MATHEMATISCHES FORSCHUNGSINSTITUT OBERWOLFACH

Tagungsbericht 43/1998

Experimental Design: Theory and Applications

November 1st–November 7th 1998

The conference was organized by Norman R. Draper (University of Wisconsin-Madison), Norbert Gaffke (Universität Magdeburg), and Friedrich Pukelsheim (Universität Augsburg), and attended by 45 participants from 14 countries.

Various topics of experimental design were presented in 35 talks, covering e.g. optimal designs for models with correlated errors, fractional factorial designs, crossover designs, blocked designs, robust designs, and Bayesian designs. Applications ranged over various fields such as chemical engineering, computer experiments, and adaptive control. Each talk was followed by intensive discussions.

Modeling quantitative structure activity relationships using stochastic processes

Markus Abt
Astra Hässle AB, Biostatistics, 43183 Mölndal, Sweden
markus.abt@hassle.se.astra.com

One of the first steps in the development of new drugs is the identification of lead compounds. Novelty in terms of chemical structure as well as high biological activity are hereby looked for. Random synthesis of compounds generally leads to very few "hits". It is therefore desirable to use the available data on tested compounds to derive models that allow to predict the potency of yet untested molecules based on their chemical structure. Large numbers of descriptors are quite often available while the number of compounds can be small to moderate. For this reason, partial least squares has become a popular analysis tool. However, the true relationships are often highly non-linear and an alternative method based on stochastic processes is therefore proposed. Selection of compounds for testing according to some "coverage" criterion appears to be of great practical importance but has not yet received much attention in practice.

Optimal design under the restriction of different observation points with applications in change-point regression

Wolfgang Bischoff and Frank Miller
Institut für Mathematische Stochastik, Universität Karlsruhe, 76128 Karlsruhe, Germany
wolfgang.bischoff@math.uni-karlsruhe.de, frank.miller@math.uni-karlsruhe.de

We consider optimal designs in the set of measures dominated by the Lebesgue measure and with bounded Lebesgue density. This is motivated by the practical restriction of different observation points: We assume that there is a minimal distance between the observation points. We handle this problem in an asymptotic way for c-optimality in this talk. The equivalence theorem for this problem allows a nice geometrical interpretation of the optimal design. We can evaluate the optimal design explicitly and this is done for an example in change-point regression. It is possible to generalize the results to other optimality criteria.

Nonlinear design problems in a chemical kinetic model

Barbara Bogacka
School of Mathematical Sciences, Queen Mary and Westfield College, London, UK
b.bogacka@qmw.ac.uk

We consider a chemical kinetic model of a consecutive reaction

$$aA \xrightarrow{k} bB.$$

The kinetic equation for the concentration of a chemical compound A is

$$-\frac{d[A]}{dt} = k[A]^{\lambda},$$

where k is a rate constant and λ denotes the order of the reaction. Both k and λ are unknown parameters. The solution of the differential equation for the initial conditions of concentrations of

A and B at time $t = 0$ equal to 1 and 0 respectively, which gives a statistical model of expected observations

$$\mathrm{E}(y_i) = [1 - (1 - \lambda)kt_i]^{1/(1-\lambda)} \xrightarrow{\lambda \to 1} e^{-kt}.$$

We assume that errors of observations are independently and normally distributed random variables with constant or non-constant variance. The design problem is to find a set of points of time at which the measurements should be taken to optimize estimation of the parameters. We calculated the designs minimizing the area of confidence ellipse for the parameters. This is called D-optimality criterion and is widely used in linear models. However, in a nonlinear case it is a linear approximation of the area. So, we also considered the criterion proposed by Hamilton and Watts (1985, *Technometrics*) which gives a second order approximation of the area and takes also into account the curvature of the model. Furthermore, we calculated D-optimum designs for this model assuming non-constant variance as well as Bayesian D-optimum designs assuming prior normal distribution for the parameters. We discussed the results in terms of the percentage of conversion of the chemical compound A and we pointed out the importance of (sometimes taken carelessly) assumptions and optimality criteria for estimation of the parameters of the model.

Bayesian experimental design for a frequentist's analysis

Kathryn M. Chaloner
School of Statistics, 352 Classroom Office Building, 1994 Buford Avenue,
St. Paul, MN 55108, USA
kathryn@stat.umn.edu

Lindley and Singpurwalla (1991, 1993) and Etzioni and Kadane (1993) developed an approach to design for another decision maker's analysis. Their approach is that of a single decision maker designing optimally for a suboptimal estimate. Our alternative approach assumes that the designer designs to optimize the inference of the other decision maker. An important special case of our approach is when the evaluator uses a vague or improper prior distribution: this leads to Bayesian design for a frequentist analysis.

This work is joint with Chin-Pei Tsai, also of the University of Minnesota.

Blocking of fractional factorial designs

Ching-Shui Cheng
Department of Statistics, University of California, Berkley, CA 94920-3860, USA
cheng@stat.berkley.edu

Blocking is an effective method for improving the efficiency of an experiment by reducing the experimental error due to differences among the experimental units. One has to choose k independent interactions (called treatment defining effects) to define a $1/2^k$ fraction of a design with n two-level factors. To divide the 2^{n-k} treatment combinations into 2^p blocks, p more independent interactions (called block defining effects) are required. Once the k treatment defining effects and the p block defining effects have been chosen, the design can be constructed easily. But the issue of how to choose these $k+p$ independent interactions and to measure the "goodness" of the resulting design has not been properly addressed. Minimum aberration (Fries and Hunter, 1980, *Technometrics*) is a popular criterion for constructing good unblocked fractional factorial designs. It will be shown that several recent attempts to extend Fries' and Hunter's criterion to blocked designs are not entirely satisfactory. A major difficulty is due to the different roles played by the treatment and block defining effects. An alternative criterion will be proposed and shown to lead to designs with desirable properties.

Equivalence of fractional factorial designs

Angela M. Dean
Statistics Department, The Ohio State University, 1958 Neil Avenue,
Columbus, Ohio 43210, USA
amd@stat.ohio-state.edu

Two designs for a fractional factorial experiment are equivalent if one can be obtained from the other by reordering the treatment combinations, relabeling the factors and relabeling the factor levels. Designs are viewed as a set of points in n-dimensional space, where n is the number of treatment combinations in the design. In this setting, two designs are equivalent if the Hamming distances between the points are the same in all possible dimensions. This representation leads to an algorithm which detects distinct designs without a complete search of all relabelings and reorderings in the fraction. In addition, if two designs are equivalent, the algorithm gives a set of permutations which map one design to the other.

Convex optimal designs for compound polynomial extrapolation

Holger Dette
Institut für Mathematik, Ruhr-Universität Bochum, 44780 Bochum, Germany
holger.dette@ruhr-uni-bochum.de

In this work, the extrapolation design problem for polynomial regression models on design space $[-1,1]$ is considered when the degree of the underlying polynomial model is with uncertainty. The model robust extrapolation design criteria proposed in Dette and Wong (1996) is adopted here to find the corresponding optimal designs. It is usually not easy to find the optimal designs analytically for the general criteria above; we investigate here optimal compound extrapolation designs with two specific polynomial models, i.e. those with degrees $\{m, 2m\}$. We prove that to extrapolate at a point z, $|z| > 1$, the optimal convex combination of the two optimal extrapolation designs $\{\xi_m^*(z), \xi_{2m}^*(z)\}$ for each model separately is an optimal compound extrapolation design to extrapolate at z. The results

are applied to derive nonlinear extremal properties for polynomials which generalize classical properties of the Chebyshev polynomials of the first kind.

On the construction of experimental designs for correlated data

John Eccleston
Department of Mathematics, The University of Queensland, Brisbane QLD 4213, Australia
jae@maths.uq.edu.au

The design of experiments when the data are correlated is an area of increasing research interest, particularly since methods for the analysis of experiments which incorporate a correlation structure for errors or over the plots are becoming more generally applied. In this talk the design of experiments of a two-dimensional layout (rows and columns) with a spatial process is investigated. Computational algorithms which search for optimal designs for experiments with correlated data are presented. Since in general the correlation structure is unknown or only partially known, the robustness of the designs produced by the algorithms is examined. A number of examples are considered which illustrate the effect of correlation over the plots on the design layout and indicate that properties such as neighbor balance, corner and edge plot balance, and diagonal neighbors are important but to varying degrees and depending on particular circumstances. Two practical examples are given: (i) a field trial, and (ii) an early generation variety trial.

Theory and applications of uniform designs

Kai-Tai Fang
Department of Mathematics, Hong Kong Baptist University, Kowloon Tong,
Hongkong, China
ktfang@hkbu.edu.hk

The uniform design is one experimental design method. It can be applied in design of computer experiments and factorial designs. In this talk a brief introduction to the uniform design will be given. Advantages of the uniform design will be discussed. Some up-to-date results are involved in the talk.

Extracting new information from crossover trial data

Subir Ghosh
Department of Statistics, University of California, Riverside, CA 92521-0138, USA
ghosh@ucrac1.ucr.edu

The crossover designs are widely used for comparing treatments in clinical trials. The standard analysis of such design is well-known. The talk will present a new analysis using a proposed model different from the standard model. It presents a method of dividing the subjects in a group into subgroups using the idea of post-stratification so that the differences in responses for treatments are very similar for subjects in a subgroup. Then the new analysis provides the inference for the treatment-by-subgroup within group interactions. The significance of such interactions shows the importance of the fact that the treatments do act differently on subjects. This fact should not be ignored in the analysis. The new analysis is illustrated with several data sets. The research is done jointly with Dr. Lisa Fairchild.

Multi-stratum response surface designs

Steven Gilmour
Department of Applied Statistics, The University of Reading, PO Box 240, Early Gate,
Reading RG6 6FN, UK
s.g.gilmour@reading.ac.uk

In many industrial and laboratory based experiments there are some factors which are difficult to control, e.g. where they require a reconfiguration of the equipment. To overcome this problem, these factors can be varied less often than the others, e.g. if there are two runs per day, their levels maybe set only once per day. It is recognized that this restriction to the randomization has implications for the analysis of data from such experiments—they have a multi-stratum structure, the simplest example of which is split-plot design. When trying to fit response surface models, the use of an unbalanced split-plot structure may be unavoidable. The data from such structures can now be routinely analyzed using the widely available routines for mixed models. However, little attention has been paid to how to design such experiments. In this paper, some suggestions will be made as to how these experiments can be designed to ensure the best possible use of information at all strata in the analysis. An algorithm for producing multi-stratum designs will be presented, which makes use of a generalization of the contrast scores.

Experimental design for polynomial spline regression

Berthold Heiligers
Institut für Mathematische Stochastik, Universität Magdeburg,
39016 Magdeburg, Germany
berthold.heiligers@mathematik.uni-magdeburg.de

We discuss design aspects in d-th degree polynomial spline regression with prescribed knots over a compact regression interval. The B-basis provides an advantageous parameterization, ensuring equivariance of designs under rescaling of the setup. Numerical results for Kiefer's ϕ_p-criteria indicate that highly efficient designs are among those supported by the alternation points of the equi-oscillating spline.

Experimental design and analysis of a study on disinfection mechanisms of wastewater marine environment by using Escherichia coli as an index

Mong-Na Lo Huang
Department of Applied Mathematics, National Sun Yat-sen University,
Kaohsiung, Taiwan, 80424 R.O.C.
lomn@math.nsysu.edu.tw

In a study on disinfection mechanisms of wastewater marine environment, rotatable central composite designs are employed to investigate the effects of solar intensity, salinity, mixing ratio of wastewater and sea water, and existence of predatory microorganisms in wastewater, to the rate of declination of Escherichia coli (E. coli in short). More explicitly, under each factor combination the numbers of E. coli are measured at the beginning of each experiment and every half hour later on for two hours, then the effects of the factors to the rate of declination of E. coli are estimated through statistical analysis. Issues related to the experimental design and statistical analysis arose in the study are discussed, and complementary designs of experiments are considered for more accurate estimation of the rate of declination of the E. coli under different environment conditions.

Efficient estimation of the prevalence of multiple rare traits

Jacqueline M. Hughes-Oliver
Department of Statistics, North Carolina State University, Raleigh, NC 27695-8203, USA
hughesol@stat.ncsu.edu

When traits are rare, testing in groups can dramatically improve efficiency of estimation and cost of data collection. The optimum group size, however, is a function of the unknown prevalence of the trait. I use an adaptive strategy to overcome this deficiency, and I use the compound D-optimum criterion to develop a framework for estimating the prevalences of multiple rare traits. The limiting behavior of the stochastic group sizes is discussed and the resulting impact on the prevalence estimators presented. As a consequence, optimum designs are obtained for nonlinear models applied to heteroscedastic binary data that are correlated. These methods are applied to the problem of estimating the prevalences of HIV and chlamydia in Ethiopian women.

Optimal designs for estimating continuous functions

Lorens Imhof
Institut für Statistik, RWTH Aachen, 52056 Aachen
imhof@stochastik.rwth-aachen.de

Suppose a regression function $f \in C[-1,1]$ to be estimated. If f is known to be a polynomial of a given degree, it is often natural to use the least squares estimators associated with the G-optimal designs derived by Guest. If f is not known to be a polynomial, one may argue that f can be approximated by polynomials, so that it is reasonable to use the same design and estimators. This amounts to estimating the unknown function by certain Lagrange interpolation polynomials. However, a fundamental result of approximation theory implies that even though there are polynomials that are close to f, the interpolation polynomials need not be close to f.

In this talk, a new estimator is introduced, which accommodates this fact. The estimator is based on Hermite rather than Lagrange interpolation polynomials. This leads to new design problems. Two optimality criteria are suggested and optimal designs are given.

Exact D_K-optimal designs for quadratic regression

Olaf Krafft
Institut für Statistik, RWTH Aachen, 52056 Aachen, Germany

For quadratic regression $y(x) = a_0 + a_1 x + a_2 x^2$ on $[-1,1]$ exact D-optimal designs are presented for the six elementary cases of subparameter systems, namely (a_0), (a_1), (a_2), (a_0, a_1), (a_0, a_2), and (a_1, a_2). Although the corresponding designs are of quite different structure, it is aimed (but until to date successlessly) to base the proofs of optimality on a general inequality.

On the determination of optimal designs for interference models

Joachim Kunert
Fachbereich Statistik, Universität Dortmund, 44221 Dortmund, Germany
kunert@statistik.uni-dortmund.de

This paper generalizes Kushner's (*Annals of Statistics* 1997) method for finding optimal repeated measurements designs to the case of interference models. In the model considered here we assume that there are different left and right neighbor effects. The paper shows the usefulness of Kushner's approach, as it avoids the use of matrix inverses. The optimal designs which can be determined, however, may need many blocks. The practical use of the results lies in the determination of lower bounds for the optimality criteria. Additionally, the form of the optimal designs may suggest efficient smaller designs.

Assessing the robustness of crossover designs to subjects dropping out

Sue M. Lewis
Department of Mathematics, University of Southampton, Southampton SO17 1BJ, UK
sml@maths.soton.ac.uk

In some crossover experiments, particularly in medical applications, subjects may fail to complete their sequences of treatments for reasons unconnected with the treatments received. A method is described of assessing the robustness of a planned crossover design, to subjects leaving the study prematurely. The method involves computing measures of efficiency for every possible design that can result, and using summaries of these measures to choose between competing designs. To make the computations manageable an equivalence relation amongst the implemented designs can be exploited. Examples of designs for studies involving four treatments and four periods will be discussed that are more robust to subjects dropping out in the final period than those currently favored in medical trials. Questions for future research will be raised.

Designing computer experiments

Dennis K.J. Lin
Dept. of Management Science and Information Systems, 314 Beam Building,
Pennsylvania State University, University Park, PA 16802-1913, USA
dkl5@psu.edu

Computer models can describe complicated physical phenomena such as performance characteristics of integrated circuits. However, to use these models for scientific investigation, their generally long running times and mostly deterministic nature require specially designed experiments. Standard factorial designs are inadequate; in the absence of one or more main effects, their replication cannot be used to estimate error but instead produces redundancy.

This talk presents a new class of designs developed from the rotation of a factorial design. These rotated factorial designs are very easy to construct and preserve many of the attractive properties of standard factorial designs: they have equally-spaced projections to univariate dimensions and uncorrelated regression effect estimates (orthogonality). They also rate comparably to maximum Latin hypercube designs by the minimum inter-point distance criterion used in the latter's construction.

Numerical computation of optimal experimental designs for constrained least-squares problems

Thomas Lohmann
Technische Universität München, Zentrum Mathematik, 80355 München, Germany
lohmann@appl-math.tu-muenchen.de

An optimal experimental design approach to constrained least-squares problems is proposed which determines optimal frequency factors for measurements to be taken at given observation points. These frequency factors can be computed cheaply by optimization methods of higher order using semi-analytic differentiation of an approximation of the covariance matrix of the parameter estimates with respect to the frequency factors. The design approach is illustrated by an application from biology, the fast fluorescence rise of photo-synthesis.

Design and analysis in the presence of trends

Dibyen Majumdar
Dept. of Mathematics, Statistics, and Computer Science, m/c 249,

University of Illinois Chicago, 851 S. Morgan, Chicago IL 60607, USA
dibyen@uic.edu

This talk will consist of some recent results on nearest neighbor analysis and efficient designs when a random trend is present. These designs can also be viewed as efficient designs for the LV correlation structure. Some related situations will also be considered.

Some results on optimal designs when the observations are correlated

Richard J. Martin
Dept. of Probability & Statistics, University of Sheffield, Sheffield S3 7RH, UK
r.j.martin@sheffield.ac.uk

It is difficult to show designs are optimal under dependence unless, when all contrasts among the t treatments are of equal interest, the C-matrix is completely symmetric and has maximal trace. In general, for a block design with blocks of size k, the conditions for complete symmetry imply that the number of blocks must be a multiple of $t(t-1)/2$. Assume that the number of blocks allows such optimal designs to exist. It is then interesting to know what structure the designs have for different correlation assumptions. This is known for some simple processes, e.g. the AR(1), for any k. Here, the structures are investigated for small k (up to 6), but for a wide range of correlations, and for both ordinary and generalized (assuming the correlations are known) least-square estimation.

The analytical approach to optimal design for linear and nonlinear models

Viatcheslav B. Melas
Dept. of Mathematics & Mechanics, St. Petersburg State University, Bibliotechnaya sq. 2,
St. Petersburg, 198904, Russia
viatcheslav.melas@pobox.spbu.ru

The analytical approach consists of the consideration of optimal design points and weights as functions of some magnitudes. In the case of regression models linear in the parameters we can take some metrical characteristics of the design domain as such magnitudes, i.e. we can take the length of the segment as an independent variable. For nonlinear models the true values of the estimated parameters can be considered under the local optimality approach. In previous papers of the author two particular models—polynomial and exponential models—were considered from the functional approach point of view. Now we suggest a general theory. This theory is based on the implicit function theorem. It allows to represent optimal design points as Taylor series.

Applications of necessary and sufficient conditions for maximin efficient designs

Christine Müller
Institut für Mathematische Stochastik, Georg-August-Universität Göttingen, 37073 Göttingen, Germany
chmuelle@math.uni-goettingen.de

General sufficient and necessary conditions for minimax designs are reconsidered in a form allowing application in various optimal design problems. In combination with the Elfving theorem they are used to find maximin efficient designs for a two-dimensional linear extrapolation, and to find the optimum design for estimating the maximum point of a quadratic response function with intercept. The approach also provides an alternative proof of a recently published relation between D-optimality and maximin efficiency. Additionally, it is shown that for exponential growth curve models with one parameter, maximin efficient designs cannot be one point designs. A similar result is obtained for growth curve models with two parameters.

Statistical Problems in Finite Element Model Updating

Robert Offinger
Institut für Mathematische Stochastik, Universität Magdeburg,
39016 Magdeburg, Germany
offinger@mail.math.uni-magdeburg.de

Finite Element Models (FEM) are very common in engineering, e.g. for locating noise sources in cars. We give a very simple example of a FEM (coming from a mechanical system in dynamics) with some structural parameter ϑ, leading to the so-called structural eigenvalue problem $K(\vartheta)x_j = \lambda_j M(\vartheta)x_j$. After a few words about FEMs in general and the unknown parameter which we try to identify, we present the statistical modeling given by a nonlinear regression model $Y = \eta(\vartheta) + \varepsilon$ with $\eta(\vartheta) = (\lambda_I(\vartheta), x_{LI}(\vartheta))'$, $\varepsilon \sim (0, \sigma^2\Sigma)$ and prior distribution for ϑ.

We explain some problems concerning smoothness and regularity of $\eta(\vartheta)$ and present a formula for the Jacobian. At last we investigate some properties of methods given in engineering literature for getting a Bayesian estimator of ϑ:

$$\begin{array}{ccc} \text{prior} & \xrightarrow{\text{linearize in } \vartheta_i} & \text{posterior} \\ (\vartheta_i, B_i) & \xleftarrow[\text{iterate } i := i+1]{} & (\vartheta_{i+1}, B_{i+1}) \end{array}$$

It turns out that the idea of iterative linearization and taking the posterior as new prior does not work as intended even after a modification. Indeed we can show that the process converges to the LS-estimator in both cases.

The analysis of ranked data in blocked factorial experiments

Philip Prescott
Faculty of Mathematical Studies, University of Southampton, SO17, IBJ, UK
pp@soton.ac.uk

A non-parametric method for the analysis of blocked factorial experiments, based on ranking within blocks, is proposed and shown to be equivalent to partitioning Friedman's test statistic into a set of contrasts reflecting polynomial components of the main effects and interaction. A slightly modified version of the procedure is suggested to overcome the problem of loss of power to detect one component when the model includes other components. This alternative procedure is shown to be equivalent to applying a standard normal theory analysis of variance to the ranks. The null distributions and power comparisons are investigated using simulation methods, and it is shown that the non-parametric methods are almost as powerful as the analysis of variance.

Experimental design in adaptive optimization

Luc Pronzato
Lab. i35, Bât. Euclide, "Les Algorithmes", 2000 Route des Lucioles,
Sophia Antipolis, 06410 Biot, France
pronzato@i35.unice.fr

We consider the situation where one has to maximize a function $\eta(\bar{\theta}, \underline{x})$ with respect to $\underline{x} \in \mathbb{R}$, where $\bar{\theta}$ is unknown and estimated by least squares through observations $y_k = f(\underline{x}_k)^\top \bar{\theta} + \varepsilon_k$, with ε_k some random error. Classical applications are regulation and extremum control problems. At each step k, $\underline{x}_{k+1}$ is chosen so as to maximize the sum of the current estimated objective and a penalization for poor estimation:

$$\underline{x}_{k+1} = \arg\max_{\underline{x}} \eta(\hat{\theta}^k, \underline{x}) + \frac{\alpha_k}{k} d_k(\underline{x}),$$

with $\hat{\theta}^k$ the estimated value of $\bar{\theta}$ at step k and $d_k(.)$ the penalization function, corresponding here to the dispersion function used for the sequential construction of a D-optimal design. Sufficient conditions on $\{\alpha_k\}$ for strong consistency of $\hat{\theta}^k$ and for almost sure convergence of $\frac{1}{k}\sum_{i=1}^{k} \eta(\bar{\theta}, \underline{x}_i)$ to the maximum value of $\eta(\bar{\theta}, \underline{x})$ are derived. A classical choice for $\{\alpha_k\}$ used in adaptive control does not satisfy these conditions. An illustrative example is presented.

De la Garza phenomenon revisited

Kirti Shah
Dept. of Statistics & Acturial Science, University of Waterloo, Waterloo, ON N2L3G1, Canada
kshah@math.uwaterloo.ca

We discuss here what is known as "De la Garza phenomenon" which gives an essentially complete class of designs in many settings. The settings considered here are linear and quadratic regression with homoscedastic errors, linear regression with heteroscedastic errors, and regression with two independent variables with homoscedastic errors. While some interesting results have been obtained, the work is still in its early stages and much further work in this direction is anticipated.

Optimality study in covariates' models

Bikas K. Sinha
Stat.-Math. Division, Indian Statistical Institute, 203 BT Road, Calcutta–700035, India
bksinha@isical.ac.in

Troya (*Journal of Statistical Planning and Inference* **6** (1982), 373–419) provided a thorough study on the nature of optimal allocations in one-way ANOVA models in the presence of covariates with assignable values in closed intervals. The emphasis was on joint estimation of the factors (in the absence of the general mean) as well as of the regression coefficients involving the covariates, assuming a first order regression model.

We propose to initiate a study along similar lines on the nature of optimal designs in the presence of one or two blocking factors. We only examine the "regular" case. It turns out that the study reveals a characterization of the so-called "error functions" in a nice form.

Rotation designs: Orthogonal first-order designs with high order projectivity

David Steinberg
Department of Statistics & OR, Tel Aviv University, Tel Aviv 69978, Israel
dms@math.tau.ac.il

In many settings bias due to inexact models is a serious concern. In particular, this is true for computer experiments. Further, it may prove that a small fraction of the experimental factors account for most of the variation in the response. In these settings, it may be desirable to use a first-order orthogonal design to screen out the important factors, which can simultaneously be used to fit a higher-order model in those factors. I show how to obtain such designs by rotating standard two-level fractional factorials. Special classes of rotations are developed to achieve certain symmetry goals. Several methods of design assessment are discussed.

The theory of causal regression and its implications for experimental design

Rolf Steyer
Institut für Psychologie, Friedrich-Schiller-Universität Jena, Am Steiger 3, Haus 1,
07743 Jena, Germany
s6stro@rz.uni-jena.de

The motivation for the theory of causal regression is the distinction between the *unbiased mean* of a response variable Y in a given treatment condition x and the *conditional expected value* of Y given treatment condition x. To define these two concepts, consider the random experiment of drawing an observational unit u from the population of units, assigning it (or observing its assignment) to one of several treatment conditions x, and observing the response Y. Then the *unbiased mean (unbiased expected value) of Y in treatment condition x* may be defined as

$$\text{Mean}_u(x) := \sum_u \text{E}(Y|X = x, U = u)\,, \tag{1}$$

and the *conditional expected value $\text{E}(Y|X = x)$ of Y given treatment condition x* may be defined to be a value of the conditional expectation $\text{E}(Y|X)$. It is well-known that

$$\text{E}(Y|X = x) = \sum_u \text{E}(Y|X = x, U = u)P(U = u|X = x)\,. \tag{2}$$

Comparing the two equations reveals that the computation of $\text{Mean}_u(x)$ only involves the unconditional probabilities $P(U = u)$ whereas computation of $\text{E}(Y|X = x)$ involves the conditional probabilities $P(U = u|X = x)$. Equation (1) reveals that $\text{Mean}_u(x)$ is unaffected or unbiased by the mechanism of assigning the unit u to the treatment condition x whereas the conditional expected value $\text{E}(Y|X = x)$ *does* depend on this assignment mechanism. Equation (2) shows that the assignment mechanism (such as self selection vs. random assignment) characterized by the conditional probabilities $P(X = x|U = u)$ (and therefore also by the conditional probabilities $P(U = u|X = x)$) may seriously affect the conditional expected values $\text{E}(Y|X = x)$.

Hence, we may define the conditional expected values $\text{E}(Y|X = x)$ to be (causally) *unbiased* if $\text{E}(Y|X = x) = \text{Mean}_x(Y)$.

Comparing equations 1 and 2 it is easily seen that the conditional expected values $E(Y|X = x)$ are causally unbiased if U and X are stochastically independent which implies $P(U = u|X = x) = P(X = x)$ for any value u of U. It is also easily seen that the conditional expected values $E(Y|X = x)$ are causally unbiased if $E(Y|X = x, U = u) = E(Y|X = x)$ for any value u of U (*homogeneity*). Hence, both independence of U and X and homogeneity imply causal unbiasedness of the conditional expected values $E(Y|X = x)$.

Compound orthogonal arrays

John Stufken
Department of Statistics, Iowa State University, Ames, IA 50011, USA
jstufken@iastate.edu

Compound orthogonal arrays have recently been introduced as an alternative to Taguchi's product arrays for studying location and dispersion effects simultaneously. In this talk we will give an introduction to compound orthogonal arrays and a justification for their use. We will also discuss the existence of two-level compound orthogonal arrays with maximum possible strengths, and show how these arrays can be obtained.

Advances in construction of optimizing distributions

Ben Torsney
Dept. of Statistics, University of Glasgow, Mathematics Building, 15 University Gardens, Glasgow GI2 89W, UK
bent@stats.gla.ac.uk

We consider the following problem: Maximize a differentiable criterion $\phi(p)$ over $\mathcal{P} = \{p = (p_1, \ldots, p_n) : p_j \geq 0, \sum p_j = 1\}$. This requires the calculation of an optimizing distribution. Examples arise in optimal design, maximum likelihood estimation and stratified sampling problems. One class of algorithms for this problem has the following multiplicative form:

$$p_j^{(r+1)} \propto p_j^{(r)} f(d_j^{(r)})$$

where $d_j^{(r)} = \partial\phi/\partial p_j$ at $p = p^{(r)}$ and $f(.)$ satisfies some suitable properties and may depend on one or more free parameters.

The performance of the algorithm is first investigated in constructing D-optimal designs. In this case we can be searching for an optimal distribution on a continuous compact (design) interval or space. In practice we must discretize these spaces. The optimum design then often turns out to be a distribution defined on disjoint clusters of points. These clusters begin to 'form' early on in the above iterations. We explore extensions of the algorithm which exploit this and enjoy improved convergence rates.

This is joint work with Saumendranath Mandal.

Experimental design and accelerated life testing

Reinhard Viertl
Institut für Statistik, TU Wien, Wiedner Hauptstr. 8/107, 1040 Wien, Austria
viertl@tuwien.ac.at

Life time data in ALT have to be transformed by acceleration functions to use stress. This transformation increases imprecisions in the data (described by fuzzy numbers) dramatically. The resulting estimates are fuzzy extensions of probability distributions. For this kind of estimates it is necessary to adapt optimality criteria. Related problems are discussed in the talk. A new look at results of measurements of continuous quantities is necessary since errors and fuzziness of single measurements are different things.

Optimal exact designs on circular and elliptical arcs

Huaiqing Wu
Department of Statistics, 111C Snedecor, Iowa State University, Ames, IA 50011, USA
isuhwu@iastate.edu

Fitting a circle or an ellipse to a set of data points on a plane frequently occurs in many areas, such as manufacturing and image processing. An important practical problem is how to choose the locations of measurements on a circular or an elliptical feature. In this talk, for Berman's bivariate four-parameter circular model, all ϕ_p-criteria (with $-\infty \leq p < 1$) are shown to be equivalent. Then ϕ_p-optimal exact designs are obtained on a circle or any circular arc for any given sample size. The equivalence of ϕ_p-criteria in the circular case extends to the case of complete ellipses, in which ϕ_p-optimal exact designs are given. However, the equivalence is no longer true for the case of general elliptical arcs, for which some related issues are outlined and for which research is being conducted.

(Conference report compiled by Wolfgang Bischof and Thomas Klein.)

E-Mail Addresses of all Participants

Abt M.	markus.abt@hassle.se.astra.com
Bischof W.	wolfgang.bischof@math.uni-augsburg.de
Bischoff W.	wolfgang.bischoff@math.uni-karlsruhe.de
Bogacka B.	b.bogacka@qmw.ac.uk
Chaloner K.	kathryn@stat.umn.edu
Cheng C.	cheng@stat.berkley.edu
Dean A.	amd@stat.ohio-state.edu
Dette H.	holger.dette@ruhr-uni-bochum.de
Draper N.	draper@stat.wisc.edu
Eccleston J.	jae@maths.uq.edu.au
Fang K.	ktfang@hkbu.edu.hk
Gaffke N.	norbert.gaffke@mathematik.uni-magdeburg.de
Ghosh S.	ghosh@ucrac1.ucr.edu
Gilmour S.	s.g.gilmour@reading.ac.uk
Heiligers B.	berthold.heiligers@mathematik.uni-magdeburg.de
Huang M.	lomn@math.nsysu.edu.tw
Hughes-Oliver J.	hughesol@stat.ncsu.edu
Imhof L.	imhof@stochastik.rwth-aachen.de
Klein T.	thomas.klein@math.uni-augsburg.de
Krafft O.	-
Kunert J.	kunert@statistik.uni-dortmund.de
Lewis S.	sml@maths.soton.ac.uk
Lin D.	dkl5@psu.edu
Lohmann T.	lohmann@appl-math.tu-muenchen.de
Majumdar D.	dibyen@uic.edu
Martin R.	r.j.martin@sheffield.ac.uk
Melas V.	viatcheslav.melas@pobox.spbu.ru
Miller F.	frank.miller@math.uni-karlsruhe.de
Müller C.	chmuelle@math.uni-goettingen.de
Müller-Funk U.	funk@uni-muenster.de
Nordström K.	nordstro@noppa.helsinki.fi
Offinger R.	offinger@mail.math.uni-magdeburg.de
Pazman A.	andrej.pazman@math.uni-augsburg.de (until 2/28/99) pazman@fmph.uniba.sk (after 3/1/99)
Prescott P.	pp@soton.ac.uk
Pronzato L.	pronzato@i35.unice.fr
Pukelsheim F.	pukelsheim@uni-augsburg.de
Schwabe R.	rschwabe@zedat.fu-berlin.de
Shah K.	kshah@math.uwaterloo.ca
Sinha B.	bksinha@isical.ac.in
Steinberg D.	dms@math.tau.ac.il
Steyer R.	s6stro@rz.uni-jena.de
Stufken J.	jstufken@iastate.edu
Torsney B.	bent@stats.gla.ac.uk
Viertl R.	viertl@tuwien.ac.at
Wu H.	isuhwu@iastate.edu

Tagungsteilnehmer

Dr. Markus Abt
Lehrstuhl für Stochastik und ihre Anwendungen
Institut für Mathematik der Univ.

86135 Augsburg

Wolfgang Bischof
Institut für Mathematik
Universität Augsburg

86135 Augsburg

Wolfgang Bischoff
Institut für Mathematische Stochastik
Universität Karlsruhe

76128 Karlsruhe

Dr. Barbara Bogacka
School of Mathematical Sciences
Queen Mary and Westfield College
University of London
Mile End Road

GB-London , E1 4NS

Prof.Dr. Kathryn M. Chaloner
Department of Applied Statistics
School of Statistics
University of Minnesota
Classroom Office Building 352

St. Paul , MN 55108-6042
USA

Prof.Dr. Ching-Shui Cheng
Department of Statistics
University of California
367 Evans Hall

Berkeley , CA 94720-3860
USA

Prof.Dr. Angela M. Dean
Department of Statistics
Ohio State University
Cockins Hall 205 A
1958 Neil Ave.

Columbus , OH 43210-1247
USA

Prof.Dr. Holger Dette
Institut f. Mathematik
Ruhr-Universität Bochum
Gebäude NA

44780 Bochum

Prof.Dr. Norman R. Draper
Department of Statistics
University of Wisconsin
1210 W. Dayton Street

Madison , WI 53706
USA

Prof.Dr. John A. Eccleston
Department of Mathematics
University of Queensland

Brisbane Qld. 4072
AUSTRALIA

Prof.Dr. Kai-Tai Fang
Department of Mathematics
Hong Kong Baptist College
224 Waterloo Road

Kowloon
HONG KONG

Prof.Dr. Norbert Gaffke
Institut für Mathematische
Stochastik
Universität Magdeburg
Postfach 4120

39016 Magdeburg

Prof.Dr. Subir Ghosh
Department of Statistics
University of California

Riverside , CA 92521-0138
USA

Dr. Steven Gilmour
Department of Applied Statistics
The University of Reading
PO Box 240, Earley Gate

GB-Reading RG6 &FN

Dr. Berthold Heiligers
Fakultät für Mathematik
Otto-von-Guericke-Universität
Magdeburg
Postfach 4120

39016 Magdeburg

Prof.Dr. Mong-Na Lo Huang
Dept. of Applied Mathematics
National Sun Yat-sen University
Kaohsiung

Taiwan 80424
TAIWAN

Prof.Dr. Jacqueline Hughes-Oliver
Dept. of Statistics
North Carolina State University

Raleigh , NC 27695-8203
USA

Dr. Lorens Imhof
Institut für Statistik und
Wirtschaftsmathematik
RWTH Aachen
Wüllnerstr. 3

52062 Aachen

Thomas Klein
Institut für Mathematik
Universität Augsburg

86135 Augsburg

Prof.Dr. Olaf Krafft
Institut für Statistik und
Wirtschaftsmathematik
RWTH Aachen

52056 Aachen

Dr. Joachim Kunert
Fachbereich Statistik
Universität Dortmund

44221 Dortmund

Dr. Sue M. Lewis
Department of Mathematics
University of Southampton

GB-Southampton SO9 5NH

Prof.Dr. Dennis K.J. Lin
Dept. of Management Science
and Information Systems
Pennsylvania State University
314 Beam Building

University Park , PA 16802-1913
USA

Dr. Thomas Lohmann
Zentrum Mathematik
TU München
Arcisstr. 21

80333 München

Prof.Dr. Ulrich Müller-Funk
Institut für Wirtschaftsinformatik
Universität Münster
Steinfurterstr. 107

48159 Münster

Prof.Dr. Dibyen Majumdar
Dept. of Mathematics, Statistics
and Computer Science, M/C 249
University of Illinois at Chicago
851 St. Morgan

Chicago , IL 60607-7045
USA

Dr. Kenneth Nordström
Department of Statistics
University of Helsinki
P.O.Box 54

SF-00014 Helsinki

Prof.Dr. Richard J. Martin
Dept. of Probability and Statistics
University of Sheffield

GB-Sheffield S3 7RH

Dr. Timothy O'Brien
Offenburgerstr. 45

CH-4057 Basel

Prof.Dr. Viatcheslav B. Melas
Dept. of Mathematics and Mechanics
St. Petersburg University
Petrodvorets
Bibliotechnaya sq. 2

198904 St. Petersburg
RUSSIA

Robert Offinger
Institut für Mathematische
Stochastik
Universität Magdeburg
Postfach 4120

39016 Magdeburg

Frank Miller
Institut für Mathematische
Stochastik
Universität Karlsruhe

76128 Karlsruhe

Prof.Dr. Andrej Pazman
Dept. of Probability and Statistics
Comenius University
Mlynska Dolina

84215 Bratislava
SLOVAKIA

Dr. Christine Müller
Institut für Mathematische
Stochastik
Universität Göttingen
Lotzestr. 13

37083 Göttingen

Dr. Philip Prescott
Faculty of Mathematical Studies
University of Southampton

GB-Southampton , SO17 1BJ

Dr. Luc Pronzato
Lab. I3S, CNRS-UPRES A 6070
Bat. Algorithms
Sophia Antipolis
2000 route des Lucioles

F-06410 Biot

Prof.Dr. Friedrich Pukelsheim
Lehrstuhl für Stochastik und ihre Anwendungen
Institut für Mathematik der Univ.

86135 Augsburg

Dr. Rainer Schwabe
Fachbereich Mathematik
Universität Mainz

55099 Mainz

Prof.Dr. Kirti R. Shah
Department of Statistics and Actuarial Science
University of Waterloo
Faculty of Mathematics

Waterloo Ontario N2L 3G1
CANADA

Prof.Dr. Bikas K. Sinha
Stat-Math. Division
Indian Statistical Institute
203 Barrackpore Trunk Road

Calcutta 700 035
INDIA

Prof.Dr. David M. Steinberg
Department of Stat. and Operat. Res
Tel Aviv University

Tel Aviv 69978
ISRAEL

Prof.Dr. Rolf Steyer
Institut für Psychologie
Friedrich-Schiller Universität
Am Steiger 3, Hs.1

07743 Jena

Prof.Dr. John Stufken
Department of Statistics
Iowa State University
102 F.Snedecor Hall

Ames , IA 50011-1210
USA

Prof.Dr. Ben Torsney
Department of Statistics
University of Glasgow
University Gardens

GB-Glasgow , G12 8QW

Prof.Dr. Reinhard K.W. Viertl
Institut für Statistik und Wahrscheinlichkeitslehre
Technische Universität
Wiedner Hauptstraße 8 - 10

A-1040 Wien

Dr. Huaiqing Wu
Department of Statistics
Iowa State University
102 F.Snedecor Hall

Ames , IA 50011-1210
USA