need to consider the type of unimodality assumed as well as the definition of "neighbour" states and allowable moves.

Additional Reference

Hardwick, J., Meyers, M. and Stout, Q.F. (2000). Directed walks for dose response problems with competing failure modes. (Submitted).

Chapter 20

FURTHER RESULTS ON OPTIMAL AND EFFICIENT DESIGNS FOR CONSTRAINED MIXTURE EXPERIMENTS

R.J. Martin
Department of Probability and Statistics, University of Sheffield
Sheffield, S3 7RH, UK
r.j.martin@sheffield.ac.uk

M.C. Bursnall
Department of Statistics, University of Leeds
Leeds, LS2 9JT, UK
matt@amsta.leeds.ac.uk

E.C. Stillman
Department of Probability and Statistics, University of Sheffield
Sheffield, S3 7RH, UK
e.c.stillman@sheffield.ac.uk

Abstract Many practical experiments on mixtures (where the components sum to one) include additional lower or upper bounds on components. This paper gives some additional results on the form of the continuous D- and V-optimal designs for 3 and 4 components.

Keywords: mixture experiments, optimal design

Introduction

In many practical mixture experiments, components, as well as summing to one, have tighter lower or upper bounds. Martin *et al.* (1999) (MBS henceforth) discussed some D- and V-optimality results for 3-component mixtures with a quadratic model and obtained numerically the form of the continuous D-

optimal design for several different situations. Additional background material on constrained mixture experiments and further references are given in MBS.

In this paper, we discuss several extensions, for quadratic models, of the results in MBS. We give two general results, one on optimal designs for the general parallelepiped and one on vertex-centroid designs when equal apices are removed. We derive a condition for a weight to be constant as a constraint changes. For 3-component mixtures, we consider some further examples of continuous D-optimal designs and give a more complete solution for the hexagon case. We also show in the unconstrained case that the $\{q, 2\}$-simplex lattice design is not V-optimal and consider continuous V-optimal designs in other cases. For 4-component mixtures, we consider continuous D-optimal designs, when imposing constraints on one, two, three or four components.

1. PRELIMINARIES

Let the q components have proportions $x_1, x_2, ..., x_q$. The x_j lie in a **regular (unit-) simplex**, being constrained by $0 = a_j \leq x_j \leq b_j = 1 \; \forall j$; and $\sum_{j=1}^{q} x_j = 1$. There are often tighter **lower** and/or **upper bounds** $a_j > 0, \;\; b_j < 1$, on x_j, giving the **constrained region** R. We do not consider here the more general multicomponent constraints.

We can simplify the constraints by assuming the (lower) L-transform $x_j \rightarrow (x_j - a_j) \,/\, (1 - \sum a_j)$ has been used, so we can take $a_j = 0 \; \forall j$. If originally $1 < \sum b_k < 2$, we can also assume the (upper) U-transform, $x_j \rightarrow (b_j - x_j) \,/\, (\sum b_k - 1)$, has been used first.

We assume a quadratic model for the mean response $E(y_i)$, which is often a reasonable assumption at the design stage. A **standard** quadratic surface can be formed from (any) $q - 1$ components, but we mainly use here barycentric coordinates and the q-component **canonical** (or Scheffé) **quadratic** model:

$$E\left(y_i\right) = \sum_{j=1}^{q} \beta_j x_{ji} + \sum_{j<k=1}^{q} \beta_{jk} x_{ji} x_{ki} \;\text{ for }\; i = 1, \ldots, n.$$

Let $p = q(q+1)/2$ be the number of model parameters. Let **point** i be $(x_{1i}, \ldots, x_{qi})'$ and let x_i be its augmented p-vector, including all cross-products. Let X be the $n \times p$ regression (design) matrix, with i^{th} row x_i' and assume rank$(X) = p$. Assuming $var\left\{(y_1, \ldots, y_n)'\right\} = I_n \sigma^2$, the (scaled) **prediction variance** at a general x is $v_x = x'(X'X)^{-1}x$.

If n is not small, insight into useful design features can be obtained from the **continuous** design, for which there are s **support points** (s.p.) with non-zero weight w_i and $\sum w_i = 1$. The finite design with the same s.p. and with replicate numbers n_i equal to the rounded nw_i will usually be quite efficient. If n is small and (some of) the nw_i are not close to integers, the finite optimal design may have different design points from these s.p.

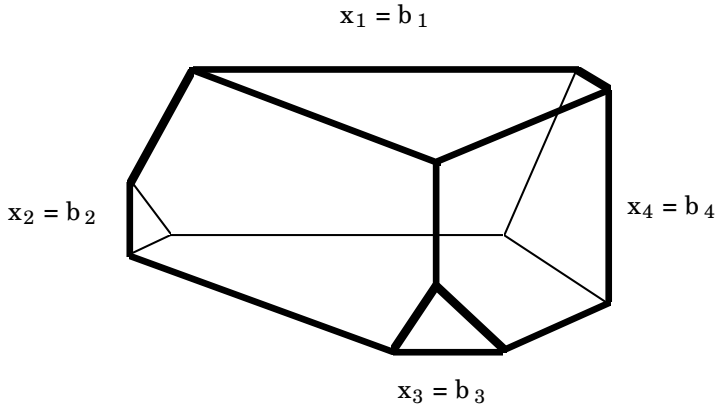


Figure 20.1 The $q = 4$ simplex with the four apices above $x_i = b_i,\ 1 \leq i \leq 4$, removed, where $b_1 + b_4 < 1$ and $b_i + b_j > 1$ otherwise

Let $W = \text{diag}\{w_1, \ldots, w_s\}$. The D-**criterion** here maximizes the D-value $|X'WX|^{1/p}$. The G-criterion, which minimizes the G-value, the maximum of $v_x = x'(X'WX)^{-1}x$ over R, gives the same optimal designs (see, e.g., Atkinson & Donev, 1992, §9.2). The G-value is at least p. The V- **(or I-) criterion**, minimizes the V-value, the average of v_x over R, which is $tr\left\{(X'WX)^{-1}M\right\}$, where $M = \left(\int_R xx'dx\right) / \left(\int_R dx\right)$ is the **moment matrix**.

The $\{q, 2\}$-**weighted simplex lattice design** for $b_j = 1\ \forall j$ uses equally weighted vertices (w_v) and equally weighted mid-edges (w_m). The $\{q, 2\}$-**weighted simplex centroid** design includes the weighted overall centroid (w_c). We define a **vertex- centroid** design as one whose support points are the vertices and the centroid.

In Sections 3 to 5, we consider removing $t = 1, \ldots, q$ apices (simplices) from the unit simplex, i.e. the constraints $x_j \leq b_j < 1,\ j \leq t$. Fig. 20.1 shows an example when $t = q = 4$.

Unless otherwise stated, we assume the **equal-apices** case: $b_j = b$ for $j \leq t$ and 1 for $j > t$. Using the U-transform, the lower limit of b is $1 - (q - 2)/t$ for $t > q - 2$. We obtain continuous D- or V-optimal designs over designs containing the commonly-used candidate points: vertices, mid-edges, face centroids and the overall centroid. We mainly use the average vertex (a.v.) centroid, which may differ from the centre of mass (c.o.m.) centroid. We have used matlab (Math Works, 1992) to obtain the weights to 4 decimal places and, for the D-criterion, the obtained G-value over the candidate points is at most $p/0.9999$. For a polynomial model, formulae for M can be obtained when simplices are removed. The weights usually depend on the value of b; we investigate how they and the s.p. vary as b decreases from 1. Note that in MBS

we investigated support points over R. Small changes to the edge points or centroids may improve the design efficiency.

The regular simplex has q vertices, $q(q-1)/2$ edges and is the limiting region in all equal-apices cases as $b \to 1$. The $\{q, 2\}$-weighted simplex lattice design with $w_v = w_m = 1/p$ is D-optimal for a quadratic surface. The parallelepiped (see §2) is often a limit in other cases. In sections 3 to 5, weights of limiting designs are, where appropriate, divided between the new vertices which arise when b is close to the limit. In the symmetric cases considered, $w_v,\ w_m,\ w_c$ denote the weights of the vertices, mid-edges and the a.v. centroid.

2. SOME GENERAL RESULTS

In this section we show why a weight can be constant as b varies and give two results for general q. We begin with two technical lemmas. Sketch proofs are given in the Appendix.

Lemma 1. Suppose $X'WX$ is $p \times p$ with rank $r < p$, but $X'WX + uZ'Z$ has rank p. Then $|X'WX + uZ'Z|$ has a factor u^{p-r}. If, additionally, $Z = [Z_1\ Z_2]$ is $(p-r) \times p$ with Z_2 square, then $|X'WX + uZ'Z|$ is $(uc^2_{X,Z})^{p-r} \times f_X(w)$, where $f_X(w)$ is the product of the positive eigenvalues of $X'WX = \sum w_i x_i x_i'$, and $c_{X,Z}$ does not depend on u and the $\{w_i\}$. If also $X = [X_1\ 0]$ for an r-column X_1, then $c_{X,Z} = |Z_2|$.

Lemma 2. Consider minimization of the separable function $f(x)g(u)$ for a vector x and a scalar u, subject to $d_u u + \sum d_j x_j = d_0$. If $f(x) = \sum c_i \left\{ \Pi_j x_j^{a_{ji}} \right\}$ is a homogeneous polynomial of degree $a,\ a_{ji} \neq 0\ \forall i, j$ and $\sum_j a_{ji} = a\ \forall i$, then at the minimum u satisfies $ad_0 g(u) = g'(u)\,(d_0 - d_u u)$. If $g(u) = u^h$ for $h \neq 0$, then at the minimum $u = hd_0 / \{d_u(h+a)\}$.

Result 1. Constant Weight as b Varies. If Lemma 1 holds for a design with design matrix $[X'Z']'$ and weight matrix $diag\,(W, uI_{p-r})$, where the w_i are unrelated to u, then Lemma 2 holds in the special case $d_0 = 1,\ \ d_u = h = p - r,\ \ a = r$, so that the D-optimal design has $u = 1/p$ (and hence $\sum w_i = r/p$). If this is true for b in some interval, then the $p - r$ Z-points have constant weight $1/p$ in that interval in the D-optimal designs. Note that provided Lemma 1 holds, the D-optimal weights do not depend on Z.

Result 2. Optimal Designs on a Parallelepiped. This generalizes the results in §3.4 of MBS on the $q = 3$ parallelogram. Suppose that $b_q = 1$ and $\sum_{j=1}^{q-1} b_j = 1$. Then, whatever the individual $b_j,\ j = 1, \ldots,\ q-1,\ R$ is a parallelepiped in the simplex and a $(q-1)$-cuboid in the standard space.

Hence, known results for optimal response surfaces on the $(q-1)$-dimensional cuboid can be used. In particular, D-optimal designs on the cube (and hence on the general cuboid) are known for low- order polynomial surfaces. Atkinson & Donev (1992, Table 11.5) give the weights for a D-optimal design for a quadratic surface for $q = 3, 4, 5, 6$. For example, when $q = 4$ the D-optimal design for a quadratic surface has weights $0.063775,\ 0.03535,\ 0.06560$ for the 8 vertices, 12 mid-edges and centroid, respectively. Although the constraints leading to a parallelepiped may be rare in practice, they may arise as an approximation in limiting cases - see MBS and §§3,4.

Result 3. Vertex-Centroid Designs for $t = q$. Consider the equal-apices case when $t = q,\ 2/q < b < 1$. For $q = 3$ this gives the hexagon symmetric under a third turn (MBS, §3.2). The vertex-centroid design has $q(q-1)$ vertices. Some algebra then shows that $|X'WX| = c_b w_v^{p-1} w_c$, where

$$c_b = q^{q-4} 2^{p-2q-1} \left\{b(1-b)\right\}^{2(p-q-1)} (2b-1)^{2(q-1)} \left\{2qb(1-b) - (q-1)\right\}^2 .$$

It therefore follows that, whatever b, the D-optimal design in this class has $w_v = (p-1)/\left\{pq(q-1)\right\},\ w_c = 1/p$. This is a special case of Result 1, which shows $(p - r = 1)$ that the D-optimal $t = q$ vertex-centroid design has $w_c = 1/p$ in the general case of unequal apices. The D-optimal weights are unchanged if the centroid is replaced by any interior point, although $|X'WX|$ will differ.

3. ADDITIONAL RESULTS FOR 3-COMPONENT *D*-OPTIMAL DESIGNS

Consider D-optimal designs when $q = 3$ $(p = 6)$, firstly for the trapezium $(t = 1)$. If $b \geq 0.66$, the only non-zero weights are for the 4 vertices, the two mid-edges with $x_1 = b/2$ and the $x_1 = 0$ mid-edge. The first 6 points give rank$(X) = 5$ since they have $x_2 \times x_3 = 0$, while all 7 have rank$(X) = 6$. Thus, by Result 1 $(p - r = 1)$, the weight for the $x_1 = 0$ mid-edge is 1/6. For any $b \geq 0.66$, the D-optimal weights are the same if this mid-edge is changed to any interior point of the $x_1 = 0$ edge, but $|X'WX|$ is maximised by using the mid-edge.

Now consider hexagons $(t = 3)$. By Result 3, vertex-centroid designs have $w_c = 1/6$. It appears (see below) that this design is D-optimal over designs including mid-edges when the b_j are "close" to $2/3$.

The equal-apices case, $2/3 \leq b \leq 1$, was considered in MBS (§3.2). We add two further details here. Firstly, the weight of the centroid becomes zero for $b \geq 0.89$. Secondly, Result 3, with the numerical result that $w_m = 0$ for b in the interval $(2/3, 0.74661)$ shows that the D-optimal design for $b = 2/3$

b	$v1$ (4)	$v2$ (2)	$me1$ (4)	$me2$ (2)	avc (1)
1	146	80	0	80	96
0.9	143	92	0	61	120
0.8	145	114	0	17	157
0.75	146	124	0	0	167
0.7	142	132	0	0	167
2/3	139	139	0	0	167
0.6	132	153	0	0	167
0.55	100	146	42	0	141
0.53	83	143	65	0	121
0.5	73	146	80	0	96

Table 20.1 Weights ($\times 10^3$) for the support points (numbers of each in brackets) of the rht-hexagon as b varies: $v1$ & $v2$ are vertices; $me1$ & $me2$ mid-edges and avc the centroid

(vertex-centroid design), with $w_v = 5/36,\ w_c = 1/6$, is also D-optimal for b up to 0.74661.

Another one-parameter hexagon, the **rht-hexagon**, has $b_1 = b_2 = b$ and $b_3 = 2(1-b),\ 1/2 < b < 1$. It is reflection symmetric about $x_1 = x_2$ and symmetric under a half-turn. The limits when $b = 1$ and $b = 1/2$ are both parallelograms (MBS, §3.4; Result 2) and $b = 2/3$ gives the regular hexagon. The case $b = 12/13$ was considered in MBS. For some b between 1 and $1/2$, the weights are given in Table 20.1 for: $v1$: the four $x_3 = 2(1-b)$ and $x_3 = 0$ vertices; $v2$: the two $x_3 = 1-b$ vertices; $me1$: the four $v1$ to $v2$ mid-edges with $x_3 = 3(1-b)/2$ or $x_3 = (1-b)/2$; $me2$: the two $v1$ to $v1$ mid-edges with $x_1 = x_2$; avc: the centroid.

The vertex-centroid design (with $w_c = 1/6$) is D-optimal for b between 0.572 and 0.777, when the weight of the $v1$ vertices is given by $192(3b-1)(b-1)w^2 - 32(2b^2-4b+1)w - 5b^2 = 0$, and varies between 0.1306 and 0.1482. The $v1$ weight falls, apart from a slight rise around $b = 0.75$, as b decreases, while the reverse holds for $v2$. For b outside $(0.572,\ 0.777)$ the $me1$ (low b) or $me2$ (high b) weights become non-zero.

Some other hexagons where the vertex-centroid design (with $w_c = 1/6$) is D-optimal are: $b_1 = b_2 = b,\ b_3 = 0.7,\ 0.59 \le b \le 0.74$; and $b_1 = 0.8, b_2 = 0.7,\ b_3 = 0.6$.

4. RESULTS FOR 4-COMPONENT D-OPTIMAL DESIGNS

We consider for $q = 4$ ($p = 10$) the cases of $t = 1, 2, 3, 4$ apices (regular simplices) removed in sections 4.1, 4.2, 4.3, 4.4, respectively. Note that a major

change in the shape of R occurs at $b = 1/2$ when $t = 2$ or 3 (see Fig. 20.1 for the general cases $b_i + b_j < 1$ and $b_i + b_j > 1$).

In this section, the following notation is used for the support points, where $S_1 = \{1, ..., t\}$ and, for $t < 4,\ S_2 = \{t + 1, \ldots, 4\}$:

$v1$: $4 - t$ intact vertices $x_i = 1\ i \in S_2(t < 4)$;
$v2$: $t(4 - t)$ vertices with $x_i = b,\ i \in S_1,\ x_j = 1 - b,\ j \in S_2(t < 4)$;
$v3$: $t(t - 1)$ vertices with $x_i = b, x_j = 1 - b,\ i, j \in S_1(t > 1,\ b \geq 1/2)$;
$v4$: $t(t - 1)$ vertices with $x_i = x_j = b,\ i, j \in S_1(t = 2,\ b \leq 1/2)$;
$v5$: $t(t - 1)/2$ vertices with $x_i = x_j = b,\ i, j \in S_1,\ x_k = 0,\ k \in S_2,\ (t = 3,\ b \leq 1/2)$;
$v6$: $t(t - 1)/2$ vertices with $x_i = x_j = b,\ x_k = 0,\ i, j, k \in S_2,\ (t = 3,\ b \leq 1/2)$;
$me1$: mid-edges between $v1$ vertices with $x_i = x_j = 1/2,\ i, j \in S_2\ (t < 3)$;
$me2$: mid-edges between $v1$ and $v2$ vertices with $x_i = b/2,\ i \in S_1,\ x_j = 1 - b/2, j \in S_2(t < 4)$;
$me3$: mid-edges between $v2$ vertices with $x_i = b,\ i \in S_1,\ x_j = x_k = (1 - b)/2,\ j, k \in S_2(t < 3)$;
$me4$: mid-edges between $v2$ and $v3$ vertices with $x_i = b,\ x_j = (1-b)/2,\ i, j \in S_1,\ x_k = (1 - b)/2,\ k \in S_2(t = 2, 3,\ b \geq 1/2)$;
$me5$: mid-edges between $v2$ and $v4$ or $v6$ vertices with $x_i = b,\ x_j = b/2,\ i, j \in S_1,\ x_k = 1 - 3b/2,\ k \in S_2(t = 2, 3,\ b \leq 1/2)$;
$me6$: mid-edges between $v3$ vertices along original simplex edges with $x_i = x_j = 1/2,\ i, j \in S_1(t > 1,\ b \geq 1/2)$;
$me7$: mid-edges between $v4$ vertices with $x_i = x_j = b,\ i, j \in S_1,\ x_k = x_l = 1/2 - b, k, l \in S_2(t = 2,\ b \leq 1/2)$;
$me8$: mid-edges between $v5$ and $v6$ vertices with $x_i = x_j = b,\ x_k = 1/2 - b,\ i, j, k \in S_1\ (t = 3, b \leq 1/2)$;
$me9$: mid-edges between $v3$ vertices along $x_i = b$ edges, or between $v5$ vertices $(b \leq 1/2)$, with $x_i = b,\ x_j = x_k = (1 - b)/2,\ i, j, k \in S_1\ (t > 2)$;
$fc1$: face a.v. centroid on the $x_i = 0,\ i \in S_2$ faces $(t < 4)$;
$fc2$: face a.v. centroid on the $x_i = 0,\ i \in S_1$ faces;
$fc3$: face a.v. centroid on the $x_i = b,\ i \in S_1$ faces;
avc: a.v. centroid (equal to the c.o.m. centroid when $t = 4$).

The ratios of the 9 edge lengths (when positive) are $1\ :\ b\ :\ 1 - b\ :\ 1 - b\ :\ b\ :\ 2b - 1\ :\ 1 - 2b\ :\ 1 - 2b\ :\ 1 - b$ (between $v3$ vertices) or $3b - 1$ (between $v5$ vertices). Except for the trapezium and the pentagon, the a.v. face centroid is equal to the c.o.m. face centroid.

b	$v1$ (3)	$v2$ (3)	$me1$ (3)	$me2$ (3)	$me3$ (3)	$fc1$ (3)
1	100	33	100	100	0	0
0.8	99	47	100	87	0	0
0.6	97	76	100	60	0	0
0.4	95	90	91	31	0	25
0.2	92	91	78	24	6	42
0	85	85	47	25	47	46

Table 20.2 Weights ($\times 10^3$) for the support points (numbers of each in brackets) when one apex is removed as b varies: $v1$ & $v2$ are vertices; $me1$, $me2$ & $me3$ mid-edges and $fc1$ a face centroid

4.1. ONE APEX REMOVED

If $t = 1,\ \ R$ is symmetric in x_2, x_3 and x_4 (under rotations of a third-turn about $x_2 = x_3 = x_4$, or reflections about $x_j = x_k, j, k > 1$). Then R has 6 vertices, 9 edges and 5 faces (2 non-congruent equilateral triangles, $x_1 = 0, x_1 = b$; 3 congruent trapezia, $x_i = 0,\ i > 1$). For b close to $0,\ R$, in standard coordinates, is close to a triangular wedge. Table 20.2 gives the weights for the s.p. as b decreases from 1 (simplex) to 0 (triangular wedge). For $b \geq 0.60$, the only non-zero weights are for the 6 vertices, the 3 $me2$ mid-edges and the 3 $me1$ mid-edges. The first 9 points give rank$(X) = 7$ since they have $x_i \times x_4 = 0$ for $i < 4$, while all 12 have rank$(X) = 10$ and Result 1 holds $(p - r = 3,\ c_{X,Z} = 1/4)$. Thus the weight for the $me1$ mid-edges is $1/10$. The weight for $fc1$ starts to become non-zero around $b = 0.59$ and, that for the $me3$ mid-edges, around $b_1 = 0.22$. The face centroids $fc2,\ \ fc3$ and the a.v. centroid appear to have zero weight throughout. The number of s.p. varies from 10 to 12 to 15 to 18 as b decreases.

4.2. TWO EQUAL APICES REMOVED

If $t = 2,\ \ R$ is reflection symmetric about both $x_1 = x_2$ and $x_3 = x_4$. Then, for $b \neq 1/2,\ \ R$ has 8 vertices, 12 edges and 6 faces. For $1/2 < b < 1$, the faces comprise 2 congruent equilateral triangles $(x_i = b,\ i = 1, 2)$, 2 congruent trapezia $(x_i = 0,\ i = 1, 2)$ and 2 congruent reflection symmetric pentagons $(x_i = 0,\ i = 3, 4)$. For $0 < b < 1/2$, they are 2 sets of 2 congruent trapezia $(x_i = 0,\ i = 1, 2$; and $x_i = b,\ i = 1, 2)$ and 2 congruent rhombi $(x_i = 0, i = 3, 4)$. When $b = 1/2$, there are 7 vertices, 11 edges and 6 faces (2 congruent equilateral triangles $(x_i = b,\ i = 1, 2)$, 2 congruent trapezia $(x_i = 0,\ i = 1, 2)$ and 2 congruent rhombi $(x_i = 0,\ i = 3, 4)$. For b close to $0,\ R$ is close to a parallelepiped (see §2). Table 20.3 shows how the weights for the s.p. change as b decreases from 1 (simplex) to 0 (parallelepiped). Similarly

b	$v1$	$v2$	$v3/$ $v4$	$me1$	$me2$	$me3$	$me4/$ $me5$	$me6/$ $me7$	$fc1$	$fc2$	avc
	(2)	(4)	(2)	(1)	(4)	(2)	(4)	(1)	(2)	(2)	(1)
1	100	38	23	100	100	0	0	100	0	0	0
0.8	98	51	38	100	87	0	0	74	0	0	0
0.6	93	78	71	100	49	0	0	0	32	0	0
0.5	91	81	43	96	36	0	21	0	34	7	0
0.4	90	81	47	81	24	0	27	0	32	27	0
0.2	84	76	65	49	14	4	29	0	32	53	0
0	64	64	64	35	35	35	35	35	0	0	66

Table 20.3 Weights ($\times 10^3$) for the support points (numbers of each in brackets) when two apices are removed as b varies: $v1$, $v2$, $v3$ & $v4$ are vertices; $me1$ to $me7$ mid-edges and avc the centroid

to $t = 1$ (§4.1), Result 1 $(p - r = 1)$ holds for $b \geq 0.56$. This case is more complicated as, apart from $me1$, there are, as b varies in [0.56,1), four different sets of points which have non-zero weights. However, all these points (vertices, $me2$, $me4$, $me6$, $fc1$) and their relevant subsets give rank$(X) = 9$ since they have $x_3 \times x_4 = 0$, while including $me1$ gives rank$(X) = 10$. Thus the weight for the $me1$ mid-edge is $1/10$. Changes in the weights for the $v2$ and $v3/v4$ vertices as b decreases reflect the change in the shape of R at $b = 1/2$. The midpoints $me3$ and $me4/me5$, and the a.v. centroid become weighted at around $b = 0.2,\ 0.55$ and 0.15, respectively. The face centroids $fc1,\ \ fc2$ become weighted at around $b = 0.75$ and 0.5, respectively, and their weights decrease to 0 at $b = 0$, while $fc3$ has zero weight throughout. The number of s.p. varies from 10 $(b = 1)$ through 14 up to 21 (around $b = 0.5$) and 25 (b around 0.1), tending to 21 $(b = 0)$.

4.3. THREE EQUAL APICES REMOVED

If $t = 3,\ b \geq 1/3,\ R$ is symmetric under a third turn about $x_1 = x_2 = x_3$ and reflections about $x_i = x_j,\ i, j < 4$. When $b = 1/3,\ R$ is a parallelepiped (see §2), with 8 vertices, 12 edges and 6 faces $(x_i = b,\ x_i = 0, i < 4)$. Then, for $b > 1/3,\ b \neq 1/2,\ R$ has 10 vertices, 15 edges and 7 faces. There are major shape changes at $b = 1/2$ and $b = 1/3$. For $1/2 < b < 1$, the faces comprise 3 congruent equilateral triangles $(x_i = b,\ i < 4)$, 3 congruent reflection symmetric pentagons $(x_i = 0,\ i < 4)$ and 1 hexagon symmetric under a third turn $(x_4 = 0)$. For $1/3 < b < 1/2$, they are 1 equilateral triangle $(x_4 = 0)$, 3 congruent rhombi $(x_i = 0,\ i < 4)$ and 3 congruent reflection symmetric pentagons $(x_i = b,\ i < 4)$. When $b = 1/2$, there are 7 vertices, 12 edges and 7 faces (4 congruent equilateral triangles $(x_i = b,\ i < 4,\ x_4 = 0)$, 3 congruent rhombi $(x_i = 0,\ \ i < 4)$). Note that the a.v. centroid tends to $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ as

b	$v1$	$v2$	$v3/$ $v4$	$v3/$ $v6$	$me2$	$me4/$ $me5$	$me6/$ $me8$	$me9$	$fc1$	$fc2$	avc
	(1)	(3)	(3)	(3)	(3)	(6)	(3)	(3)	(1)	(3)	(1)
1	100	40	30	30	100	0	100	0	0	0	0
0.8	98	53	44	44	87	0	67	0	18	0	0
0.6	87	82	74	74	49	0	0	0	0	0	79
0.5	96	84	43	43	7	8	0	57	0	51	0
0.4	94	78	66	63	0	19	0	0	0	55	8
1/3	88	76	17	65	0	21	47	0	0	55	4

Table 20.4 Weights ($\times 10^3$) for the support points (numbers of each in brackets) when three apices are removed as b varies: $v1$ to $v6$ are vertices; $me2$ to $me9$ mid-edges; $fc1$ & $fc2$ face centroids and avc the centroid

b tends to $1/2$ from above or below, but is $(\frac{3}{14}, \frac{3}{14}, \frac{3}{14}, \frac{5}{14})$ at $b = 1/2$. Similarly, it tends to $(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{2}{5})$ as b tends from above to $1/3$, but is $(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{2})$ at $b = 1/3$. This means that although the geometric limit as b tends to $1/3$ is the parallelepiped, the limiting weights are not those of the parallelepiped, and the limiting design is slightly less efficient (D-efficiency 0.9944). Although the limiting weights look quite different from those of the parallelepiped (see §2), using the c.o.m. centroid gives a very similar set of limiting weights with the same D-value as for the symmetrically weighted parallelepiped. Table 20.4 shows how the weights for the s.p. change as b decreases from 1 (simplex) to $1/3$ (parallelepiped). Apart from edges getting longer or shorter, changes in the weights as b decreases reflect the shape and definition changes at $b = 1/2,\ 1/3$. Some weights are only non-zero in small intervals - the mid-edge $me9$ in the region of $b = 0.5$ and the face centroid $fc1$ in $[0.65, 0.82]$. The a.v. centroid becomes moderately highly weighted around $b = 0.6$. The number of s.p. varies from 10 (simplex) through 15 up to 25 (b around 0.51), and down to 21 ($b = 1/3$).

4.4. FOUR EQUAL APICES REMOVED

If $t = 4,\ 1/2 \leq b < 1,\ R$ is an octahedron symmetric in the x_i, under a third turn about each of $x_i = x_j = x_k$ and reflections about $x_i = x_j$. For $1/2 < b < 1,\ R$ has 12 vertices, 18 edges and 8 faces (4 congruent equilateral triangles $(x_i = b,\ \forall i)$, 4 congruent hexagons $(x_i = 0,\ \forall i)$). The 12 vertices are equivalent and the edges are of two types: 6 between $x_i = b$ faces and 12 on them, with length ratio $(2b-1)/(1-b)$. The lengths are equal when $b = 2/3$ and the 6 edges are longer for $b > 2/3$. When $b = 1/2,\ R$ is the regular octahedron, with 6 vertices, 12 edges and 8 faces (8 congruent equilateral triangles).

Table 20.5 shows how the weights for the s.p. change as b decreases from 1 (simplex) to $1/2$ (regular octahedron). For $0.5984 < b < 0.7119$ only w_v

b	$v3$ (12)	$me6$ (6)	$me9$ (12)	avc (1)
1	33	100	0	0
0.9	37	93	0	0
0.8	47	56	0	45
0.75	59	35	0	79
0.72	71	9	0	96
0.7	75	0	0	100
0.6	75	0	0	100
0.59	68	0	7	92
0.57	56	0	21	72
0.54	45	0	35	42
0.52	41	0	41	20
0.5	40	0	43	1

Table 20.5 Weights ($\times 10^3$) for the support points (numbers of each in brackets) when four apices are removed as b varies: $v3$ is a vertex; $me6$ & $me9$ mid-edges and avc the centroid

and w_c are non-zero, so that then, by Result 3, the vertex-centroid design has $w_c = 1/10, \;\; w_v = 3/40$. As b decreases from 1 to $1/2, \; w_v$ rises from 0.033 to 0.075 at $b = 0.7119$ and then from $b = 0.5984$ drops to 0.040, whilst w_c rises from 0 to 0.100 and then drops back to nearly 0. The weight of the $me6$ mid-edges drops from 0.100 to 0 at $b = 0.712$ and the weight of the $me9$ mid-edges rises from 0 at $b = 0.598$ to 0.043. The face centroids $fc2$, $fc3$ have zero weight throughout. The number of s.p. varies from 10 to 18 to 19 to 13 to 25 to 19 as b decreases.

5. V-OPTIMAL DESIGNS

Consider first the regular simplex. Laake (1975) showed that the V-optimal weights for the $\{q, 2\}$-weighted simplex lattice design satisfy

$$w_v/w_m = \left\{(q^2 - 7q + 18)/32\right\}^{1/2}.$$

Note that this ratio, 1 for the D-optimal design, varies with q, rising from 0.433 at $q = 3$, to over 1 when $q = 9$, and for large q is approximately $q/5.66$. However, for the $\{q, 2\}$-weighted simplex centroid design, some algebra shows that $q(q+1)(q+2)(q+3)$ times the V-value is

$$\begin{aligned} & \left[2q^4 w_v w_m \left\{16(q-1)w_v + (q^2 - 7q + 18)w_m\right\}\right. \\ + \;& w_c \left\{64(q-1)(q-2)(3q-1)w_v^2\right. \\ + \;& 32(q-1)(q^3 - 7q^2 + 19q - 10)w_v w_m \\ + \;& \left.\left.(q-2)^2(q^3 - 12q^2 + 49q - 38)w_m^2\right\}\right] / \\ & \left\{w_v w_m (q^3 w_v w_m + 8(q-1)w_v w_c + (q-2)^2 w_m w_c)\right\}, \end{aligned}$$

q	w_v	w_m	w_c	efficiency
3	1002	2002	949	0.9863
4	506	1108	1329	0.9503
5	304	691	1564	0.9108
6	217	469	1665	0.8863
7	178	337	1669	0.8767
8	159	254	1627	0.8753
9	147	197	1573	0.8775
10	139	158	1519	0.8807
20	87	37	1236	0.9042
50	40	6	1067	0.9203

Table 20.6 V-optimal weights ($\times 10^4$) for the $\{q, 2\}$-weighted simplex centroid design and relative V-efficiency of the optimal simplex lattice design: w_v vertices; w_m and w_c the centroid

with $qw_v + \{q(q-1)/2\}\, w_m + w_c = 1$.

It is then clear that a $\{q, 2\}$-weighted simplex centroid design with a small $w_c > 0$ has a lower V-value than the V-optimal $\{q, 2\}$-weighted simplex lattice design. Table 20.6 gives, for some q, the V-optimal weights for the former design and the relative V-efficiency of the latter design. Note that w_c is at least 0.09 for $3 \leq q \leq 50$, and is a maximum at $q = 7$. The relative efficiency is a minimum at $q = 8$. For $q = 3$, runs of `gosset` (Hardin & Sloane, 1993) suggest that the $\{3, 2\}$-weighted simplex centroid design is the continuous V-optimal design over the whole simplex. Since the ratio of the weights $w_v : w_m : w_c$ is very close to $1 : 2 : 0$ or $1 : 2 : 1$ for the weighted simplex lattice or centroid designs, respectively, some very efficient finite designs exist – for example $n = 9,\ \ n_v = 1,\ \ n_m = 2\ (0.982)$ and $n = 10,\ \ n_v = 1,\ \ n_m = 2,\ \ n_c = 1$ (1.000 to 3 d.p.). As noted in MBS, when $n = 6$ the $\{3, 2\}$-simplex lattice design is not V-optimal.

Now consider for $q = 3$ the regular hexagon, $t = 3,\ \ b_j = 2/3\ \forall j$. The vertex-centroid design for the quadratic model has V-value $(36w_v + 23w_c)/90w_vw_c$. Thus the optimal design has $w_c = (\sqrt{138} - 6)/17 \simeq 0.3381$, and $w_v \simeq 0.1103$. This design appears to be V-optimal over R and, similarly to Result 3, also for b up to 0.697, although the weights change slightly with b ($0.3374,\ 0.1104$ at $b = 0.697$). For higher b, w_v decreases, w_m increases and w_c decreases, reaching $0.050,\ \ 0.202,\ \ 0.095$, respectively, at $b = 1$ (simplex).

Result 2 applies for the parallelogram $(q = 3)$. The continuous V-optimal design for the quadratic model has weights $w_v \simeq 0.0911,\ \ w_m \simeq 0.0912$, and $w_c \simeq 0.2709$. This is the limiting design in the equal-apices case for the trapezium $(t = 1)$ and the symmetric pentagon $(t = 2)$ for low b, and for the rht-hexagon (see §3) for both low and high b. Evaluations at different b in these 3 cases show that patterns are similar to those for the D-criterion (see MBS and

§3), except that the centroid is much more highly weighted. Its weight is never zero, being at least 0.095 for the trapezium and symmetric pentagon, rising to 0.271 for the trapezium and 0.296 for the symmetric pentagon. For the rht-hexagon its weight ranges between 0.271 and 0.338. Because the a.v. centroid does not have the appropriate limits, it is more V-efficient for the symmetric pentagon and high b to use the c.o.m. centroid - the relative V-efficiency as b tends to 1 is 0.969. A similar but smaller effect occurs with the trapezium and the symmetric pentagon as b tends to 0 (relative V-efficiency 0.991).

6. DISCUSSION AND CONCLUSIONS

Result 3 explains why, when $x_j \leq b \; \forall j$, there are constant weights in an interval of b in the $q = 3$ hexagon case (§3), and $q = 4$ in the octahedron case (§4). It is much harder to show that the only s.p. are the vertices and centroid. Note that the relevant shape for constant weights is that when $b = 2/3$ (when the edge lengths are equal but there are 2 different planar face shapes), not, as might have been expected from the $q = 3$ result, the regular region for which $b = 2/q$. For $q = 5$ and $x_j \leq b \; \forall j$ the same property holds for $0.6363 \leq b \leq 0.6745$. However, for $q = 6$, it does not even hold for $b = 2/3$ (although the vertex-centroid design is highly G-efficient, 0.987, over the design points).

We have also considered, when $q = 3$, continuous D-optimal designs for some irregular shapes (differing $\{b_j\}$). The D-optimal designs reinforce the conclusions that the vertices almost always have positive support. Standard advice is to include the centroid, plus midpoints of long edges. The results here suggest that mid-edges often have higher weight than the overall centroid, and some face centroids can have also have higher weight. Except when R is close to a parallelepiped, mid-edges of relatively short sides usually have zero weight. If $t = q$, the D-optimal design for some $\{b_j\}$ may be the vertex-centroid design for which the mid-edges have zero weight.

In all considered cases of V-optimal designs, the weights for the vertices are lower than those for the D-optimal design, while those for the mid-edges and centroid are higher, particularly so for the centroid, which never has zero weight. Apart from the possible inclusion of the centroid, in the cases considered the s.p. for the V-optimal design are very similar to those for the D-optimal design.

Acknowledgements

We are grateful to N.J.A. Sloane of AT&T Labs Research for supplying the `gosset` program and for help in its implementation. Financial assistance from Johnson Matthey plc and the EPSRC is gratefully acknowledged.

References

Atkinson, A.C. and Donev, A.N. (1992). *Optimum Experimental Designs*. Oxford: University Press.

Hardin, R.H. and Sloane, N.J.A. (1993). A new approach to the construction of optimal designs. *J. Statist. Plann. Inference* **37**, 339–369.

Laake, P. (1975). On the optimal allocation of observations in experiments with mixtures. *Scandinavian Journal of Statistics* **2**, 153–157.

Martin, R.J., Bursnall, M.C. and Stillman, E.C. (1999). Efficient designs for constrained mixture experiments. *Statistics and Computing* **9**, 229–237.

Math Works (1992). *Matlab version 3.5*. Natick MA: The Math Works Inc.

Appendix: Sketch Proofs

Lemma 1 (§2). Write $X'WX$ as $P\Lambda P'$, with $\Lambda_1 = \text{diag}\,(\lambda_1, \ldots, \lambda_r)$ and $\Lambda = \text{diag}\,(\Lambda_1, 0I_{p-r})$ (spectral decomposition). Let $S = ZP = [S_1 S_2]$. Then $|X'WX + uZZ'| = |\Lambda + S'S| = |uS_2'S_2| \times |\Lambda_1|$ using the usual formula for the determinant of a partitioned matrix, with S_2 square and non-singular. Then $c_{X,Z} = |S_2|$. If $X = [X_1\ 0]$, treating $X'WX + uZZ'$ directly as a partitioned matrix gives $|X'WX + uZZ'| = |uZ_2'Z_2| \times |X_1'WX_1|$.

Lemma 2 (§2). Differentiate $h(x,u) = f(x)g(u) + \lambda\,(d_u u + \sum d_j x_j - d_0)$ with respect to the x_i and u, and set to 0. Then
$0 = \sum_i [x_i\,\{d_0 \times (dh/dx_i) - d_i \times (dh/du)\}] = ad_0 f(x)g(u) - f(x)g'(u) \times (\sum d_j x_j)$.

Result 1 (§2). Since $X'WX = \sum w_i x_i x_i'$, the product of its positive eigenvalues is a homogeneous polynomial in the w_i.

Result 3 (§2). In the equal-apices case, the form of $|X'WX|$ follows from Result 1. For c_b, partition X into linear and quadratic parts $[X_1\ X_2]$. Partitioning $X'WX$ similarly, the 4 sub-matrices are highly structured.

In the unequal-apices case, consider the first 3 removed apices, and their 6 vertices. Let the row of X corresponding to the vertex with $x_i = b_i,\ x_j = 1 - b_i$ be $x_{ij}{}'$. Then the 6 rows with $i \neq j,\ i,\ j = 1, 2, 3$ are linearly dependent, since $\sum [(1 - b_k)\,\{b_j\,(1 - b_j)\,x_{ij} - b_i\,(1 - b_i)\,x_{ji}\} / (b_i + b_j - 1)] = 0$, where the sum is over $(i,j) = (1,2), (2,3), (3,1)$, and $k \neq i, j$. All sets of triples from 1 to q can be obtained from those which include 1, so the number of linear dependencies in the rows of X is ${}^{q-1}C_2$. Thus the vertices have rank$(X) = q(q-1) - (q-1)(q-2)/2 = p - 1$.

Rht-hexagon weights (§3). Using X for the 4 $v1$ and 2 $v2$ vertices (weights w_1, w_2, respectively) and Result 1, the symmetry in $X'WX$, and that the product of the 5 positive eigenvalues is the sum of the principal minors of $X'WX$, sh ows that $|X'WX + uzz'|$ is proportional to $\left\{2(2b-1)^2 w_1 + b^2 w_2\right\} w_1^3 w_2 u$. Then maximize this subject to $u = 1/6$, and $2w_1 + w_2 = 5/12$.

V-value for the simplex centroid design (§5). M can be written down. Use X for the q vertices and $q(q-1)/2$ mid-edges, and z for the centroid (weights $w_v,\ w_m,\ u$, respectively). Then X is square and non-singular, so that $(X'WX + uzz')^{-1} = \frac{(X'WX)^{-1} - u(X'WX)^{-1}zz'(X'WX)^{-1}}{1+uz'(X'WX)^{-1}z}$, with $(X'WX)^{-1} = X^{-1}W^{-1}X'^{-1}$.

V-value for the regular hexagon (§5). M can be written down. $(X'WX)^{-1}$ follows easily from the structure of $X'WX$.

Discussion

Bertold Heiligers and Richard Martin

The results of Draper et al. (2000) and Draper & Pukelsheim (1999) on the second-order mixture model allow some stronger statements to be made about V-optimality on the simplex than those in §5 above. Firstly, for 3 components the 3, 2-weighted simplex-centroid design given there is V-optimal over the whole simplex.

Secondly, the results of Draper et al. (2000) show that for 4 components the V-optimal design over the whole simplex is a weighted centroid design. Numerical maximization then shows that the 4, 2-weighted simplex-centroid design given above is not V-optimal, although its efficiency is 0.9955. The V-optimal design has weights 0.0515, 0.0947, 0.0564 and 0 for the vertices, mid-edges, face centroids and centroid, respectively.

Thirdly, the results of Draper et al. (2000) show that for 5 components a weighted centroid design is V-optimal over the whole simplex, with support on the barycentres of depth 1, 2, 3, 5 or 1, 3, 4, 5. Numerical maximization shows that the 5, 2-weighted simplex-centroid design of the paper is not V-optimal, although its efficiency is 0.9830. The V-optimal centroid design has weights 0.0314, 0.0493 and 0.0350 for the vertices, mid-edges, and face centroids, respectively (barycentres of depth 1, 2, 3), with the higher-dimensional centroids having weight zero.

Additional References

Draper, N. R., Heiligers, B. and Pukelsheim, F. (2000). Kiefer ordering of simplex designs for second-degree mixture models with four and more ingredients. *Annals of Statistics*. (To appear).

Draper, N. R. and Pukelsheim, F. (1999). Kiefer ordering of simplex designs for first- and second-degree mixture models. *J. Statistical Planning Inf.* **79**, 325–348.

Chapter 21

COFFEE-HOUSE DESIGNS

Werner G. Müller
Department of Statistics
University of Economics and Business Administration
Augasse 2-6, A-1090 Vienna, Austria
werner.mueller@wu-wien.ac.at

Abstract Designs that attempt to cover the experimental region as uniformly as possible, so called space-filling designs, have regained much interest after their successful linkage to D-optimum type designs by Johnson *et al.* (1990). The proposed maximin distance designs, however, are very difficult to generate, especially when the region is irregular or/and high dimensional and the number of design points is large.

(Coffee-house) Designs (see Müller, 1998) which are constructed by sequentially adding maximin distance points are asymptotically equivalent and reasonably efficient. A variant with good projective properties will be proposed and compared to the approach by Morris and Mitchell (1995). A relation to a design algorithm for random coefficient regression, cf. Fedorov and Müller (1989), is revealed.

Keywords: Space-filling designs; spatial sampling; Latin hypercubes.

"In Wien geht man ins Café, um sich zurückzuziehen, und jeder setzt sich, inselbildend, soweit wie möglich von jedem anderen - [...]"
von Doderer (1951)

Introduction

Designs selected with only parameter estimation in mind (such as optimum designs) have been criticized for leaving large unsampled 'holes' in the design region $\mathcal{X}$. Particularly in the absence of prior knowledge about the structure of the response, we would want a design to achieve good overall coverage of the area being studied.

The most intuitive approach to space-filling is taken by Johnson *et al.* (1990). They propose minimax and maximin distance designs, where the latter are

defined as

$$\xi_n^* = \arg\max_{\xi_n} \min_{x_i, x_{i'} \in \xi_n} h_{ii'}. \tag{21.1}$$

Here $\xi_n = \{x_1, \ldots, x_n\}$ denotes a replication-free n-point design, and $h_{ii'} = ||x_i - x_{i'}||$ the interpoint distance in a suitable norm.

Maximin distance designs, even for small sample problems, may be very difficult to construct (especially when the design region is irregular and/or the dimensionality is high) and can have multiple structurally different solutions. Tobias (1995) writes that "space-filling designs can be very time-consuming to produce; search times of days are not uncommon for realistic candidate regions of interest." He (and others, e.g. Royle and Nychka, 1998) suggest using heuristic point exchange algorithms, whereas Morris and Mitchell (1995) give an extended definition of (21.1) to avoid multiple global optima and employ a simulated annealing algorithm for the computation of their optimum designs.

1. A SIMPLE CONSTRUCTION RULE

To avoid all these difficulties let us instead consider designs that are generated as follows:

a) First the two points

$$\xi_2 = \{x^{(1)}, x^{(2)}\} = \arg\max_{x_i, x_{i'} \in \mathcal{X}} h_{ii'}$$

with maximum distance among all pairs in the design region $\mathcal{X}$ are found.

b) Then the design is subsequently supplemented such that $\xi_{i+1} = \xi_i \cup \{x^{(i+1)}\}$ by

$$x^{(i+1)} = \arg\max_{x_0 \in \mathcal{X}} \min_{x_i \in \xi_i} h_i,$$

where $h_i = \|x_i - x_0\|$, the point which maximizes the minimum distance to all points in ξ_i, until the desired number of sites n is reached.

This construction rule is in accordance with the description in the epigraph above from a very popular Austrian novel of how Viennese select their table in a coffee-house. We may therefore call designs constructed by this rule coffee-house designs (for a typical table setup see Figure 21.3 in the Appendix).

The above construction algorithm is simple and quick, and the coffee-house designs seem to (at least asymptotically) share some desirable properties of maximin distance designs. Not only does

$$\lim_{n \to \infty} P(\max_{x_i^*} \min_{x_{i'}} \|x_i^* - x_{i'}\| < \epsilon) = 1$$

hold (for an arbitrarily small positive number ϵ and x_i^* being the points in a maximin distance design), which is also true for simple random sampling.

Moreover, the ratio $\frac{h_{\min}}{h^*_{\min}}$ of the minimum distance in a coffee-house and the one in a maximin design, although not tending to 1, seems to be bounded from below.

1.1. THE ONE-DIMENSIONAL CASE

In the one-dimensional setup it is evident that the regular equispaced sequence of points provides the maximin distance designs and we therefore have

$$h^*_{\min} = \frac{1}{n-1}.$$

The corresponding minimum distance in coffee-house designs is

$$h_{\min} = \max_{i \in \mathbb{N}} \frac{1}{2^i} \quad \text{s.t. } 2^i \geq n-1,$$

and the ratio is thus given by

$$\frac{h_{\min}}{h^*_{\min}} \geq \frac{2^{i-1}-1}{2^i}.$$

The above equality holds at $n = 2^{i-1}+2$ and therefore by resubstituting n in the second term we have as a lower bound

$$\frac{h_{\min}}{h^*_{\min}} \geq \frac{n-1}{2(n-2)} > \frac{1}{2},$$

i.e. the minimum distance in a coffee-house design is in the worst case twice as large as the optimum.

1.2. THE HIGHER-DIMENSIONAL CASE

Here an explicit evaluation of this lower bound is not possible due to the irregular sequence of maximin distance designs even for simply shaped design regions. However, the effectiveness of coffee house designs can be evaluated by comparing the achieved minimum distances to the average of minimum random spacings on the same region.

For the two-dimensional unit square, for instance, the coffee-house rule leads (after the vertices) to an alternating sequence of 4^i distances $\frac{\sqrt{2}}{2^{i+1}}$ ($i = 0, 1, \dots$) and $2^{j+1}(2^j+1)$ distances $\frac{1}{2^{j+1}}$ ($j = 0, 1, \dots$). The ratio of this sequence to the respective average minimum spacings, which is clearly an increasing function, is depicted for the first few distances in Figure 1.2.

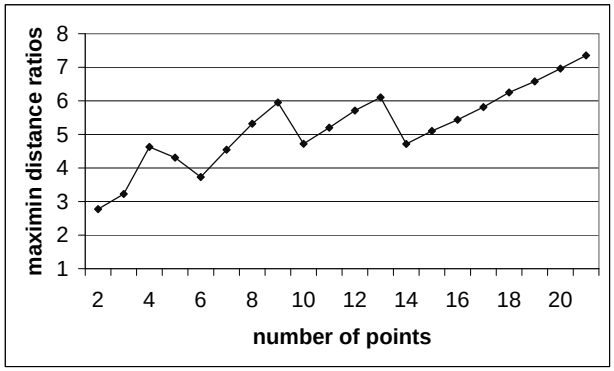


Figure 21.1 The ratio of the minimum distances in two-dimensional coffee-house designs to the respective average minimum random spacing.

2. A MODIFICATION WITH GOOD PROJECTIVE PROPERTIES

Maximin distance designs, especially when defined on hypercubic regions, usually do not exhibit good overall coverage when projected onto lower dimensional representations of the design space. For this reason Morris and Mitchell (1995) have suggested searching for maximin distance designs within a class of Latin hypercube arrangements. Specifically, they only allow designs for which each column of the design matrix contains some permutation of the elements of the set $\{0, \frac{1}{n-1}, \frac{2}{n-1}, \ldots, 1\}$, and they consequently call their resulting designs Minimax distance Latin hypercube designs.

There are now basically two variants of coffee-house designs that will have similar projective properties.

A) The Latin hypercube "grid" is fixed in advance and subsequently filled with the points according to the coffee-house rule.

B) At each step the maximin distance point(s) is/are found for all one-dimensional projections, and from all the resulting point combinations the one with the maximin distance to all others is selected.

For irregular regions we might be content with a point that is as close as possible (perhaps on a border) to the above defined candidate point. To see the difference between the two specifications a graphical representation for 3 to 10 point designs on a square is given in Figure 21.2.

Although variant A) at the first glance seems more natural, it can be seen (from Table 21.1) that its performance (with respect to the minimum distance ratio as a measure of coverage) is even worse than the one of random sampling.

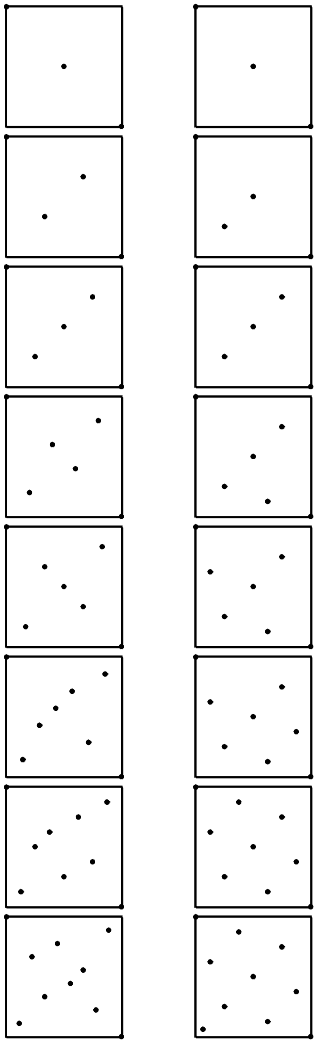


Figure 21.2 Two-dimensional variants A) and B) of coffee-house designs.

However, version B) seems to perform very well and clearly approaches the Maximin Latin hypercube designs for increasing numbers of points. Moreover, it distinctively outperforms random sampling, except for cases with very few design points.

n	random	A)	B)
3	1.000	1.000	1.000
4	0.667	0.632	0.475
5	0.674	0.633	0.633
6	0.673	0.633	0.792
7	0.546	0.501	0.752
8	0.542	0.500	0.876
9	0.478	0.448	0.896
10	0.484	0.447	0.755
11	0.481	0.446	0.839
12	0.421	0.393	0.604
13	0.425	0.392	0.658
14	0.372	0.344	0.625
15	0.369	0.342	0.671
16	0.367	0.342	0.720
17	0.355	0.332	0.747
18	0.356	0.332	0.792
19	0.360	0.335	0.839
20	0.354	0.332	0.888

Table 21.1 Minimum intersite distance ratios (to Maximin Latin Hypercube designs) for random design (average), A), and B).

3. THE RELATIONSHIP TO OTHER DESIGNS

Interest in space-filling designs has been renewed, especially for spatial applications, by the paper of Johnson *et al.* (1990). They have shown that if the data generating process is assumed to be a random field

$$y(x_i) = \eta(x_i, \beta) + \varepsilon(x_i), \tag{21.2}$$

with $E[\varepsilon(x_i)] = 0$ and given covariance function

$$E[\varepsilon(x_i), \varepsilon(x_{i'})] = c(h_{ii'}) = [c'(h_{ii'})]^{\kappa},$$

with $c'(h_{ii'})$ being itself a decreasing covariance function, a design according to (21.1) coincides with

$$\xi_n^* = \arg\max_{\xi_n} |C|,$$

if one lets $\kappa \to \infty$. Here C denotes the theoretical covariance matrix consisting of entries $C_{ii'} = c(h_{ii'})$.

This is an optimum design rule, which is intuitively in accordance with the heuristic design reduction procedure of Der Megréditchian (1985), which

is based upon the empirical counterpart of C. Namely, from a given design $\xi_n = \{x_1, \ldots, x_n\}$ with an observed covariance matrix $\hat{K}$ delete the point x_{i^-}, where

$$i^- = \arg\max_i \hat{K}_{ii}^{-1}.$$

There is even some relationship between this procedure and my joint work with the jubilarian we are honouring at this conference. Fedorov and Müller (1989) have shown that if a random coefficient specification of (21.2) is assumed, a deletion step in a classical design algorithm approximately coincides with the above procedure. Further relations, e.g. to maximum entropy sampling are given in Müller (1998).

Acknowledgement

I am grateful to I. Molchanov for a remark at the presentation, that led to an improvement of the paper.

Appendix

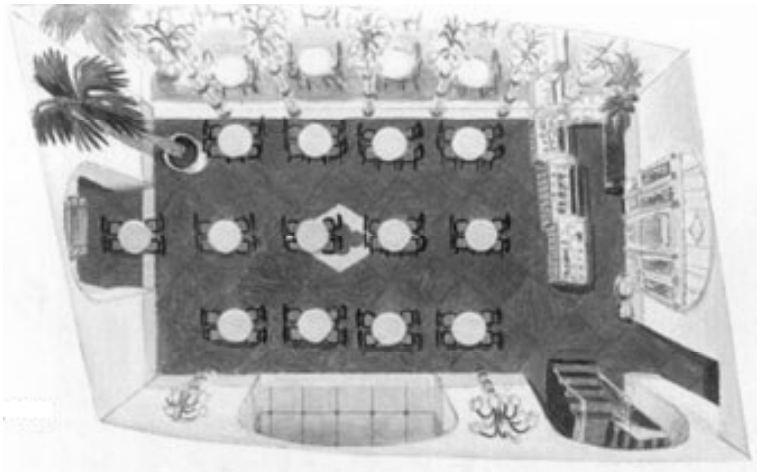


Figure 21.3 The table setup in a Viennese coffee-house (Demel).

References

Der Megréditchian, G. (1985). *Methodes Statistiques d'Analyse et d'Interpolation des Champs Meteorologiques*. Geneve: Organisation Meteorologique Mondiale.

von Doderer, H. (1951). *Die Strudlhofstiege*. München: Biederstein Verlag.

Fedorov, V.V. and Müller, W.G. (1989). Comparison of two approaches in the optimal design of an observation network. *Statistics* **20**, **3**, 339–351.

Johnson, M.E., Moore, L.M. and Ylvisaker, D. (1990). Minimax and maximin distance designs. *Journal of Statistical Planning and Inference* **26**, 131–148.

Morris, M.D. and Mitchell, T.J. (1995). Exploratory designs for computational experiments. *Journal of Statistical Planning and Inference* **43**, 381–402.

Müller, W. G. (1998). *Collecting Spatial Data*. Heidelberg: Physica Verlag.

Royle, J.A. and Nychka, D. (1998). An algorithm for the construction of spatial coverage designs with implementation in SPLUS. *Computers & Geosciences* **24**, **5**, 479–488.

Tobias, R. (1995). *SAS QC Software, Vol. 2: Usage and Reference.* Cary, NC: SAS Institute.

Chapter 22

(D_T, C)-OPTIMAL RUN ORDERS

Lieven Tack and Martina Vandebroek
Department of Applied Economics, Katholieke Universiteit Leuven
Naamsestraat 69, B-3000 Leuven, Belgium
lieven.tack@econ.kuleuven.ac.be and martina.vandebroek@econ.kuleuven.ac.be

Abstract Cost considerations have rarely been taken into account in optimum design theory. A few authors consider measurement costs, i.e. the cost associated with a particular factor level combination. A second cost approach results from the fact that it is often expensive to change factor levels from one observation to another. We refer to these costs as transition costs. In view of cost minimization, one should minimize the number of factor level changes. However, there is a substantial likelihood that there is some time order dependence in the results. Consequently, when considering both time order dependence and transition costs, an optimal ordering is not easy to find. There is precious little in the literature on how to select good time order sequences for arbitrary design problems and up to now, no thorough analysis of both costs is found in the literature. Our proposed algorithm incorporates both costs in optimum design construction and enables one to compute cost-efficient and nearly trend-free run orders for arbitrary design problems. The results show that cost considerations in the construction of trend-resistant run orders entail considerable reductions in the total cost of an experiment and imply a large increase in the amount of information per unit cost.

Keywords: Exchange Algorithm, $\mathcal{D}$-optimality, Trend-Robustness, Cost, Run Order

Introduction

In optimum design theory designs are constructed that maximize the information on the unknown parameters of the response function. Although such constructed designs have good statistical properties, they may not be fit for use because of economical reasons. Cost considerations have rarely been dealt with in the construction of optimum experimental designs. Generally speaking, two cost approaches are found in the literature.

A few authors deal with measurement costs, i.e. the costs associated with the particular factor level combinations. Measurement costs include the equipment cost, the cost of material, the cost of personnel, the cost for spending time during the measurement, etc. Kiefer (1959) and Neuhardt and Bradley (1971) respectively suggest complete and partial enumeration of appropriate designs in order to select the minimum cost design. Yen (1985) uses integer programming and Pignatiello (1985) finds cost-efficient fractional factorial designs which permit the estimation of specified main and interaction effects. Rafajłowicz (1989) seeks a minimum cost design between all designs with specified desired information matrix.

However, it is usually expensive to alter some of the factor levels, such as oven temperature or line set-up, from one observation to another. We refer to these costs as transition costs. An interesting approach to the minimization of the total transition cost comes from Ju (1992), Anbari (1993) and Anbari and Lucas (1994) who show how proper blocking on hard-to-change factors achieves efficient designs.

One drawback is that the latter authors do not take into account trend effects. Variables that often affect observations obtained in some specific order are equipment wear-out, learning, fatigue, etc. The relative cost-effectiveness of any sequence is a function of the cost of changing factor levels and the protection afforded against time order dependence. Minimization of factor level changes is no longer the only design issue of interest. The objective is to construct a run order such that the estimates of the important factorial effects are orthogonal or nearly orthogonal to the postulated polynomial trend. If the least-squares estimator of a factorial effect is the same as when the time trend of qth order is not present, that effect is said to be q-trend-free or orthogonal to the time trend. Cheng (1985), Coster and Cheng (1988) and Coster (1993) formulate the Generalized Foldover Scheme (GFS) for generating run orders which minimize the cost equal to the number of factor level changes and for which all main effects are orthogonal to a polynomial time trend. Based on Daniel and Wilcoxon (1966), Cheng and Jacroux (1988) derive a method for the construction of run orders that yield trend-free main effects and 2-factor interactions by assigning the main effect contrasts to higher order interaction contrasts in the standard order. A more extensive review of the construction of trend-free run orders can be found in Tack and Vandebroek (1999).

With the exception of Atkinson and Donev (1996), there is precious little in the literature on how to select good time order sequences for arbitrary design problems. However, Atkinson and Donev (1996) do not take into account cost considerations. Our concern is about cost-efficient run orders with maximal protection against time order dependence for arbitrary design problems, polynomial time trends of any order and arbitrary cost functions.

1. COSTS IN EXPERIMENTAL DESIGN

Before passing on to the construction of optimal run orders, this section elaborates our generic cost approach.

We define the measurement cost at design point $\mathbf{x}_i$ as $c^m(\mathbf{x}_i) = \mathbf{m}'(\mathbf{x}_i)\boldsymbol{\varsigma}$, where $\mathbf{m}(\mathbf{x}_i)$ is a column vector with p_m elements, representing the polynomial expansion of the design point for the measurement cost and $\boldsymbol{\varsigma}$ is a $(p_m \times 1)$ vector of coefficients. Assuming d design points, the total measurement cost C^m of an experiment equals

$$\begin{aligned} C^m &= \textstyle\sum_{i=1}^{d} n_i c^m(\mathbf{x}_i), \\ &= \mathbf{1}' \mathbf{N} \mathbf{M} \boldsymbol{\varsigma}, \end{aligned} \tag{22.1}$$

where n_i denotes the number of observations at design point i and $\mathbf{N}$ equals diag$(n_1, \ldots, n_d)$. $\mathbf{1}$ is a $(d \times 1)$ vector with elements 1 and the $(d \times p_m)$ matrix $\mathbf{M}$ equals $(\mathbf{m}'(\mathbf{x}_1) \ \cdots \ \mathbf{m}'(\mathbf{x}_d))'$. In practice, it frequently happens that cost information is available at only a subset of all treatment combinations. To deal with this problem, the calculation of the measurement cost at any design point is based on an interpolation technique.

In contrast with the measurement costs, the total transition cost of an experiment depends on the sequence in which the observations are taken. The transition cost $c^t(\mathbf{x}_i, \mathbf{x}_j)$ from design point $\mathbf{x}_i$ to design point $\mathbf{x}_j$ is the cost for changing the factor levels of design point $\mathbf{x}_i$ in the previous run to the factor levels of design point $\mathbf{x}_j$ in the next run. This transition cost is defined as $c^t(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{t}'(\mathbf{x}_i, \mathbf{x}_j)\boldsymbol{\tau}$, where $\mathbf{t}'(\mathbf{x}_i, \mathbf{x}_j)$ is a $(p_t \times 1)$ vector representing the polynomial expansion of design points $\mathbf{x}_i$ and $\mathbf{x}_j$ for the transition cost and $\boldsymbol{\tau}$ is a column vector with p_t coefficients. The total transition cost C^t of a run order equals

$$\begin{aligned} C^t &= \textstyle\sum_{i=1,j=1}^{d} n_{i,j} c^t(\mathbf{x}_i, \mathbf{x}_j), \\ &= \mathbf{1}' \mathbf{L} \mathbf{T} \boldsymbol{\tau}, \end{aligned} \tag{22.2}$$

where $n_{i,j}$ denotes the number of transitions from design point $\mathbf{x}_i$ to design point $\mathbf{x}_j$ in the considered run order. The column vector $\mathbf{1}$ contains d^2 1-elements. $\mathbf{L}$ is the $(d^2 \times d^2)$ matrix diag$(n_{1,1}, \ldots, n_{1,d}, \ldots, n_{d,1}, \ldots, n_{d,d})$ and the $(d^2 \times p_t)$-matrix $\mathbf{T}$ is written as $(\mathbf{t}'_{1,1} \ \cdots \ \mathbf{t}'_{1,d} \cdots \mathbf{t}'_{d,1} \ \cdots \ \mathbf{t}'_{d,d})'$. Based on the given cost information, a two-dimensional interpolation technique is used to calculate the transition costs. The first dimension refers to the factor levels of the previous run whereas the second dimension refers to the factor levels of the next run.

The total cost C of a run order is defined as the sum of the total measurement cost (22.1) and the total transition cost (22.2), or $C = \mathbf{1}' \mathbf{N} \mathbf{M} \boldsymbol{\varsigma} + \mathbf{1}' \mathbf{L} \mathbf{T} \boldsymbol{\tau}$.

2. (D_T, C)-OPTIMAL RUN ORDERS

This section deals with the incorporation of measurement costs and transition costs in the construction of designs that yield maximal protection against time order dependence. Attention will be drawn to arbitrary design problems, arbitrary cost functions, polynomial time trends of any order and whether the time points are equally spaced or not. The aim is the construction of the best run order in terms of information about the unknown parameters of the response function and costs.

Henceforth y denotes the response of interest and $\mathbf{x}' = (x_1 \ \cdots \ x_f)$ is the vector of f control variables presumed to influence the response. Denote by $\mathbf{f}(\mathbf{x})$ the $(p \times 1)$ vector representing the polynomial expansion of $\mathbf{x}$ for the response model and by $\mathbf{g}(t)$ the $(q \times 1)$ vector representing the polynomial expansion for the time trend, expressed as a function of time t. With $\boldsymbol{\alpha}$ the $(p \times 1)$ vector of important parameters and $\boldsymbol{\beta}$ the $(q \times 1)$ vector of parameters of the polynomial time trend, let the model for the response be of the form

$$y = \mathbf{f}'(\mathbf{x})\boldsymbol{\alpha} + \mathbf{g}'(t)\boldsymbol{\beta} + \varepsilon = \mathbf{z}'(\mathbf{x}, t)\boldsymbol{\gamma} + \varepsilon. \tag{22.3}$$

The independent error terms ε are assumed to have expectation zero and constant variance σ^2. It is convenient to write (22.3) as

$$\mathbf{Y} = \mathbf{F}\boldsymbol{\alpha} + \mathbf{G}\boldsymbol{\beta} + \boldsymbol{\varepsilon} = \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon},$$

where $\mathbf{Y}$ is an $(n \times 1)$ vector of responses and $\mathbf{F}$ and $\mathbf{G}$ represent the $(n \times p)$ and the $(n \times q)$ extended design matrices respectively. In the absence of trend effects, the $\mathcal{D}$-optimal design $\delta_{\mathcal{D}}$ is found by minimizing the determinant of the information matrix or $\mathcal{D} = |\mathbf{F}'\mathbf{F}|$. Now we consider three additional optimality criteria.

We define a run order to be $(\mathcal{D}, C)$-optimal if it maximizes the amount of information per unit cost. The construction of the $(\mathcal{D}, C)$-optimal run order is based on maximization of

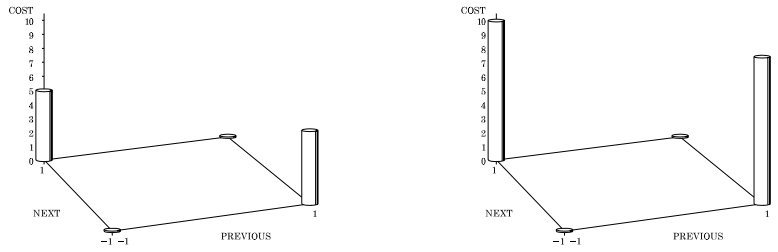
$$(\mathcal{D}, C) = \frac{|\mathbf{F}'\mathbf{F}|^{\frac{1}{p}}}{C}.$$

The power $1/p$ results in a measure of information per parameter. The efficiency of the $(\mathcal{D}, C)$-optimal design $\delta_{(\mathcal{D},C)}$ compared with the $\mathcal{D}$-optimal design $\delta_{\mathcal{D}}$ in terms of the amount of information per unit cost equals

$$\left(\frac{\mathcal{D}(\delta_{(\mathcal{D},C)})}{\mathcal{D}(\delta_{\mathcal{D}})}\right)^{\frac{1}{p}} \frac{C(\delta_{\mathcal{D}})}{C(\delta_{(\mathcal{D},C)})}. \tag{22.4}$$

Raising the determinants to the power $1/p$ results in an efficiency measure which is nearly proportional to design size. Tack and Vandebroek (1999) show that this linearity is only precise for the average transition cost per run order.

The benefit of incorporating cost information in the construction of optimal designs is illustrated by the following example. An imaginary two-factor experiment with $n = 15$ observations is set-up and the assumed model is described by $\mathbf{f}'(\mathbf{x}) = (1 \;\; x_1 \;\; x_2 \;\; x_1x_2)$. The design points constitute the full 2^2-factorial. The measurement costs are given as $c^m(\mathbf{x}) = 15 - 2.5x_1 + 2.5x_2$ and the transition costs are shown in Figure 22.1. For instance, changing factor x_1 from the low level to the high level or vice versa amounts to a cost of 5. Even when there is no factor level change from the previous run to the next one, the factor levels have to be adjusted again in order to preserve the assumption of uncorrelated error terms in model (22.3). Consequently, transition costs have to be taken into account. In practice, this is done by a little perturbation of the factor level of the previous run and then by adjusting it again at the same level of the previous run. As shown in Figure 22.1 these transition costs are rather low. For each possible design, the amount of information $|\mathbf{F}'\mathbf{F}|^{\frac{1}{p}}$ and the total cost of the cheapest run order belonging to that design is calculated (Figure 22.2). The $\mathcal{D}$-optimal design is shown in the middle of the three points on the right. The ($\mathcal{D}, C$)-optimal design is indicated just below the $\mathcal{D}$-optimal design. The decrease in information of the ($\mathcal{D}, C$)-optimal design compared with the $\mathcal{D}$-optimal design is negligible and the decrease in the total cost amounts to 2%. The resulting increase in the amount of information per unit cost of the ($\mathcal{D}, C$)-optimal design amounts to 2%.



(a) factor x_1 (b) factor x_2

Figure 22.1 Transition Costs

In the presence of time trends when no costs are calculated, designs are constructed that maximize the information on the important parameters $\boldsymbol{\alpha}$, whereas the q parameters modeling the time dependence are treated as nuisance parameters. The corresponding $\mathcal{D}_t$-optimal design $\delta_{\mathcal{D}_t}$ is found by maximization

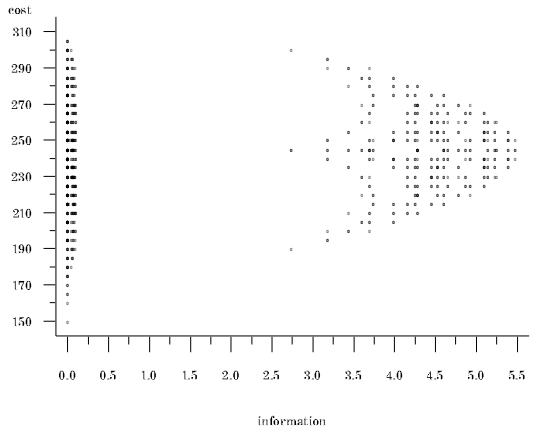


Figure 22.2 Cost and Information

of

$$\mathcal{D}_t = \frac{|\mathbf{Z}'\mathbf{Z}|}{|\mathbf{G}'\mathbf{G}|}.$$

A run order is called trend-free if the least-squares estimates of the factorial effects of interest are free of bias that might be introduced from the unknown trend effects in $\boldsymbol{\beta}$. Otherwise stated, trend-robustness is obtained when the columns of $\mathbf{F}$ are orthogonal to the columns of $\mathbf{G}$ or, equivalently, when $\mathbf{F}'\mathbf{G} = \mathbf{G}'\mathbf{F} = \mathbf{0}$. Then $|\mathbf{Z}'\mathbf{Z}| = |\mathbf{F}'\mathbf{F}||\mathbf{G}'\mathbf{G}|$. To compare the $\mathcal{D}$- and $\mathcal{D}_t$-optimal design for information about the important parameters $\boldsymbol{\alpha}$, the generalized variance of $\boldsymbol{\alpha}$ is compared through

$$\left(\frac{\mathcal{D}_t(\delta_{\mathcal{D}_t})}{\mathcal{D}(\delta_{\mathcal{D}})}\right)^{\frac{1}{p}}, \tag{22.5}$$

denoting the trend-resistance of the $\mathcal{D}_t$-optimal design.

Finally, in the presence of trend effects and when both measurement costs and transition costs are taken into account, the $(\mathcal{D}_t, C)$-optimal run order maximizes

$$(\mathcal{D}_t, C) = \frac{1}{C}\left(\frac{|\mathbf{Z}'\mathbf{Z}|}{|\mathbf{G}'\mathbf{G}|}\right)^{\frac{1}{p}}.$$

Analogously to the trend-resistance of the $\mathcal{D}_t$-optimal run order (22.5), the trend-resistance of the $(\mathcal{D}_t, C)$-optimal run order is defined as

$$\left(\frac{\mathcal{D}_t(\delta_{(\mathcal{D}_t,C)})}{\mathcal{D}(\delta_{\mathcal{D}})}\right)^{\frac{1}{p}}. \tag{22.6}$$

In the next section, we propose an algorithm for the construction of $(\mathcal{D}, C)$-, $\mathcal{D}_t$- and $(\mathcal{D}_t, C)$-optimal run orders.

3. THE DESIGN ALGORITHM

The aim of our proposed exchange algorithm is to construct optimal run orders by allocating n observations selected from a candidate list of d design points to n out of h available time points in such a way as to maximize the value of the optimality criterion used. The user specified optimality criterion may be $\mathcal{D}$-, $(\mathcal{D}, C)$-, $\mathcal{D}_t$- or $(\mathcal{D}_t, C)$-optimality.

Sometimes the experimenter wishes to include $n_1 > 0$ design points with corresponding time points. This can be done in the first phase of the algorithm. Next, a starting run order is constructed by allotting n_2 randomly chosen treatment combinations from the candidate list to n_2 randomly chosen time points from the list of available time points. This starting run order is then augmented to n trials in the second phase, by sequentially adding $n - n_1 - n_2$ treatment combinations at time points still available so that these additions lead to the largest improvement of the optimality criterion. Finally, the trials are subject to iterative improvement in the third phase. This improvement consists of alternate exchange and interchange of design points. The effect is investigated of the deletion of a design point $\mathbf{x}_i$ at time point t_k and the addition of a new design point $\mathbf{x}_j$ from the list of candidate points still available at a time point t_l. The interchange of design points $\mathbf{x}_i$ and $\mathbf{x}_j$ from $(\mathbf{x}_i, t_k)$ and $(\mathbf{x}_j, t_l)$ to $(\mathbf{x}_i, t_l)$ and $(\mathbf{x}_j, t_k)$ is also investigated. The process continues as long as an exchange or interchange increases the value of the optimality criterion used. In order to avoid being stuck at a local optimum, the probability of finding the global optimum can be increased by repeating the search several times from different starting designs or tries.

4. APPLICATIONS

This section illustrates the benefits of incorporating cost information in the construction of optimal designs. Two examples clarify practical utility in industrial environments.

4.1. EXAMPLE 1: POLISHING IN VLSI

This example is based on an experiment reported by Freeny and Lai (1997). In chemical mechanical polishing in very large scale integration, a wafer is held by a rotating carrier and is polished by pressing the wafer face down onto a polishing pad on a rotating platen. The important parameters for the polishing process are platen and wafer rotation frequencies, x_1 and x_2 respectively. Fifteen polishing conditions combining every platen frequency (11, 15 and 19 rpm) with every wafer frequency (12, 22, 32, 42 and 52 rpm) form a full facto-

rial experiment. A tendency of the polisher removal rate to drift lower through time had previously been noticed. An important design issue was to choose the order of the fifteen combinations to estimate the effects of the design parameters independently of the linear drift. In the experiment mentioned by Freeny and Lai (1997), the run sequence of polishing conditions is chosen to confound the effect of the linear drift with interaction component $x_1^2x_2^3$ (Table 22.1). In our model, the interaction term $x_1^2x_2^3$ is therefore omitted and a linear trend described by **g**(t) = t is introduced. Due to functional marginality, the higher-order interaction $x_1^2x_2^4$ is also omitted.

							Freeny and Lai								
run	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x_1	15	19	11	15	11	19	11	15	19	11	19	15	19	11	15
x_2	22	42	42	52	12	12	32	32	22	52	52	12	32	22	42
							$\delta_{\mathcal{D}_t}$								
run	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x_1	15	11	19	15	11	19	15	19	11	19	11	19	11	15	15
x_2	42	22	22	12	52	52	32	32	32	12	12	42	42	52	22

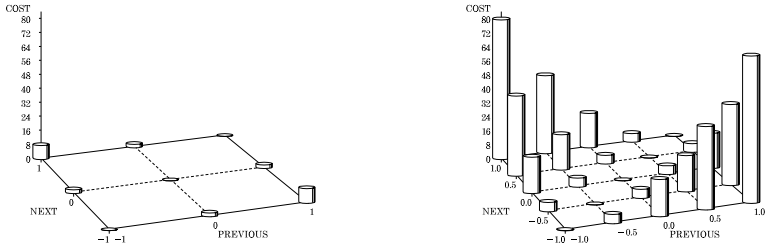
Table 22.1 Run Orders

The assumed response function is given by

$$\mathbf{f}'(\mathbf{x}) = (1 \;\; x_1 \;\; x_2 \;\; x_1^2 \;\; x_1x_2 \;\; x_2^2 \;\; x_1^2x_2 \;\; x_1x_2^2 \;\; x_2^3 \;\; x_1^2x_2^2 \;\; x_1x_2^3 \;\; x_2^4 \;\; x_1x_2^4).$$

Previous work indicated that the polisher may not reach equilibrium immediately after a change in the parameter settings. We turn this knowledge into a transition cost c^t and we assume an increasing cost function of factor level changes. Moreover, $c^t = c_1^t + c_2^t$ where c_i^t refers to the transition cost associated with changing the levels of factor i. We computed the $\mathcal{D}_t$- and $(\mathcal{D}_t, C)$-optimal run orders $\delta_{\mathcal{D}_t}$ and $\delta_{(\mathcal{D}_t,C)}$ for several ratios $\frac{c_1^t}{c_2^t}$. For instance, $\frac{c_1^t}{c_2^t} = 0.1$ means that factor x_1 is ten times cheaper to change than factor x_2. The transition costs for $\frac{c_1^t}{c_2^t} = 0.1$ are shown in Figure 22.3. Again, consecutive replicate factor levels are assigned a non-zero transition cost. The optimal run orders are compared in Table 22.2 with the run order of Freeny and Lai. The results presented relate to the quadratic cost functions of Figure 22.3 but are similar for other increasing cost functions.

The computed $\mathcal{D}_t$-optimal run order $\delta_{\mathcal{D}_t}$ is shown in Table 22.1 and one observes from Table 22.2 that this run order is a little more trend-resistant than the run order proposed in Freeny and Lai (1997). A decrease in trend-robustness is observed when transition costs are allowed for. The $\mathcal{D}_t$-optimal run orders outperform the run order of Freeny and Lai in terms of cost per unit information



(a) c_1^t for Platen Rotation Frequency

(b) c_2^t for Wafer Rotation Frequency

Figure 22.3 Transition Costs for $\frac{c_1^t}{c_2^t} = 0.1$

	trend-resistance (%)			cost per unit information		
$\frac{c_1^t}{c_2^t}$	Freeny and Lai	$\delta_{\mathcal{D}_t}$	$\delta_{(\mathcal{D}_t,C)}$	Freeny and Lai	$\delta_{\mathcal{D}_t}$	$\delta_{(\mathcal{D}_t,C)}$
1000	98.69	99.60	83.45	43,960	46,275	3,279
100	98.69	99.60	80.58	4,414	4,642	346
10	98.69	99.60	80.58	460	479	44
1	98.69	99.60	89.03	65	63	10
0.1	98.69	99.60	83.49	250	210	37
0.01	98.69	99.60	74.87	2,104	1,679	210
0.001	98.69	99.60	74.87	20,640	16,373	1,839

Table 22.2 Comparison of Run Orders

for low ratios $\frac{c_1^t}{c_2^t}$. Besides, the $(\mathcal{D}_t, C)$-optimal run orders imply considerable decreases in the cost per unit information. These percentage decreases range from 82% to 93%.

4.2. EXAMPLE 2: FLAME SPECTROSCOPY

This example was mentioned by Joiner and Campbell (1976). An experiment is set-up in order to evaluate the sensitivity of a spectrophotometer. Because of a linear drift due to carbon build-up, it is necessary to remove all of the built up carbon after every 20 observations. The number of levels per factor and the times needed to change the factor levels are given in Table 22.3.

We assume the total measurement cost to be fixed and the transition costs are supposed to be proportional to the times needed to change the factor levels. Table 22.4 shows the comparison between the computed $\mathcal{D}_t$- and $(\mathcal{D}_t, C)$-optimal run orders for the following response models:

	factor	number of levels	time to change (sec)
x_1	lamp position	2	1
x_2	burner position	2	60
x_3	burner height	3	1
x_4	type of flame	3	60
x_5	flow rate	3	120

Table 22.3 Description of the Flame Spectroscopy Experiment

	transition cost			trend-resistance (%)			cost per information		
$\mathbf{f}'(\mathbf{x})$	$\delta_{\mathcal{D}_t}$	$\delta_{(\mathcal{D}_t,C)}$	red.	$\delta_{\mathcal{D}_t}$	$\delta_{(\mathcal{D}_t,C)}$	red.	$\delta_{\mathcal{D}_t}$	$\delta_{(\mathcal{D}_t,C)}$	red.
(1)	$\geq$ 2,669	394	85	100	45.11	55	$\geq$ 134	46	66
(2)	$\geq$ 2,537	623	75	99.99	64.94	35	$\geq$ 245	95	61
(3)	4,170	694	83	82.70	66.03	20	266	57	79
(4)	3,864	1,357	65	77.81	72.51	6.8	422	161	62

Table 22.4 Comparison of Optimal Run Orders

(1) $\mathbf{f}'(\mathbf{x}) = (1\ x_1\ x_2\ x_3\ x_4\ x_5)$
(2) $\mathbf{f}'(\mathbf{x}) = (1\ x_1\ x_2\ x_3\ x_4\ x_5\ x_3^2\ x_4^2\ x_5^2)$
(3) $\mathbf{f}'(\mathbf{x}) = (1\ x_1\ x_2\ x_3\ x_4\ x_5\ x_1x_2\ x_1x_3\ x_1x_4\ x_1x_5\ x_2x_3\ x_2x_4\ x_2x_5\ x_3x_4\ x_3x_5\ x_4x_5)$
(4) $\mathbf{f}'(\mathbf{x}) = (1\,x_1\,x_2\,x_3\,x_4\,x_5\,x_3^2\,x_4^2\,x_5^2\,x_1x_2\,x_1x_3\,x_1x_4\,x_1x_5\,x_2x_3\,x_2x_4\,x_2x_5\,x_3x_4\,x_3x_5\,x_4x_5)$

The reduction in terms of percentage in the total transition cost, trend-resistance and the cost per unit information of the $(\mathcal{D}_t, C)$-optimal run order with respect to the $\mathcal{D}_t$-optimal run order is also mentioned. Note that, for the first two models of Table 22.4, more than one $\mathcal{D}_t$-optimal run order is found. As a consequence, different transition costs are obtained for these $\mathcal{D}_t$-optimal run orders. Table 22.4 shows that allowing for cost considerations implies large reductions in the total cost of an experiment. However, this reduction is obtained at the cost of the degree of trend-resistance of the computed $(\mathcal{D}_t, C)$-optimal run orders. The decrease in trend-resistance especially comes true for the first two models. But roughly speaking, the $(\mathcal{D}_t, C)$-optimal run orders outperform the $\mathcal{D}_t$-optimal ones in terms of the cost of information. The reductions in the cost of information range from 61% to 79%.

5. CONCLUSION

Economical reasons often limit the usefulness of experimental designs computed on the basis of alphabetic optimality criteria. However, the incorporation of cost considerations in optimum design theory is a topic about which the literature is conspicuously silent. This paper provides an analysis of cost con-

siderations in the construction of optimum designs and presents an algorithm for the construction of cost-efficient run orders that are optimally protected against specific time trends. Arbitrary design problems and arbitrary cost models can be studied. The results show that incorporating cost information implies a considerable increase in the amount of information per unit cost attended by a loss in the degree of trend-resistance.

References

Anbari, F. T. (1993). *Experimental Designs for Quality Improvement when there are Hard-to-Change Factors and Easy-to-Change Factors*. PhD Thesis, Drexel University.

Anbari, F. T. and Lucas, J. M. (1994). Super-efficient designs: How to run your experiment for higher efficiency and lower cost. *ASQC 48th Annual Quality Congress Proceedings*, 852–863.

Atkinson, A. C. and Donev, A. N. (1996). Experimental designs optimally balanced for trend. *Technometrics* **38**, 333–341.

Cheng, C. S. (1985). Run Orders of Factorial Designs. In *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer*, Vol. II, Eds L. LeCam and R. A. Olshen, pp. 619–633. Belmont, CA: Wadsworth.

Cheng, C. S. and Jacroux, M. (1988). The construction of trend-free run orders of two-level factorial designs. *Journal of the American Statistical Association* **83**, 1152–1158.

Coster, D. C. (1993). Tables of minimum cost, linear trend-free run sequences for two- and three-level fractional factorial designs. *Computational Statistics and Data Analysis* **16**, 325–336.

Coster, D. C. and Cheng, C. S. (1988). Minimum cost trend-free run orders of fractional factorial designs. *The Annals of Statistics* **16**, 1188–1205.

Daniel, C. and Wilcoxon, F. (1966). Fractional 2^{p-q} plans robust against linear and quadratic trends. *Technometrics* **8**, 259–278.

Freeny, A. E. and Lai, W. Y.-C. (1997). Planarization by chemical mechanical polishing: a rate and uniformity study. In *Statistical Example Studies for Industrial Process Improvement*, Eds V. Czitrom and P. D. Spagon, pp. 251–264. Philadelphia and Alexandria: SIAM and ASA.

Joiner, B. L. and Campbell, C. (1976). Designing experiments when run order is important. *Technometrics* **18**, 249–259.

Ju, H. L. (1992). *Split Plotting and Randomization in Industrial Experiments*. PhD Thesis, University of Delaware.

Kiefer, J. (1959). Optimum experimental designs. *Journal of the Royal Statistical Society* B **21**, 272–319.

Neuhardt, J. B. and Bradley, H. E. (1971). On the selection of multi-factor experimental arrangements with resource constraints. *Journal of the American Statistical Association* **66**, 618–621.

Pignatiello, J. J. (1985). A minimum cost approach for finding fractional factorials. *IIE Transactions* **17**, 212–218.

Rafajłowicz, E, (1989). Minimum cost experimental design with a prescribed information matrix. *SIAM Theory of Probability and its Applications* **34**, 367–370.

Tack, L. and Vandebroek, M. (1999). $(\mathcal{D}_t, C)$-Optimal run orders. *Submitted*.

Yen, V. (1985). Cost optimal saturated regression designs. *ASA Proceedings of Business and Economic Statistics Section*, 286–288.

Chapter 23

OPTIMAL DESIGN IN FLEXIBLE MODELS, INCLUDING FEED-FORWARD NETWORKS AND NONPARAMETRIC REGRESSION

D.M. Titterington
Department of Statistics, University of Glasgow
Glasgow, G12 8QQ
mike@stats.gla.ac.uk

Abstract Feed-forward networks, also known as multilayer perceptrons, are the most frequently implemented type of neural network. In statistical terminology, they can be regarded as a class of nonlinear regression or classification models, depending on the context, nonlinear in terms of both the explanatory variables and the parameters. An attempt at optimal design therefore leads to a nonlinear design problem. In principle, statistical work in this area can be applied to this context, and a major aim of the paper will be to survey relevant material that has appeared in the neural-computing literature, where it is described under headings such as 'active learning', as well as 'optimal design'. This part of the chapter will reinforce the contribution of Haines (1998).

A major reason for the attraction of feed-forward networks is that they can provide parametric but flexible regression models. One can consider going further and discuss the question of optimal design in nonparametric regression scenarios. The chapter discusses this issue and in particular the approach taken by Cheng et al. (1998) in the context of local linear smoothing.

Keywords: active learning, Bayesian design, local linear smoothing, neural networks, nonlinear, nonparametric regression, sequential design

1. MULTILAYER PERCEPTRONS AND 'ACTIVE LEARNING'

Multilayer perceptrons, otherwise known as feed-forward networks or back-propagation networks, are the most commonly used class of so-called artificial neural network models. Figure 8 of Cheng and Titterington (1994) depicts

the simplest and most often applied version of this structure. From the point of view of mainstream statistics, one can interpret this simply as a pictorial representation of a regression function

$$E(y) = g\{w_0 + \sum_{k=1}^{m} w_k f(v_{k0} + \sum_{j=1}^{p} v_{kj} x_j)\}, \qquad (23.1)$$

where the parameters $\{w_j\}$ and $\{v_{jk}\}$ are referred to as 'weights' in the neural-network literature. The architecture of the network is that of a set of 'nodes', joined up in a directed way. There are input nodes, corresponding to inputs/covariates x, an output node or nodes, corresponding to the (expected) response(s) y, and a layer of 'hidden nodes' that can be thought of as associated with hidden or latent variables whose introduction enriches the resulting model. The function g is the so-called activation function at the output node and f is the common activation function at each of the hidden nodes. If there was no hidden layer and if g and f were identity functions then the model would correspond to the standard multiple linear regression model. Often g is indeed taken to be the identity function, but f is traditionally nonlinear and usually a sigmoid, such as the logistic distribution function or the tanh function. This level of nonlinearity, together with the hidden layer, creates an extremely flexible, albeit still parametric, class of models for prediction. In other applications g is taken to be the sign function and the right-hand side of (23.1) then creates a two-class nonlinear (in the data) discriminant function.

There is a vast and continually growing literature concerning the choice of architecture ('model selection') and subsequent network training ('estimation of parameters'); see Cheng and Titterington (1994), Ripley (1994) and Titterington (1999) for statistically oriented reviews. When all the variables are continuous, parameter estimation is usually carried out by nonlinear least squares, so that the mechanics correspond to the assumptions underlying nonlinear regression with additive Gaussian errors. In most contexts there is no scope for preliminary manipulation of the data, i.e. for design, so that the estimation exercise has been termed 'passive learning'. However, the neural-network community have recently come alive to the possibility of 'active learning', and to the realisation that judicious choice of inputs can greatly influence the efficiency of estimators or otherwise lead to some form of economy. In other words, ideas from optimal design of experiments have found new application in this particular class of nonlinear problems, as we describe in Section 3. First, we briefly outline other ideas with an active-learning flavour from the neural-network literature.

2. QUERY-BASED LEARNING

The approaches mentioned in this section are mainly quite different in character from what statisticians usually think of as optimal design, but they ap-

peared in the neural-network journals and were precursors of the recognition of statistical optimal design within that community.

As remarked in Section 1, many feed-forward networks are used to create flexible discriminant or classification functions. In such contexts, the item of crucial interest is the discriminant boundary, and in general one wants to 'design' so as to gain maximum information about this boundary. In Baum (1991), it is envisaged in the context of binary discrimination that there is an 'oracle' that will, if asked, identify on which side of the true discriminant boundary a given set of inputs x lies. (This implies therefore that the two subpopulations underlying the classification problem are perfectly separable, which of course limits the applicability of this work.) Thus one is at liberty to select sets of inputs, exhibit them to the oracle as 'queries' and be given essentially noise-free responses. The objective is to establish good strategies for choosing the queries, and the spirit of Baum's algorithm is to home in on the decision boundary as effectively as possible. The algorithm is not based on any optimality criterion, but it is shown to be provably effective according to the tenets of computational learning theory in being able to PAC (probably almost correct(ly), see Anthony and Biggs, 1992) learn the discriminant boundaries associated with feed-forward networks with up to and including four hidden variables in the single hidden layer. Of course, 'four' is rather a small number in this context, but good empirical performance is also reported in examples based on up to 200 hidden variables and up to 200 input variables. The same sort of goal is pursued by Hwang et al. (1991), again based on setting queries before an error-free oracle, but using a different type of algorithm to home in on the discriminant boundary. Nothing is actually proved about the algorithm, but it does perform much better than random choice of inputs when applied to a real power system security problem; in fact the network used for this application contained an extra layer of complexity, compared with (23.1), in the form of a second layer of hidden variables.

Sollich (1994) comments on the heuristic and problem-specific nature of algorithms such as those considered by Baum (1991) and Hwang et al. (1991), and instead aims for an approach that is both more general and based on some concept of optimality. There is a much stronger relationship with optimal design and with the version of it applied by MacKay (1992); see Section 3.2. In particular, one of the criteria for optimality he adopts is that of entropy, or information gain, also used by MacKay (1992), but, at least in the general formulation, Sollich allows the model adopted by the experimenters to be different from the true model. Sollich also considers an alternative optimality criterion, the so-called generalisation error, which measures the performance of an estimated model in predicting the responses of the universe of possible future observations. He considers truly sequential design in which 'queries' are made one by one, and the relevant posterior distributions are updated and the

network retrained at each stage. The query, i.e. the design point, selected is that which optimises the value of the chosen optimality criterion, averaged over the current posterior distributions associated with the experimenters' and the true models. Sollich gives a detailed account of the application of the approach to two fairly simple problems, in both of which the true and assumed models are taken to be the same. One of the problems is distinctly nonlinear, but the other is linear, with distributions chosen to be Gaussian so that exact computations are typically feasible. Sollich's experimental results showed that, compared with a strategy of selecting new observations at random, his method significantly improved the performance in terms of generalisation error in the case of the nonlinear problem, but much less so in the linear example.

Plutowski and White (1993) consider a noise-free version of the problem, in which they choose sequentially from a set of candidate design points so as to optimise the stage-by-stage improvement in integrated squared bias incurred if the best-fitting feedforward network is used to interpolate the true response surface. They start off with a simple network with one hidden unit and enrich the architecture as necessary, as evidence comes in from the new 'data' that a more sophisticated network is required in order to give an adequate fit. The calculation of the change in the integrated squared bias is eased by using a Taylor-type approximation, and care is taken that at each stage the sum-of-squares function that measures the effectiveness of the fitting procedure is sufficiently small; in particular, this dictates whether or not the number of hidden units needs to be increased. Integrated squared bias is the noise-free analogue of integrated mean squared error, so the spirit of the algorithm is much the same as that of prediction-based optimal design, with the design space discretised into the set of 'candidate points'. Plutowski and White (1993) illustrate the algorithm with two very simple examples, indicating the gains to be made over uniform designs; with their method, effective learning is possible using many fewer design points than if uniform designs are used.

For another query-based active learning scheme see Krogh and Vedelsby (1995).

3. APPLICATION OF STATISTICAL OPTIMAL DESIGN TO FEED-FORWARD NETWORKS

3.1. CLASSICAL METHODS

Clearly, equation (23.1) is a special form of a nonlinear regression function for which the general case can be written

$$E(y) = \eta(x, \theta),$$

where θ represents all the parameters in w and v. If independent, constant-variance Gaussian errors are assumed, then the approximate information matrix

for θ based on a design ξ is of course proportional to

$$M(\xi,\theta) = \int \left(\frac{\partial\eta(x,\theta)}{\partial\theta}\right)\left(\frac{\partial\eta(x,\theta)}{\partial\theta}\right)^T \xi(dx), \tag{23.2}$$

and the approximate variance of the expected predicted response at design point x, based on data from a design ξ, is proportional to

$$d(x,\xi,\theta) = \left(\frac{\partial\eta(x,\theta)}{\partial\theta}\right)^T \{M(\xi,\theta)\}^{-1} \left(\frac{\partial\eta(x,\theta)}{\partial\theta}\right). \tag{23.3}$$

As in the setting of general nonlinear problems (see for example Ford et al., 1989) the expressions in (23.2) and (23.3) form the basis of equivalence theorems about locally optimal designs, and of algorithms, usually sequential in practice, for generating designs that one hopes will be optimal.

Haines (1998) applied this approach to the calculation of D-optimal designs for a very simple multilayer perceptron, with a scalar covariate x, i.e. $p = 1$, two hidden units, i.e. $m = 2$, logistic-sigmoid activation function f and activation function g equal to the identity function. The design obtained was generated by maximising numerically the determinant of $M(\xi,\hat{\theta})$, where $\hat{\theta}$ was calculated by least squares from data from a preliminary set of 25 observations at equally spaced design points. Even with this very simple example, the potential problems associated with multiple minima of the least squares surface were revealed; two other local minima were found, using the preliminary data, both of which led to very poorly conditioned information matrices. Haines (1998) also implemented two sequential approaches, in the first case adding to the pilot design a design point for which the d-function is maximum, and in the second case adding a general design such that the combination of this and the pilot design optimised the resulting D-criterion, evaluated at $\hat{\theta}$.

The use of standard nonlinear optimal design theory and algorithms was also used with feed-forward networks by Cohn (1996). He used mainly sequential design methods with a view to fitting multilayer perceptron models for two-dimensional arm movements. The dimensions of the investigations were much higher than in the example in Haines (1998); there were up to 6 covariates and up to 4 response variables, and $m = 20$ hidden units were included in the neural-network model. Cohn (1996) compared the sequentially obtained optimal design on the basis of prediction mean squared error with what resulted from randomly chosen design points. In general, the optimal design greatly improved variance, but ultimately not necessarily mean squared error. Cohn (1996) conjectured that the less-than-satisfactory mean-squared-error performance resulted from bias incurred by the optimal design algorithm having become trapped in an inappropriate local minimum of the sum-of-squares surface, or, equivalently, an unhelpful local maximum of the Gaussian-based loglikelihood surface.

The problems caused by multiple optima pervade discussion of the practical implementation of these neural-network models, especially when the parameter space is of high dimension. A further issue follows from the fact that in practice neural-network models are often applied to complicated real-life contexts in which the 'true' model simply does not conform to the model represented by the multilayer perceptron, flexible though that model is. This is of course an issue that pervades Statistics, and motivates the development of robust designs; see for example Chang and Notz (1996) and Wiens (1999).

3.2. BAYESIAN METHODS

A similar caveat about the realism of the models is raised by MacKay (1992), who remarks with concern that when constructing designs one makes the practically implausible assumption that the model under investigation is correct. His approach to design is a Bayesian one in which the object of central interest is the posterior distribution of the weights/parameters θ, given the data to date. His sequential-design rationale is to choose the next design point so as to achieve the greatest expected improvement in the Shannon entropy of the posterior distribution, averaged over the random variation in the response at the next observation. To achieve a tractable procedure he assumes that the posterior distributions are approximately Gaussian, so that the key quantity turns out to be the posterior covariance matrix and we are back in familiar territory; optimising the entropy improvement amounts to sequential improvement of the D-criterion. MacKay (1992) also considers criteria based on optimum prediction at one or several locations in the design space, as well as procedures for constructing designs for discriminating between models. In all cases, the criterion is expressed in terms of 'information gain', but recourse to the assumption of Gaussian approximations leads back to methods from Fedorov (1972). MacKay's development is quite general, but his illustrations involve simple versions of the multilayer perceptron given in (23.1).

A different Bayesian approach to optimal design, see for example Chaloner and Verdinelli (1995), is to base D-optimal designs, say, on the averaged criterion

$$\int \log \det\{M(\xi, \theta)\} q(\theta) d\theta, \tag{23.4}$$

where $q(\theta)$ is the 'prior' density for θ. There is an equivalence theorem based on this criterion which, along standard lines, contains a check that can be applied to any design to assess its optimality or otherwise. Haines (1998) applied this approach to her simple feed-forward network. She approximated the integral criterion in (23.4) by the sample average of values of $\log \det\{M(\xi, \theta)\}$ based on a sample of values of θ drawn from $q(\theta)$ by Markov chain Monte Carlo. The approximate integral was then maximised with respect to ξ and the resulting design checked for optimality in the usual way.

4. OPTIMAL DESIGN FOR LOCAL LINEAR SMOOTHING

4.1. INTRODUCTION

As Haines (1998) points out, neural network models represent one structured way to approach nonparametric regression, and this begs the question of how to deal with the notion of optimal design within other familiar nonparametric regression paradigms. In this section we summarise the approach taken by Cheng et al. (1998) in the context of local linear smoothing with a scalar covariate. Before we embark on this, however, a brief comment is appropriate about previous contributions. Müller (1996) considers local linear fitting with a view to optimal prediction of a response at a pre-specified finite number of design points. A single optimality criterion is created by taking a weighted average, based on allegedly prescribed weights, of the G-criteria for the q design points of interest, and assuming that the regression model is locally linear at each of those points. The nature of the optimality criterion is then such that an equivalence theorem of familiar form can by written down and used to motivate a vertex-direction algorithm for computing the design. Earlier relevant work referenced by Müller (1996) includes Müller (1984) and an unpublished 1995 paper, 'Local regression: optimal allocation of observational points' by V.V. Fedorov and C.J. Nachtsheim, which has recently appeared as Fedorov and Nachtsheim (1999). From now on we concentrate on the approach of Cheng et al. (1998), but we will comment on both Müller (1984) and Faraway (1990) within that discussion.

4.2. THE METHOD OF CHENG ET AL. (1998)

Most of the optimal design work done in parametric contexts is based on the assumption that the underlying model being fitted is correct, with the result that variance, rather than bias, dominates one's strategy. By the nature of nonparametric regression, however, it is accepted that the 'model' that is effectively fitted is incorrect, with the consequence that mean squared error is the key criterion on which optimality should be based. To be more precise, Cheng et al. (1998) aimed to minimise integrated mean squared error over a compact design space. They established an empirical, asymptotically optimal sequential rule for selecting both the bandwidth and the design density for a locally linear regression estimator, with the goal of modifying a uniform design so as to put less weight in regions of low curvature. Thus, a key issue is how to estimate the curve in places of high curvature.

It is assumed that observations are generated by the model

$$Y = g(x) + \epsilon,$$

where g is the function to be estimated, ϵ has zero mean, variance σ^2 and a distribution not depending on x, and, conditional on design points $x = X_i$, the responses $Y = Y_i$ are independent. We shall suppose that the design is restricted to the interval $\mathcal{I} = [0, 1]$. The algorithm works as follows.

Given $r > 1$, let n_k denote the integer part of r^k. Estimation of g is conducted iteratively, with step k employing information derived from the previous n_{k-1} data pairs. Step k may be conveniently broken into two parts: in part (a) we determine a design density $\hat{f}_k$ from which to draw the design points for the next $N_k = n_k - n_{k-1}$ pairs; and in part (b) we obtain these new data, adjoin them to the earlier n_{k-1} pairs to produce a new set $\mathcal{X}_k = \{(X_i, Y_i),\ 1 \le i \le n_k\}$, and use $\mathcal{X}_k$ to construct estimators $\hat{g}_k$ of g and $\hat{\sigma}_k^2$ of σ^2. We compute $\hat{f}_k$ as a histogram, define $\hat{g}_k$ using local linear smoothing, and construct $\hat{\sigma}_k^2$ using first-order differences of the responses. Thus we effectively create a sequence of piecewise-constant densities, whose parameters are chosen at a given stage k to minimise an estimate Δ of the integrated mean squared error incurred by the resulting local linear curve estimator.

For details of the implementaton of parts (a) and (b), see Cheng et al. (1998). As is traditional in nonparametric curve estimation and density estimation, mean squared error is approximated by the 'asymptotic' version, consisting of a squared-bias term and a variance term, of which the former involves the second derivative of the curve or density under investigation; see equation (23.5) below. However, it turns out that the algorithm for part (a) manages to circumvent, by its choice of the estimator Δ, the troublesome, direct estimation of the second derivative of the curve. (As with the mean squared error itself, the estimator Δ consists of two terms, one representing squared bias and the other variance.)

The objective of minimising integrated mean squared error is clearly relevant to the goals of calibration or prediction. The algorithm can be justified rigorously under the following conditions: (a) the target function g has two continuous derivatives on $\mathcal{I}$, and g'' vanishes only at a finite number of points; (b) the error distribution has all moments finite, zero mean and variance σ^2; and (c) the symmetric, nonnegative kernel K, used in constructing the local linear estimator of g, is Hölder continuous and supported on a bounded interval.

Suppose n independent observations are made of a pair (X, Y) generated as $Y = g(X) + \epsilon$, in which the design variables X are distributed with a continuous density f, and the distribution of the error, ϵ, has zero mean and variance σ^2. An estimator of g based on local linear smoothing, using kernel K and bandwidth h, has its asymptotic mean squared error at $x \in \mathcal{I}$ given by

$$H_n(x, h|f) = (nh)^{-1}\kappa_1\sigma^2\, f(x)^{-1} + \tfrac{1}{4}\, h^4\, \kappa_2\, g''(x)^2\,, \tag{23.5}$$

where $\kappa_1 = \int K^2(y)dy$ and $\kappa_2 = \{\int y^2\, K(y)\, dy\}^2$; see for example Fan (1993) and Hastie and Loader (1993). For fixed x, which we now suppress, $H_n(x, h|f)$ is minimised by taking $h = h_0 = (n\,\kappa_3\, fg''^2)^{-1/5}$, where $\kappa_3 =$

$\kappa_2/(\kappa_1\sigma^2)$. Substituting back into (23.5) we deduce that, with an optimal local choice of bandwidth, mean squared error is given asymptotically by $n^{-4/5}\,\kappa_4\,(g''^2/f^4)^{1/5}$, where the constant κ_4 depends only on K and σ^2. The minimum mean integrated squared error is obtained by integrating this quantity over $\mathcal{I}$, producing a functional proportional to $A(f) = \int_{\mathcal{I}} (g''^2/f^4)^{1/5}$.

The optimal design density f is that function which minimizes $A(f)$ subject to $\int_{\mathcal{I}} f = 1$ and $f \geq 0$. A simple calculus of variations argument shows that this is given by $f_0 = c_0\,|g''|^{2/9}$, where the constant c_0 is chosen to ensure that $\int_{\mathcal{I}} f_0 = 1$. Note particularly that, for this choice, the optimal bandwidth is inversely proportional to the square of f: $h_0 = c_1\,f^{-2}$, where c_1 is a constant. Thus, in their algorithm, Cheng et al. (1998) take the bandwidth for computing $\hat{g}_k$ to vary in proportion to $\hat{f}_k^{-2}$.

In (23.5), let the bandwidth h be $h_1 = bn^{-1/5}f^{-2}$, where b is an arbitrary positive constant and f is an arbitrary continuous density, bounded away from zero on $\mathcal{I}$. If we integrate $H_n(\cdot, h_1|f)$ over $\mathcal{I}$ the contribution from the first term may be seen to equal

$$\int_{\mathcal{I}} (nh_1)^{-1}\kappa_1\sigma^2 f^{-1} = \kappa_1\,\sigma^2\,n^{-4/5}\,b^{-1}\,.$$

Note particularly that the effect of f has disappeared. This turns out to have a simplifying effect on the variance term in the formula for the estimator Δ; see Cheng et al. (1998) for details. This term in Δ is an estimate of the integral of the second term of (23.5), again with h_1 substituted for h.

A little care is needed to accommodate awkwardness caused by the fact that the function g will most likely have points of inflection, i.e. at which $g'' = 0$, thereby suggesting a zero optimal design density. Cheng et al. (1998) provide a suitable resolution of this by bounding the design density away from zero and prove that the underlying conditions and the structure of the algorithm combine to the effect that

$$\left\{\int_{\mathcal{I}} (\hat{g}_k - g)^2\right\} \Big/ \left\{\int_{\mathcal{I}} \inf_h H_n(\cdot, h|f_\eta)\right\} \to 1$$

with probability 1 as $k \to \infty$. In other words, $\hat{g}_k$ ultimately achieves the minimum integrated mean squared error, optimised over all permissible design densities.

Cheng et al. (1998) provided numerical evidence to support their work. Recall that the mean integrated squared error associated with a design f and an optimally chosen bandwidth is proportional to $A(f) = \int_{\mathcal{I}} (g''^2/f^4)^{1/5}$, and the optimal design is given by $f_0 = c_0\,|g''|^{2/9}$. Thus, $A(f_0) = \{\int_{\mathcal{I}} |g''|^{2/9}\}^{9/5}$ is the minimum achievable value of $A(f)$, and it is natural to define the efficiency of design f, relative to the optimal design, by $Eff(f, f_0) = A(f_0)/A(f)$. This

was evaluated in the context of the set of fifteen Gaussian mixtures proposed by Marron and Wand (1992). These 'true' curves vary from the standard Gaussian density function, through a range of Gaussian mixtures of varying complexity. The design space was the interval $(-3, 3)$. Although for many of the fifteen curves there was not much to be gained from using the optimal design, rather than the uniform, in some cases, especially the very 'wiggly' ones, with high curvature, the gains could be considerable.

These results were based on asymptotic formulae. A supplementary simulation study was carried out, based on a piecewise-quadratic true curve, to investigate the mutual similarity of the formula for the asymptotic, integrated mean-squared error and the corresponding estimator thereof, defined by Δ, and also to explore how closely the design created after one iteration of the algorithm approximates the optimal design. The results were reassuringly positive.

It is important to note that, although the theoretical results are asymptotic in k, in practice only a small number of iterations of the algorithm would be carried out, bearing in mind the relationships between successive sample sizes: the design procedure is best described as *batch-sequential*, with large batch sizes, and anything approaching genuinely sequential design, in which the design is updated one observation at a time, does not make sense in the context of nonparametric regression, if it is intended to use direct analogues of the type of sequential algorithm used in parametric nonlinear problems; Faraway (1990)'s design algorithm, mentioned below, is truly sequential, but has an essentially different character.

Cheng et al. (1998) say a little about the issue of design for other approaches to nonparametric regression. For example, for a Nadaraya-Watson kernel estimator the asymptotic mean squared error formula, the analogue of (23.5), may be written as

$$(nh)^{-1}\,\kappa_1\sigma^2\,f(x)^{-1} + \kappa_5\,h^4\left\{g''(x) + 2\,g'(x)\,f'(x)\,f(x)^{-1}\right\}^2,$$

where the constant κ_5 depends only on the kernel. The second term here represents the squared bias contribution to mean squared error, and may be rendered equal to zero (except at zeros of g') by defining $f = f_0 = c_0\,|g'|^{-1/2}$, where $c_0^{-1} = \int |g'|^{-1/2}$. We may estimate $|g'|$ either explicitly or implicitly, and employ the estimator in the obvious way to construct an estimator of f_0. Good performance is available, in theory, but, as Cheng et al. (1998) explain, there are practical difficulties which render the approach much less feasible than in the case of local linear regression.

Finally we comment on previous related work. Müller (1984) considers the kernel-based approach to nonparametric regression, covering both the cases of constant and locally variable smoothing, and obtains results that include the optimal design density formula indicated above, namely $f_0 \propto |g''|^{2/9}$. Faraway (1990) develops a truly sequential algorithm for this approach, choosing a new

point so as to match as well as possible two sets of quantiles, one based on the consequent design and the other based on the current estimate of the optimal design, using a kernel-based estimate of g''. Empirical performance on a number of test functions is encouraging and generally superior to using instead a uniform design, although no theory is established. Faraway (1990) also proposes and evaluates empirically a version of the method to deal with surface estimation. The method includes a way of extending the idea of matching sets of quantiles to higher dimensions.

References

Anthony, M.A. and Biggs, N. (1992). *Computational Learning Theory*. Cambridge Tracts in Theoretical Computer Science, No. 30. Cambridge: Cambridge University Press.

Baum, E.B. (1991). Neural net algorithms that learn in polynomial time from examples and queries. *IEEE Trans. Neural Networks* **2**, 5–19.

Chaloner, K. and Verdinelli, I. (1995). Bayesian experimental design: a review. *Statist. Science* **10**, 273–304.

Chang, Y.-J. and Notz, W.I. (1996). Model robust designs. In *Handbook of Statistics, Volume 13* Eds S. Ghosh and C.R. Rao, pp. 1055–1098. Amsterdam: Elsevier.

Cheng, B. and Titterington, D.M. (1994). Neural networks: a review from a statistical perspective (with discussion). *Statist. Science* **9**, 2–54.

Cheng, M.Y., Hall, P. and Titterington, D.M. (1998). Optimal design for curve estimation by local linear smoothing. *Bernoulli* **4**, 3–14.

Cohn, D.A. (1996). Neural network exploration using optimal experimental design. *Neural Networks* **9**, 1071–1083.

Fan, J. (1993). Local linear regression smoothers and their minimax efficiencies. *Ann. Statist.* **21**, 196–216.

Faraway, J.J. (1990). Sequential design for the nonparametric regression of curves and surfaces. In *Proc. 22nd Symposium on the Interface*, pp. 104–110. New York: Springer.

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

Fedorov, V.V. and Nachtsheim, C.J. (1999). Design of experiments for locally weighted regression. *J. Statist. Plan. Infer.* **81**, 363–382.

Ford, I., Kitsos, C.P. and Titterington, D.M. (1989). Recent advances in nonlinear experimental design. *Technometrics* **31**, 49–60.

Haines, L.M. (1998). Optimal designs for neural networks. In *New Developments and Applications in Experimental Design* Eds N. Flournoy, W.F. Rosenberger and W.K. Wong. IMS Lecture Notes-Monograph Series Volume 34. Hayward, CA: IMS.

Hastie, T. and Loader, C. (1993). Local regression: automatic kernel carpentry (with discussion). *Statist. Science* **8**, 120–143.

Hwang, J.-N., Choi, J.J., Oh, S. and Marks II, R.J. (1991). Query-based learning applied to partially trained multilayer perceptrons. *IEEE Trans. Neural Networks* **2**, 131–136.

Krogh, A. and Vedelsby, J. (1995). Neural network ensembles, cross validation, and active learning. In *Advances in Neural Information Processing Systems, Volume 7* Eds G. Tesauro, D.S. Touretzky and T.K. Leen, pp. 231–238. Cambridge MA: MIT Press.

Mackay, D.J.C. (1992). Information-based objective functions for active data selection. *Neural Computation* **4**, 590–604.

Marron, J.S. and Wand, M.P. (1992). Exact mean integrated squared error. *Ann. Statist.* **20**, 712–736.

Müller, H.-G. (1984). Optimal designs for nonparametric kernel regression. *Statist. Prob. Lett.* **2**, 285–290.

Müller, W.G. (1996). Optimal design for local fitting. *J. Statist. Plan. Infer.* **55**, 389–397.

Plutowski, M. and White, H. (1993). Selecting concise training sets from clean data. *IEEE Trans. Neural Networks* **4**, 305–318.

Ripley, B.D. (1994). Neural networks and related methods for classification (with discussion). *J. R. Statist. Soc.* B **56**, 409–456.

Sollich, P. (1994). Query construction, entropy, and generalization in neural-network models. *Phys. Rev.* E **49**, 4637–4651.

Titterington, D.M., (1999), Neural networks. In *Encyclopedia of Statistical Science, Update Volume 3* Eds S. Kotz, C.B. Read and D. Banks, pp. 528–535. New York: Wiley.

Wiens, D.P. (1999). Robust sequential designs for nonlinear regression. Technical Report 99.05, Statistics Centre, Univ. of Alberta.

Discussion

Linda M. Haines, University of Natal Pietermaritzburg, South Africa.

This paper provides a welcome review of work in the area of optimal design for nonparametric regression and, specifically, for hidden-layer feed-forward neural networks and for local linear smoothing. The associated literature is quite diffuse and the list of references is therefore particularly valuable.

An interesting feature to emerge from the article is that there would seem to be two very different broad approaches to optimal design for nonparametric regression. First the classical optimal design theory of Kiefer can be invoked. In this case the problem of finding an optimal design measure reduces, through a series of remarkable results, to that of finding a finite set of support points and

associated weights. In addition Equivalence Theorems can be formulated and used in order to confirm the optimality or otherwise of candidate designs. This approach is adopted in the work of, for example, Cohn (1996), Müller (1996) and Fedorov et al (1999). Second optimal design measures which are continuous can be found by formulating appropriate and rather subtle criteria and by using powerful and attractive arguments based on the calculus of variations. This latter methodology is well-illustrated in the studies by Müller (1984), Faraway (1990) and Cheng et al (1998) and, within the different but related context of robust design, by Wiens (1998).

The question I have is "Where now?". It is clear that there is much scope for developing the existing methodologies and in particular for focussing on ways of accommodating model misspecification into the construction of optimal designs. Furthermore it is worthwhile considering whether a synergy between the two very different approaches to optimal design for nonparametric models described above can somehow be identified and exploited. I would be interested to hear the author's comments on these and other directions for future research.

Additional Reference

Wiens D.P. (1998). Minimax robust designs and weights for approximately specified regression models with heteroscedastic errors. *J. Amer. Statist. Assoc.* **93**, 1440–1450.

Chapter 24

ON OPTIMAL DESIGNS FOR HIGH DIMENSIONAL BINARY REGRESSION MODELS

Ben Torsney
Department of Statistics, University of Glasgow
G12 8QW, Scotland, United Kingdom
b.torsney@stats.gla.ac.uk

Necla Gunduz
Department of Statistics, University of Gazi
06500, Ankara, Turkey
gunduz@quark.fef.gazi.edu.tr

Abstract We consider the problem of deriving optimal designs for generalised linear models depending on several design variables. Ford, Torsney and Wu (1992) consider a two parameter/single design variable case. They derive a range of optimal designs, while making conjectures about D-optimal designs for all possible design intervals in the case of binary regression models. Motivated by these we establish results concerning the number of support points in the multi-design-variable case, an area which, in respect of non-linear models, has uncharted prospects.

Keywords: binary response models, binary weight functions, D-Optimal, generalized linear model, weighted linear regression.

Introduction

We consider the topic of local D-optimal designs for binary regression models depending on several design variables. In particular we focus on the number of support points of these designs.

Optimal design for binary data have been studied extensively in the literature: see Chernoff (1979), Abdelbasit and Plackett (1983) and Sitter and Wu (1993). Locally optimal designs have also been considered for generalized

linear models which include many of those commonly used for binary data by Ford, Torsney and Wu (1992). They argued that, in two-parameter models, many locally D-optimal designs have two support points and they conjectured what these would be. Sitter and Torsney (1995a, 1995b) extended this work to more than two design variables for some typical binary response models. Their results too are 'conjecture' based. As a first step to establishing formal results we focus on the number of support points of these designs.

The structure of the paper is as follows: first in Section 1 we briefly review the two parameter case, which relies crucially, on a parameter dependent transformation (conceived by Ford, Torsney and Wu, 1992) of the non-linear design problem to a weighted linear design problem; in Section 2 we consider the general case of several design variables and in Section 3 we establish results about the number of support points of D-optimal designs.

1. TWO PARAMETER CASE

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

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

So, $u \sim Bi(1, \pi(x))$. 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 $z = \alpha + \beta x$ for unknown parameters α, β. This is an example of a generalized linear model.

For the above model the information matrix of a single replicate can be written as

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

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

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

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 following 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 $F(\cdot)$. We now consider the implications of this.

A geometrical approach to the construction of D-optimal designs is sometimes useful. We therefore introduce the following set:

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

Box and Lucas (1959) called this the **design locus**. The support of D-optimal designs depends heavily on the geometry of this design locus (see Silvey, 1980). We know from Carathéodory's theorem that there must exist a design with 2 or 3 support points. Silvey's geometrical characterization can provide us with some insights about what these are or, at least, their number. The support points are the points of contact between $G(\mathcal{Z})$, 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).

The objective of Ford, Torsney and Wu (1992), Torsney and Musrati (1993), Sitter and Torsney (1995a) and Gunduz (1999) is to find D-optimal designs for all possible interval subsets $\mathcal{Z} = [a,b]$ $(= I_{ab})$ of I_w, where I_w is the widest possible design interval for z; namely the sample space of $f(z)$.

We note that for many choices of $F(\cdot)$, $g_1 = \sqrt{w(z)}$ and $g_2 = z\sqrt{w(z)}$ are both bounded functions for all z so that there need be no restrictions on z. Further $G(\mathcal{Z}_w)$ is a closed convex curve 'beginning and ending at the origin'.

This also happens for other standard weight functions e.g. $w(z) = f(z)$ or $f^2(z)$ where $f(z)$ is a density function.

Additionally all of the weight functions $w(\cdot)$ considered have similar properties. In particular they are typically unimodal with one maximal turning point at z_{max} ($w'(z_{max}) = 0$), $w'(z) > 0$ if $z < z_{max}$ and $w'(z) < 0$ if $z > z_{max}$ and $w(z) \to 0$ at the extremes of Z_w.

It is clear from plots of $G(\mathcal{Z}_w)$ that Silvey's minimal ellipsoid can only touch $G(\mathcal{Z}_w)$ at two points for many choices of $F(z)$. We formalise this later.

For the moment suppose the two support points are $c,\ d$ ($c < d$). For this to be D-optimal on this support the weights must be $1/2, 1/2$. Denote this design by ξ:

$$\xi = \begin{pmatrix} z & c & d \\ p_z & \frac{1}{2} & \frac{1}{2} \end{pmatrix}. \tag{24.1}$$

2. $K \geq 3$ PARAMETER CASE

2.1. RECTANGULAR Z

In multiple binary regression, we generally consider a model in which an observed value u 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 $\mathcal{X} \subset R^l$. The outcome is binary, with probabilities

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

Namely, $u \sim Bi(1, \pi(\underline{x}))$. 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

$$\begin{aligned} \mathbb{E}(u|\underline{x}) &= \pi(\underline{x}) = F(\alpha + \beta_1 x_1 + \cdots + \beta_l x_l), \\ \mathbb{E}(u|\underline{x}) &= \pi(\underline{x})[1 - \pi(\underline{x})]. \end{aligned}$$

The information matrix of a single replicate under the above model is

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

where $f(z) = F'(z)$.
Clearly, z_1 plays a similar role to $z = \alpha + \beta x$ in the two parameter case.

Consider the transformation

$$\begin{pmatrix} 1 \\ z_1 \\ z_2 \\ \vdots \\ z_l \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ \alpha & \beta_1 & \beta_2 & \cdots & \beta_l \\ b_{31} & b_{32} & b_{33} & \cdots & b_{3k} \\ \vdots & \vdots & \vdots & \vdots & \\ b_{k1} & b_{k2} & b_{k3} & \cdots & b_{kk} \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \\ x_l \end{pmatrix} = \mathcal{B} \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_l \end{pmatrix},$$

where $k = l+1$ and $b_{ij}, i = 3, \cdots, k,\ j = 1, \cdots, k$ are arbitrary constants to be chosen by the experimenter. 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 if $\underline{g}(\underline{z}) = \mathcal{B}\underline{v}$ then

$$\underline{g}(\underline{z}) \;=\; \frac{f(z_1)}{\sqrt{F(z_1)[1-F(z_1)]}} \begin{pmatrix} 1 \\ z_1 \\ \vdots \\ z_l \end{pmatrix}$$

and

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

So, again appealing to the invariance property of D-optimality, we consider the D-optimal linear design problem with design vectors

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

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 $w(z_1)$.

The implied design locus is

$$\begin{aligned} G \;=\; G(\mathcal{Z}) \;=\; & \{\underline{g} = (g_1, \cdots, g_k)^T : g_1 = \sqrt{w(z_1)}, g_j = z_{j-1}\sqrt{w(z_1)}, \\ & j = 2, \cdots, k,\ \ z \in \mathcal{Z}\}. \end{aligned}$$

This set G must be bounded. In practice this will be the case if $\mathcal{X}$ and hence $\mathcal{Z}$ is bounded. However $\mathcal{Z}$ need not be bounded. As in the two parameter case g_1 and g_2 can be bounded for all possible z_1 for many $w(z_1)$, but for $j \geq 3$, g_j will be unbounded if z_{j-1} is unbounded. We assume without loss of generality that $-1 \leq z_t \leq 1$, $t = 2, \cdots, l$, in the sense that any other bounded design locus $\tilde{G}$

can be linearly transformed to a subset of

$$G_w = G(\mathcal{Z}_w) = \Big\{\underline{g} = (g_1, \cdots, g_k)^T : g_1 = \sqrt{w(z_1)}, g_j = z_{j-1}\sqrt{w(z_1)}, \\ j = 2, \cdots, k, \qquad z \in \mathcal{Z}\Big\},$$

with

$$\mathcal{Z}_w = \{(z_1, \cdots, z_l) : z_1 \in I_w, \qquad -1 \le z_t \le 1, \quad t = 2, \cdots, l\}.$$

In particular we consider

$$G = G_{ab} = \Big\{\underline{g} = (g_1, \cdots, g_k)^T : g_1 = \sqrt{w(z_1)}, \quad g_j = z_{j-1}\sqrt{w(z_1)}, \\ z_1 \in I_{ab}, \quad -1 \le z_j \le 1, \quad j = 2, \cdots, k\Big\}.$$

To uncover conclusions about D-optimal designs it helps to consider the case $k = 3$. We note that for many choices of $F(z)$, the design locus $G(Z_w)$ is a vertical surface with, for a given z_2, the set $\{\underline{g} \in G : z_2 = \text{fixed}\}$ forming a closed curve. If $z_2 = 0$ it is the same closed convex curve as in the two parameter case. If $z_2 = \pm 1$ it is the upper or lower ridge to $G(Z_w)$. These arguments extend to the case $G_{ab} = G\{Z(I_{ab})\}$ which corresponds to making vertical cuts to G_w at $z_1 = a,\, b$.

It is immediately clear that any ellipsoid centred on the origin containing G_w can only touch G_w on these 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_w 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, these are split equally between $z_2 = \pm 1$; see Sitter and Torsney (1995b). This means we can focus on the marginal design on z_1.

Similarly, in the k-parameter case, a D-optimal design can only assign weight to $z_t = \pm 1$, $t = 2, \cdots, l$, i.e. on the ridges of G. Further, under one design, the total weight at a value of z_1 is distributed equally across the combinations $z_t = \pm 1$, $t = 2, \cdots, l$. Other possibilities are uniform distributions across certain subsets of these combinations e.g. sets forming a Hadamard matrix; see Sitter and Torsney (1995b). So we need only be concerned with the induced marginal design on z_1 which will also be the conditional design on a given 'ridge' consisting of design loci $\underline{g} = \sqrt{w(z_1)}(1,\, z_1,\, \pm 1, \cdots, \pm 1)$. Again we argue that $\underline{g} = A(g_1, g_2)^T$ where

$$A = \begin{pmatrix} 1 & 0 & \pm 1 & \cdots & \pm 1 \\ 0 & 1 & 0 & \cdots & 0 \end{pmatrix},$$

so that results for 2 dimensions apply; namely there must exist a conditional D-optimal design on the ridge at, at most, 3 support points. We further argue that, for many $w(\cdot)$, there will be at most two points, so that the design is of the form

$$\xi \;=\; \begin{pmatrix} z_1 & c & d \\ p_{z_1} & p_c & p_d \end{pmatrix},$$

where $p_c, p_d > 0$ and $p_c + p_d = 1$.

Again for given c and d there is an explicit solution for p_c, p_d, namely,

$$p_c^* \;=\; \frac{(k-1)(r-1) - 2 + \sqrt{(k-1)^2(r-1)^2 + 4r}}{2k(r-1)} \tag{24.2}$$

$$p_d^* \;=\; 1 - p_c^*, \tag{24.3}$$

where $r = \dfrac{w(c)}{w(d)}$.

Thus c and d must be chosen to maximize the determinant of the information matrix $\det M(\xi)$ where $c, d \in I_{ab}$ and $c < d$. We return later to the choice of these values.

2.2. POLYHEDRAL $\mathcal{Z}$

Consider now the case of an induced design space $\mathcal{Z}$ in the form of a polyhedron. This will be the image of a rectangular $\mathcal{X}$. For example, in the case $l = 2$, suppose $\mathcal{X} = \{(x_1, x_2) : c_i \le x_i \le d_i,\ i = 1, 2\}$. Then $z_1 = \alpha + \beta_1 x_1 + \beta_2 x_2$ and $z_2 = a + bx_1 + cx_2$ are confined to a polygon $\mathcal{Z}$ with at most 6 sides. The number of sides will depend on the choice of z_1 and z_2. 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 and x_2 are the component values of a mixture of two drugs. These constraints could lead to quadrilaterals, pentagons or hexagons as the form of $\mathcal{Z}$.

As before z_2 must be bounded. We assume without loss of generality that $\mathcal{Z}$ is confined within the limits $-1 \le z_2 \le 1$, so that G is a subset of G_w. Applying Silvey's (1972) minimal ellipsoid argument it is clear that a D-optimal design can only assign weight along the ridges of G.

Consider a conditional design ξ along a ridge defined by the equation $z_2 = rz_1 + s$. Then the design locus $\underline{g} = (g_1, g_2, g_3)$ satisfies

$$\underline{g} \;=\; A(g_1, g_2)^T,$$

where

$$A \;=\; \begin{pmatrix} 1 & 0 & s \\ 0 & 1 & r \end{pmatrix}^T \;=\; (\; \underline{c} \quad \underline{m} \;),$$

with $\underline{c} = (1,\, 0,\, s)^T$ and $\underline{m} = (0,\, 1,\, r)^T$. Thus the information matrix $M_3(\xi)$ of this conditional design satisfies

$$M_3(\xi) \;=\; AM_2(\xi)A^T.$$

So there must exist a conditional D-optimal design ξ with at most 3 support points along this ridge.

This argument extends to the k-parameter case. A D-optimal design can only assign weight along the ridges of $G = G(\mathcal{Z})$ where $\mathcal{Z}$ is polyhedral. Such a ridge is defined by a relationship of the form

$$(z_2, \cdots, z_l)^T \;=\; \underline{r}z_1 + \underline{s}.$$

Hence a design locus $\underline{g}$ on this ridge satisfies

$$\begin{aligned} \underline{g} &= A(g_1, g_2)^T \\ &= \sqrt{w(z_1)}(\underline{m}z_1 + \underline{c}), \end{aligned}$$

where

$$\begin{aligned} A &= (\underline{c} \quad \underline{m}), \\ \underline{c} &= (1,\, 0,\, \underline{s}^T), \\ \underline{m} &= (0,\, 1,\, \underline{r}^T). \end{aligned}$$

So, as above, there must exist a conditional D-optimal design on this ridge with at most 3 support points. We argue that for many $w(\cdot)$ there will be at most two such points. Now we establish a sufficient condition for this.

3. ON 'MINIMAL' SUPPORT DESIGNS

To prove the above conjectures we need to confirm the requirements of the Equivalence Theorem. This requires that the following necessary and sufficient condition of Kiefer and Wolfowitz (1960) be satisfied along each ridge of G by an arbitrary design $\xi(\underline{z})$. Suppose the ridge is defined by

$$\underline{g} = \sqrt{w(z_1)}(\underline{m}z_1 + \underline{c})$$

over the interval $A \le z_1 \le B$. Then we must have

$$\begin{aligned} w(z_1)(\underline{m}z_1 + \underline{c})^T M^{-1}(\xi)(\underline{m}z_1 + \underline{c}) &\le k \quad \forall \quad A \le z_1 \le B \\ &= k \quad \text{if} \quad \xi(\underline{z}) > 0. \end{aligned}$$

Then

$$\begin{aligned} w(z_1)Q(z_1) &\leq k \quad \forall \quad A \leq z_1 \leq B, \\ &= k \quad \text{if} \quad \xi(\underline{z}) > 0, \end{aligned}$$

where $Q(z_1) = (\underline{m}z_1 + \underline{c})^T M^{-1}(\xi)(\underline{m}z_1 + \underline{c})$, a quadratic function, i.e.

$$\begin{aligned} v(z_1) = \frac{1}{k}Q(z_1) - \frac{1}{w(z_1)} &\leq 0 \quad \forall \quad A \leq z_1 \leq B, \\ &= 0 \quad \text{if} \quad \xi(\underline{z}) > 0. \end{aligned}$$

So, for an optimal design, we wish to see $v(z_1) \leq 0$ over $A \leq z_1 \leq B$. Now

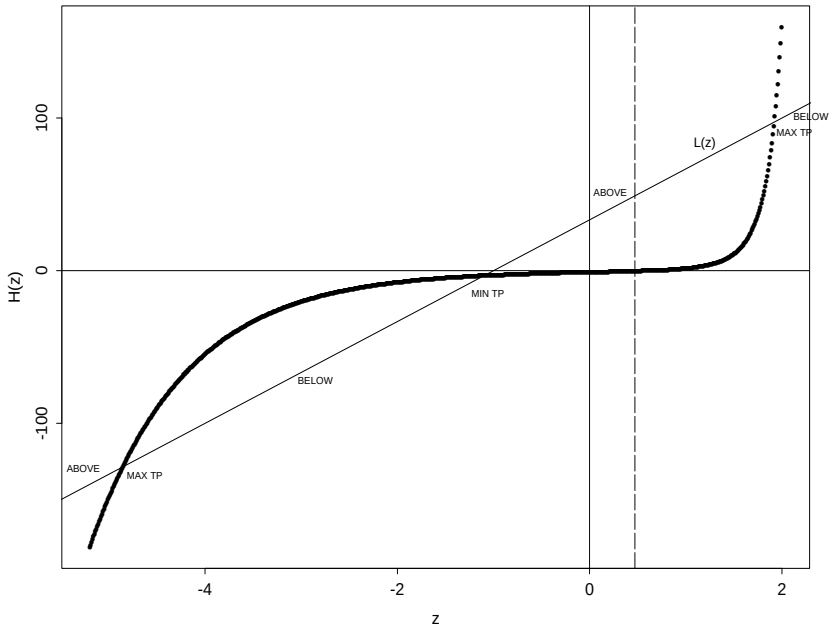


Figure 24.1 A typical plot of $H(z)$ and $L(z)$ for the complementary log-log weight function, $z_{max} = 0.47$

$$\frac{dv(z_1)}{dz_1} = L(z_1) - H(z_1),$$

where $H(z_1) = -w'(z_1)/[w(z_1)]^2$ and $L(z_1)$ is an increasing linear function of z_1 because the coefficient of z_1 is $(2\underline{m}^T M(\xi)^{-1}\underline{m})/k$. Consequently, $\frac{dv(z_1)}{dz_1} = 0$ iff $L(z_1) = H(z_1)$. That is, $\frac{dv(z_1)}{dz_1} = 0$ when the line $L(z_1)$ crosses $H(z_1)$.

For various weight functions plots of $H(z_1)$ have similar shapes and properties. In particular let $\mathcal{Z}_w = [A, B]$. Then $H(A) = -\infty$, $H(B) = +\infty$ and $H(z_1)$ is concave increasing up to some point and thereafter is convex increasing - see Fig. 24.1. Also $H'(A) = \infty$, $H'(B) = \infty$, while the second derivative of $H(z_1)$ has one change of sign for all the weight functions considered. This was observed empirically in most cases. Only a few like the logistic and the normal weight functions offer an $H(z_1)$ function whose change of sign can be seen analytically.

Given such an $H(z_1)$, an upward sloping line $L(z_1)$ can cross it, over the whole range of z_1, either one or three times. This depends on the slope of the line. This means that the derivative of $v(z_1)$ can have at most 3 zeros in $(-\infty, \infty)$. Further such a line must initially lie above $H(z_1)$. So if there is only one turning point (TP) it is a maximal one (the line crosses $H(z_1)$ from above), 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 $v(z_1)$ has only one minimum TP and at most two maximum TP's. Hence given three solutions to $v'(z_1) = 0$ the middle one must be a minimum TP. (The line crosses first from above, then from below, then from above, then from below the curve.)

Of course a further possiblity is that an upward sloping line may be tangential either to the concave or convex section of $H(z_1)$, in which case $v(z_1)$ has one maximal TP and one point of inflexion. In either case this means that a horizontal line can only cross $v(z_1)$ twice. Moreover $v(z_1)$ lies above any such line between c and d, where these are the values of z_1 at which crossing takes place. This cannot be the case if $v(z_1)$ arises under a design which is D-optimal on an interval say $[c, d]$. We must have $v(z_1) \leq 0$ on $[c, d]$.

Hence, we have established 'minimal' support D-optimal designs for many $w(z_1)$. Of course, there are some choices of $F(z)$ for which 3 support points are needed including the double exponential and double reciprocal distributions: See Ford, Torsney and Wu (1992), Musrati (1992) and Torsney and Musrati (1993).

Acknowledgements

One of the authors, Necla Gunduz, acknowledges the support of the Department of Statistics of the University of Glasgow where she did her Ph.D. She also wishes to acknowledge the financial support she received from Gazi University and the UK CVCP Overseas Research Student (ORS) Award (95 01 7020).

References

Abdelbasit, K. and Plackett, R. L. (1983). Experimental design for binary data. *Journal of the American Statistical Association*, **78** 90–98.

Box, G. E. P. and Lucas, H. L. (1959). Design of experiments in nonlinear situations. *Biometrika* **46**, 77–90.

Chernoff, H. (1979). *Sequential analysis and Optimal Design, revised edn*. Philadelphia: SIAM.

Ford, I., Torsney, B. and Wu, C. F. J. (1992). The use of a canonical form in the construction of locally optimal designs for non-linear problems. *J. R. Statist. Soc.* B **54** 569–583.

Gunduz, N.(1999). *D-Optimal Designs for Weighted Linear Regression and Binary Regression Models*. Ph.D. Thesis, Department of Statistics, University of Glasgow.

Gunduz, N. and Torsney, B. (2000). D-optimal designs for binary and weighted linear regression models: one design variable. Technical Report 00-6, Department of Statistics, University of Glasgow.

Kiefer, J. and Wolfowitz, J. (1960). The equivalence of two extremum problems. *Can. J. Math.* **12**, 363–6.

Musrati, A. K. (1992). *Locally Optimal Designs for Binary and Weighted Regression Models.* Master's Thesis, Department of Statistics, University of Glasgow.

Sebastiani, P. and Settimi, R.(1997). A note on D-optimal designs for a logistic regression model. *J. 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. Soc.* B **34**, 181–183.

Silvey, S. D. (1972). Discussion on 'Results in the theory and construction of D-optimal experimental designs '. *J. Roy. Stat. Soc.* B **34**, 174–175.

Silvey, S. D. (1980). *Optimal Design*. London & New York: Chapman and Hall.

Sitter, R. R. and Torsney, B. (1995a). D-Optimal designs for generalized linear models. In *MODA4 – Advances in Model-Oriented Data Analysis* Eds C. P. Kitsos and W. G. Müller, pp. 87–102. Heidelberg: Physica-Verlag.

Sitter, R. R. and Torsney, B. (1995b). Optimal designs for binary response experiments with two design variables. *Statistica Sinica* **5**, 405–419.

Sitter, R. R. and Wu, C. F. J. (1993). Optimal designs for binary response experiments: Fieller, D and A criteria. *Scand. J. Statist.* **20**, 329–341.

Torsney, B. and Gunduz, N. (2000). D-optimal designs for multiparameter binary regression models. Technical Report 00-7, Department of Statistics, University of Glasgow.

Torsney, B. and Musrati, A. K. (1993). On the construction of optimal designs with applications to binary response and to weighted regression models. In *Model-Oriented Data Analysis* Eds W. G. Müller, H. P. Wynn and A. A. Zhigliavsky, pp. 37–52. Heidelberg: Physica-Verlag.

Wynn, H. P. (1972). Results in the theory and construction of D-optimal experimental designs (with discussion). *J. Roy. Stat. Soc.* B **34**, 133–186.

Chapter 25

PLANNING HERBICIDE DOSE-RESPONSE BIOASSAYS USING THE BOOTSTRAP

Silvio Sandoval Zocchi and Clarice Garcia Borges Demétrio
Departamento de Ciências Exatas, ESALQ,
University of São Paulo, Piracicaba, SP, Brazil.
sszocchi@carpa.ciagri.usp.br and clarice@carpa.ciagri.usp.br

Abstract Dose response herbicide bioassays generally demand large amounts of time and resources. The choice of doses to be used is thus critical. For a given model, optimum design theory can be used to generate optimum designs for parameter estimation. However, such designs depend on the parameter values and in general do not have enough support points to detect lack of fit. This work describes the use of bootstrap methods to generate an empirical distribution of the optimum design points, based on the results of a previous experiment, and suggests designs based on this distribution. These designs are then compared to the Bayesian D-optimum designs.

Keywords: D-optimum designs, Bayesian D-optimum designs, Bootstrap, Non-linear models, Dose-response models

Introduction

The relationship between herbicide dose and plant response is of fundamental importance in understanding herbicide efficacy and its mode of action. A thorough understanding of this relationship is essential for the design and interpretation of the results of a field, greenhouse or laboratory experiment. In its simplest form, a single dose-response curve is used to quantify plant sensitivity to a herbicide. In most situations, however, scientists want to compare herbicide activities.

Successive applications of herbicide during recent decades have led to the evolutionary selection of resistant weeds, causing a problem that affects today more than 100 species in several regions of the world (Powles and Holtum, 1994). So scientists frequently want to evaluate and compare herbicide activities when applied to resistant and susceptible weeds.

The application of the herbicide haloxyfop to control *Brachiaria plantaginea*, a weed causing damage in several cultures in Brazil, for example, has given rise to resistant biotypes. In order to verify and quantify the resistance of this weed to haloxyfop, a preliminary bioassay using a theoretically resistant biotype (collected in areas where the herbicide has been extensively used) and a theoretically susceptible biotype (collected in areas where the herbicide has never been used) was conducted in a greenhouse, using 6 different doses (including the control) and 4 replicates, in a completely randomized design. The controlled percentages of resistant and susceptible *Brachiaria plantaginea* to different doses of the herbicide haloxyfop, relative to the recommended dosage, are shown in Table 25.1.

Table 25.1 Controlled percentages of *Brachiaria plantaginea* to different doses of the herbicide haloxyfop compared to the recommended dosage.

Biotype	Dose (x)	Controlled percentage			
Resistant	0.5	0	0	0	0
	1	0	0	5	0
	2	35	40	40	40
	4	100	85	80	85
	8	90	90	100	100
Susceptible	0.5	95	80	95	90
	1	100	100	100	95
	2	100	100	100	100
	4	100	100	100	100
	8	100	100	100	100

Ideally, such studies quantifying the level of resistance should employ different dose-response designs for each biotype. Additionally, to fit dose-response models to the observed data it is necessary to have doses that cover the whole range of responses from no visible effects to complete control of the weeds. These would undesirably inflate the number of different doses and, as these studies often demands large investments of time and financial resources, a bioassay with few dose levels should be utilized. So, the selection of doses is especially critical and optimum design theory can with advantage be used to generate optimum designs to estimate the parameters of the model fitted to the data.

Section 1 presents the model fitted to the *Brachiaria* data set, Section 2 describes the use of bootstrap methods to generate efficient designs based on the information given by a previous experiment and Section 3 contains some concluding remarks.

1. MODEL

The model used in the analysis of the *Brachiaria* data set was the standard logistic model

$$\eta_j(x,\theta) = \frac{1}{1+\exp[\beta(\log MD_j - \log x)]} \qquad j = R, S,$$

where x is the dose of herbicide, $\theta = (MD_R, MD_S, \beta)$ is the vector of parameters, $\eta_R(x)$, MD_R, $\eta_S(x)$ and MD_S are the proportions of weed control and the median herbicide doses (dose that controls 50% of weed) for the resistant and susceptible biotypes, respectively, and β is the slope at the inflexion point.

Table 25.2 Minimum least square and bootstrap estimates of the parameters (sample size $B = 5000$).

Estimation method	Parameters		
	β	MD_R	MD_S
Minimum least square	3.625	2.299	0.2730
Bootstrap	3.610	2.302	0.2719

Using the data presented in Table 25.1, the least square estimates of the parameters are as presented in Table 25.2. The interpretation of these estimates is very important for the agricultural researcher. For example, $MD_S = 0.2730$ means that for a control of 50% of the susceptible weeds it is necessary to use 27.3% of the recommended dose. On the other hand, $MD_R = 2.299$ means that, for a control of 50% of the resistant weeds, the required dosage is 2.299 times the recommended dose. From this we can say that the dose to control the resistant biotype is 8.4 times larger than the dose needed to control the susceptible one.

Figure 25.1 shows the fitted model and the observations; we can clearly see the lack of small doses for the susceptible biotype.

2. D-OPTIMUM DESIGNS AND THE BOOTSTRAP

The general theory of optimum experimental design has received several book-length treatments including Fedorov (1972), Silvey (1980), Pàzman (1986), Atkinson and Donev (1992) and Pukelsheim (1993).

Among several design criteria (Kiefer, 1959) we decided to work solely with the D criterion, as precise information about the parameters is of major importance. Designs obtained using this criterion are called D-optimum designs and maximize the determinant of the information matrix for the experiment, i.e.,

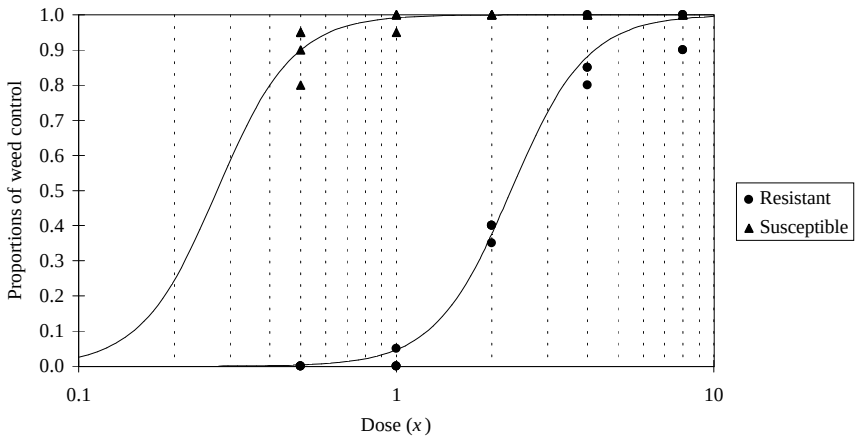


Figure 25.1 Observed and fitted (—) weed control proportions against doses of the herbicide haloxyfop, for the resistant and susceptible biotypes

minimize the generalized variance of the parameter estimates. However, they generally have insufficient support points to detect lack of fit; the design using the parameter estimates of the model adopted above only has two support points for each biotype (see Table 25.3).

Table 25.3 Locally D-optimum design, Bayesian D-optimum design based on the empirical bootstrap distribution of the parameter estimates and alternative designs

Biotype	Original		D-optimum designs Locally		Bayesian		Alternative designs $t = 3$		$t = 4$	
	x	w	x	w	x	w	x	w	x	w
	0.5	0.1	1.82	0.250	1.82	0.257	1.80	0.167	1.81	0.125
	1	0.1	2.90	0.250	2.85	0.259	1.87	0.167	1.84	0.125
Resistant	2	0.1					2.88	0.167	2.72	0.125
	4	0.1							2.92	0.125
	8	0.1								
	0.5	0.1	0.22	0.250	0.22	0.234	0.21	0.167	0.21	0.125
	1	0.1	0.34	0.250	0.35	0.250	0.30	0.167	0.25	0.125
Susceptible	2	0.1					0.36	0.167	0.33	0.125
	4	0.1							0.37	0.125
	8	0.1								
D-efficiency	14.1%		100%		99.8%		96.4%		98.2%	

A common solution for this problem is the use of Bayesian criteria (Chaloner and Verdinelli, 1995) which was not considered here. As an alternative, this

work presents the use of bootstrap methods (Efron and Tibshirani, 1993; Davison and Hinkley, 1997) to generate efficient designs.

The bootstrap is a computationally intensive technique that can be employed to generate an empirical distribution of the statistic of interest. Here our interest is in obtaining an empirical distribution of the parameters of the model and from it, an empirical distribution of locally D-optimum designs.

Briefly, the technique applied to non-linear regression models consists in generating B bootstrap samples of size n from the set of n observation pairs $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, where a bootstrap sample is, in its simplest form, a simple random sample of the values (x_i, y_i) with replacement. Despite other ways of obtaining a bootstrap sample (see Silverman, 1986; Seber and Wild, 1989; Davison and Hinkley, 1997 and Souza, 1998), Efron and Tibshirani (1993) state that sampling pairs of observations is less sensitive to the assumptions of the model.

Once the B bootstrap samples are generated, we get the estimate $\hat{\theta}^{*b}$ of θ for each sample $(b = 1, \ldots, B)$ and the corresponding D-optimum designs. Giving weight $1/B$ to each of the B estimates and their correspondent optimum designs, we get the desired empirical distributions.

From the empirical bootstrap distribution of $\hat{\theta}$, one can obtain a bootstrap estimate $\tilde{\theta}$ of θ, given by

$$\tilde{\theta} = \frac{1}{B} \sum_{b=1}^{B} \hat{\theta}^{*b}$$

and the variance covariance matrix estimate

$$\tilde{\Sigma} = \frac{1}{B-1} \sum_{b=1}^{B} (\hat{\theta}^{*b} - \tilde{\theta})(\hat{\theta}^{*b} - \tilde{\theta})^T.$$

From Table 25.1 we have 40 pairs of values (x_i, y_i) $(i = 1, \ldots, 40)$, where the first 20 are related to the resistant biotype and the last 20, to the susceptible one. Following the bootstrap procedure described in Efron and Tibshirani (1993), $B = 5000$ bootstrap samples with size $n = 40$ were generated, 20 from each biotype (see Figure 25.2), in order to obtain the desired distributions. From Table 25.2 we can see that the bootstrap and least squares estimates are very similar.

Considering the empirical bootstrap distribution of $\hat{\theta}$ we get the D-optimum Bayesian design presented in Table 25.3. This design is almost identical to the locally D-optimum design, again with not enough support points to detect any lack of fit of the model. To solve this problem, we propose generating a distribution of the locally D-optimum designs and, from it, alternative designs. Considering the fact that the B locally D-optimum designs have the same number (k) of design points x_i $(i = 1, \ldots, k)$ with equal weights $w_i = 1/k$ we can write a locally D-optimum design as $\xi = \{ \; x_1 \quad x_2 \quad \cdots \quad x_k \; \}$.

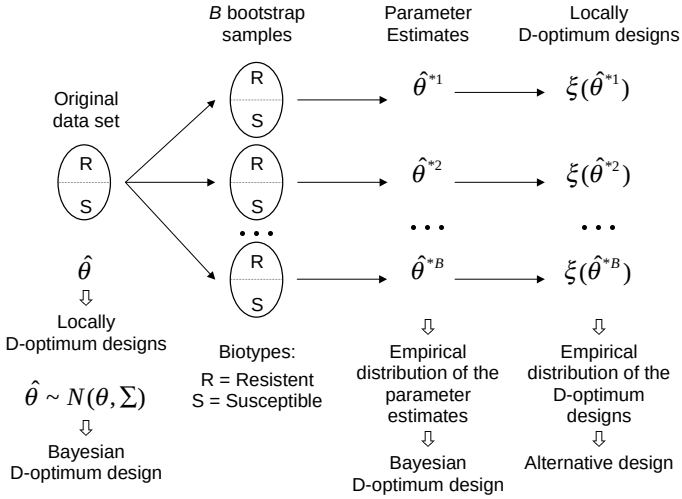


Figure 25.2 Summary of the proposed methodology.

Assume now the uniqueness of the locally D-optimum design, i.e., for all $\theta \in \Omega$, where Ω is the parameter space of θ, there is only one locally D-optimum design. This fact characterizes the design as a function of θ, which will be denoted by $\xi(\theta)$. Then, from $\{\hat{\theta}^{*1}, \ldots, \hat{\theta}^{*B}\}$, we get $\{\xi(\hat{\theta}^{*1}), \ldots, \xi(\hat{\theta}^{*B})\}$ and the joint empirical bootstrap distribution of the design points (doses) $\{x^{*1}, \ldots, x^{*k}\}$ of the locally D-optimum designs.

Finally, based on this distribution, we propose to join all the locally optimum design points in two groups, one for each biotype, and suggest, as the final design with $2 \times t$ points, the percentiles

$$P_{\left\{\frac{100 \times i}{t+1}\right\}} \qquad (i = 1, \ldots, t).$$

For the experiment discussed, the final proposed designs with, for example, $t = 3$ and $t = 4$ have the support points given in Table 25.3. We can see that, for the resistant biotype, the recommended doses are in the range of the doses used in the original experiment. On the other hand, for the susceptible biotype, the recommended doses are less than the smallest dose used in the original experiment. This confirms that it is necessary to use different doses for the different biotypes.

In order to compare the designs obtained, we use the following D-efficiency measure of a design ξ relative to the locally D-optimum one ξ^*

$$\left\{ \frac{|M(\xi)|}{|M(\xi^*)|} \right\}^{\frac{1}{p}},$$

given by Atkinson and Donev (1992), p.116, where $M(\xi)$ and $M(\xi^*)$ are the information matrices associated with the designs ξ and ξ^* and p is the number of parameters of the model. So, compared to the locally D-optimum design (see Table 25.3), the alternative designs have efficiencies much larger than the efficiency of 14.1% for the original design.

3. CONCLUDING REMARKS

The focus in this work is on the novel and easily implemented method of building alternative efficient designs using the bootstrap. There are, however, some points needing further development.

We have worked only with D-optimality, appropriate if precise information on the parameters of the model is of major importance. Other criteria would reflect other interests. As an example, we could find c-optimum designs to estimate the ratio MD_R/MD_S, an important relative resistance measure in horticulture.

In order to construct the Bayesian and alternative designs we have used a prior distribution derived from a previous experiment. Prior distributions can also be built from experimenters' opinions and from results in the literature. In this case we would have only to generate the support point distribution from the adopted prior.

Other ways of incorporating the prior information, for example, directly in the information matrix or by using a sequential design strategy could be entertained. However, as we do not have total control of the experimental material, we would not be comfortable to use any of these alternatives.

References

Atkinson, A.C. and Donev, A.N. (1992). *Optimum Experimental Designs*. Oxford: Oxford University Press.

Chaloner, K. and Verdinelli, I. (1995). Bayesian experimental design: a review. *Stat. Sci.* **10**, 273–304.

Davison, A.C. and Hinkley, D.V. (1997). *Bootstrap Methods and Their Application*. Cambridge: Cambridge University Press.

Efron, B. and Tibshirani, R.J. (1993). *An Introduction to the Bootstrap*. London: Chapman & Hall.

Fedorov, V.V. (1972). *Theory of Optimal Experiments*. London: Academic Press.

Kiefer, J. (1959). Optimum experimental designs (with discussion). *J. R. Stat. Soc.* B **12**, 363–66.

Pàzman, A. (1986). *Foundations of Optimum Experimental Design*. Bratislava: VEDA.

Powles, S.B. and Holtum, J.A.M. (Eds) (1994). *Herbicide Resistance in Plants: Biology and Biochemistry*. Boca Raton: Lewis.

Pukelsheim, F. (1993). *Optimal Design of Experiments*. New York: Wiley.

Seber, G.A.F. and Wild, C.J. (1989). *Nonlinear Regression*. New York: Wiley.

Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*. London: Chapman & Hall.

Silvey, S.D. (1980). *Optimal Design*. London: Chapman & Hall.

Souza, G.S. (1998). *Introdução aos Modelos de Regressão Linear e Não-linear*. Brasília: Embrapa.

Chapter 26

PHOTO GALLERY

Valeri

Anthony and Valeri

Henry

Misha

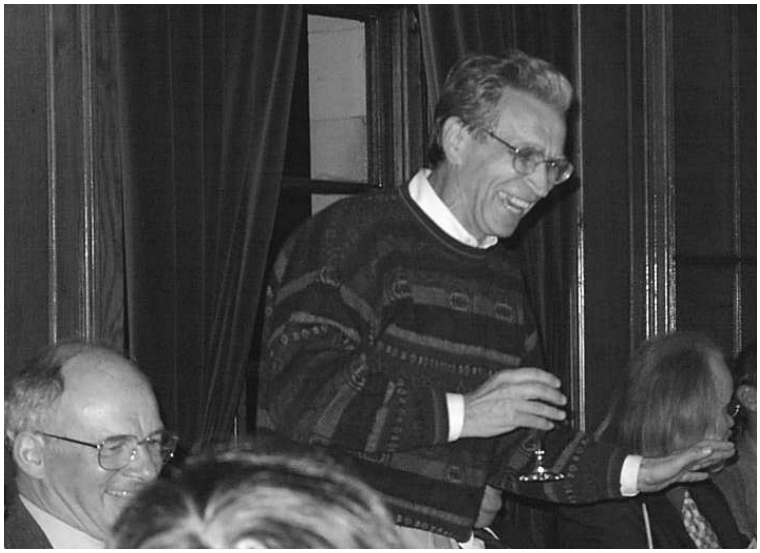
Andrej

Richard, Darryl and Janis

Henry and Ilya

Joachim and Angela

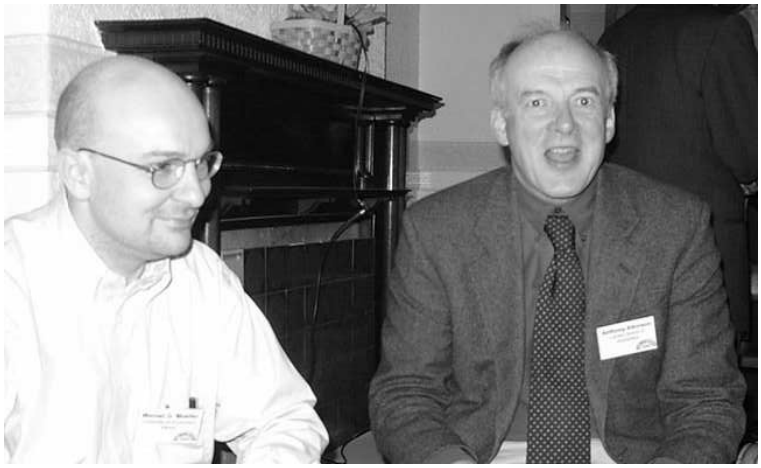
Basia and Necla



Valeri talking

Patrick

Rosemary and Janis



Werner and Anthony

Russel and Henry

OPTIMUM DESIGN 2000: LIST OF PARTCIPANTS

Kasra Afsarinejad, AstraZeneca, AstraZeneca R&D Molndal, Biostatistics, S-Molndal, Sweden, 43183
kasra.afsarinejad@astrazeneca.com

Anthony Atkinson, London School of Economics, Department of Statistics, LSE, London WC2A 2AE, UK
a.c.atkinson@lse.ac.uk

Rosemary Bailey, Queen Mary & Westfield College, School of Mathematical Sciences, Queen Mary & Westfield College, Mile End Road, London E1 4NS, UK
r.a.bailey@qmw.ac.uk

Basia Bogacka, Queen Mary & Westfield College, School of Mathematical Sciences, Queen Mary & Westfield College, Mile End Road, London E1 4NS, UK
b.bogacka@qmr.ac.uk

Tadeusz Calinski, Agricultural University of Pozan, Dept of Math. And Stat. Methods, Agricultural University of Poznan, Wojska Polskiego 28, 60-637 Poznan, Poland
calinski@owl.au.poznan.pl

Russel Cheng, University of Southampton, Dept of Math., University of Southampton, Southampton, SO17 IBJ, UK
rchc@maths.soton.ac.uk

Ching-Shui Cheng, University of California, Berkeley, Department of Statistics University of California Berkeley CA, USA 94720-3860
cheng@stat.berkeley.edu

Angela Dean, Ohio State University, Department of Statistics, The Ohio State University, Columbus, Ohio 43210 USA
dean.9@osu.edu

Darryl Downing, SmithKline Beecham Pharmaceuticals, 706 Swedeland Road, PO Box 1539, King of Prussia, PA 19406, USA
Darryl_J_Downing@sbphrd.com

Sergey Ermakov, St. Petersburg University, Dept. of Math. and Mech., St.Peterburg University, Bibliothetshnaja sq.2, St.Peterburg, Russia, 198904
Sergej.ermakov@pobox.spbu.ru

Valeri Fedorov, Smithkline Beecham Pharmaceuticals, Statistical Sciences Department
Valeri_V_Fedorov@sbphrd.com

Nancy Flournoy, American University, 4400 Massachusetts Ave NW Washingotn DC 20016-8050, USA
flournoy@american.edu

Steven Gilmour, Queen Mary & Westfield College, School of Mathematical Sciences, Queen Mary & Westfield College, Mile End Road, London E1 4NS, UK
s.g.gilmour@qmw.ac.uk

Alessandra Giovagnoli, University of Bologna, Department of Statistics University of Bologna via belle Arti 41, Bologna, Italy 40126
giovagno@stat.unibo.it

Peter Goos, Katholieke Universiteit Leuven, Department of Applied Economics, Naamsestraat 69, B-3000 Leuven Belgium
peter.goos@econ.kuleuven.ac.be

Jeff Griffiths, Cardiff University, School of Mathematics, Senghennydd Rd, Cardiff, CF24 4YH, UK
Griffiths@cardiff.ac.uk

Heiko Grossmann, Universitaet Muenster, Psychologisches Institut IV, Fliednerstr. 21, D-48149
grossman@psy.uni-muenster.de

Necla Gunduz, University of Gazi, Department of Statistics, University of Gazi, 06500, Ankara, Turkey
gunduz@quark.fef.gazi.edu.tr

Janis Hardwick, University of Michigan, 3509 Creekside Drive, Ann Arbor, Michigan, USA 48105
jphard@umich.edu

Berthold Heiligers, University of Magdeburg, Institue of Mathematical Stochastics University of Magdeburg, Universitaets str.2, D-139016 Magdeburg, Germany
berthold.heiligers@mathematik.Uni-Magdeburg.de

Ralf-Dieter Hilgers, University of Cologne, University of Cologne Joseph-Stelzmann-Str. 9 50931 K-ln, Germany
ralf.Dieter.Hilgers@medizin.Uni-Koeln.de

Matthew Jones, Cardiff University School of Mathematics, Senghennydd Rd, Cardiff, CF24 4YH, UK
JonesMC1@cardiff.ac.uk

Joachim Kunert, Universitaet Dortmund, Fachbereich Statistik Universitaet Dortmund, D 44221 Dortmund, Germany
kunert@statistik.uni-dortmund.de

Patrick Laycock, UMIST, Manchester, Mathematics dept., UMIST, Sackville st, Manchester, M60 1QD
pjlaycock@umist.ac.uk

Jesus Lopez-Fidalgo, University of Salamanca, Spain, Dpto EstadUstica Facultad de Ciencias, Universidad de Salamanca Plaza de los Caidos, s/n Salamanca Spain 37008
fidalgo@gugu.usal.es

Michael Malyutov, Northeastern University, Math. Dept., 567 Lake Hall, Northeastern University, Boston, MA 02115, USA
MLTV@neu.edu

Saumendranath Mandal, University of Glasgow, Department of Statistics, University of Glasgow, Mathematics Buildin,g Glasgow G12 8QW, UK
smandal@stats.gla.ac.uk

Polly Martin, University of Reading, Department of Applied Statistics, University of Reading, Earley Gate, Reading RG6 6FN, UK
p.martin@reading.ac.uk

Richard Martin, Sheffield University, Sheffield S3 7RH, UK
r.j.martin@sheffield.ac.uk

Ignacio Martínez, Universidad de Almería, Spain, Dpto. Estadística y Matematica Aplicada, La Cañada, Almería, Spain, E 04121
ijmartin@ualm.es

Ilya Molchanov, University of Glasgow, Department of Statistics, University of Glasgow, Mathematics Building, Glasgow G12 8QW, UK
molchanov@stats.gla.ac.uk

Hervé Monod, INRA, France, Unité de Biométrie, INRA-Versailles, Route de Saint-Cyr, F78026 Versailles CEDEX, FRANCE
monod@versailles.inra.fr

Malcolm Morrisey, Cardiff University, School of Mathematics, Senghennydd Rd, Cardiff, CF24 4YH, UK
MorrisseyMI1@cardiff.ac.uk

Valentina Moskvina, Cardiff University School of Mathematics, Senghennydd Rd, Cardiff, CF24 4YH, UK
MoskvinaV@cardiff.ac.uk

Werner Mueller, University of Economics Vienna, Augasse 2–6, Vienna, Austria A-1090
mueller@wu-wien.ac.at

Thomas Mueller-Gronbach, Free University, Berlin, Arnimalle 2–6, 14195 Berlin, Germany
gronbach@math.fu-berlin.de

Isabel Ortiz, Universidad de Almería, Spain, Dpto. Estadística y Matematica Aplicada, La Cañada, Almería, Spain, E 04120
iortiz@ualm.es

Andrej Pazman, Comenius University, Bratislava, Dept. of Probability & Statistics, Faculty of Mathematics & Physics, Comenius University, SK 84248 Bratislava, Slovak Republic
pazman@center.fmph.uniba.sk

Antonio Ponce de Leon, University of Rio-de-Janeiro, Departamento de Epidemiologia, Universidade do Estado do Rio de Janeiro, Rua Sao Francisco Xavier, 524 - Maracana, Rio de Janeiro (RJ), Brasil, 20550-013
deleon@uerj.br

Luc Pronzato, CNRS, Laboratoire I3S, Les Algorithmes/Batiment Euclide B, 2000 route des Lucioles, BP 121, 06903 Sophia-Antipolis Cedex, France
pronzato@i3s.unice.fr

Michael Sahm, University of Heidelberg, Institute for Applied Mathematics, INF 294, 69120 Heidelberg, Germany
msahm@statlab.uni-heidelberg.de

Rainer Schwabe, University of Tubingen, University of Tubingen, Institute of Medical Biometry, Westbahnhof str. 55, 72070 Tubingen, Germany
rainer.schwabe@uni-tuebingen.de

Kalin Stoyanov, University of Hull, CPACT, Department of Chemistry, Faculty of Science and the Environment, University of Hull, Hull HU6 7RX, UK
k.stoyanov@chem.hull.ac.uk

Lieven Tack, Katholieke Universiteit Leuven, Department of Applied Economics, Naamsestraat 69, B-3000 Leuven, Belgium
lieven.tack@econ.kuleuven.ac.be

Ben Torsney, University of Glasgow, Dept of Statistics, Mathematics Bdg, 15 university Gdns, Glasgow G12 8QW, UK
bent@stats.gla.ac.uk

Martina Vandebroek, Katholieke Universiteit Leuven, Department of Applied Economics, Naamsestraat 69, B-3000 Leuven, Belgium
martina.vandebroek@econ.kuleuven.ac.be

Ivan Vuchkov, University of Chemical Technology, Sophia, Euroquality Center University of Chemical Technology and Metallurgy, 1756 Sofia, Bulgaria BG-1756
vuchkov@adm1.uctm.acad.bg

Jixian Wang, University of Dundee, Medicines Monitoring Unit, Ninewells Hospital and Medical School, Dundee, DD1 9SY
jixian@memo.dundee.ac.uk

Lynda White, Imperial College, Department of Mathematics, Imperial College, London, SW7 2BZ, UK
l.white@ic.ac.uk

Henry Wynn, University of Warwick, Department of Statistics, Coventry CV4 7AL, UK
hpw@stats.warwick.ac.uk

Alexader Zaigraev, N.Copernicus University, Torun, Faculty of Mathematics and Informatics, N.Copernicus University, Chopin str. 12/18 Torun, Poland, 87-100
alzaig@mat.uni.torun.pl

Anatoly Zhigljavsky, Cardiff University School of Mathematics, Senghennydd Rd, Cardiff, CF24 4YH, UK
ZhigljavskyAA@cardiff.ac.uk

Silvio Zocchi, ESALQ, Brazil, Departamento de Ciencias Extas ESALQ/USP, Av. Padu Dias, 11 CEP 13418 900, Piracicaba, SP, Brazil, 13418 900
sszocchi@carpa.ciagri.usp.br