A Bayesian Approach to Model Robust Designs

Vincent K. Agboto, PhD
Department of Family and Community Medicine
Meharry Medical College
Nashville, TN
Joint work with Professors Christopher Nachtsheim and William Li
Operations and Management Science Department
University of Minnesota

July 28, 2009

Outline

1. Introduction and Motivation
2. Optimal Experimental Design, Model Robust Optimal Design and Bayesian Optimal Design
3. Bayesian Model Robust Design (BMRD)
4. Future Research

1. Introduction and Motivation

- Typical requirements of optimal designs:
  - Model
  - Criterion
  - Construction of the designs based on the model and criterion

1. Introduction and Motivation

- Typical requirements of optimal designs:
  - Model
  - Criterion
  - Construction of the designs based on the model and criterion
- Main criticism of optimal designs: Model dependency

1. Introduction and Motivation

- Typical requirements of optimal designs:
  - Model
  - Criterion
  - Construction of the designs based on the model and criterion
- Main criticism of optimal designs: Model dependency
- Model robust optimal designs: gaining popularity recently

1. Introduction and Motivation

- Typical requirements of optimal designs:
  - Model
  - Criterion
  - Construction of the designs based on the model and criterion
- Main criticism of optimal designs: Model dependency
- Model robust optimal designs: gaining popularity recently
- Main objective: Find designs that are efficient over a class of models
  - Model estimation: Are all models estimable?
- Goal of the presentation: **Bayesian Model Robust Designs (BMRD)**

1. Introduction and Motivation

- Typical requirements of optimal designs:
  - Model
  - Criterion
  - Construction of the designs based on the model and criterion
- Main criticism of optimal designs: Model dependency
- Model robust optimal designs: gaining popularity recently
- Main objective: Find designs that are efficient over a class of models
  - Model estimation: Are all models estimable?
- Goal of the presentation: **Bayesian Model Robust Designs (BMRD)**

2. Optimal Design, Model Robust Design and Bayesian Design

- Optimal Design
- Model Robust Design
- Bayesian Optimal Design

2.1.1. A Motivating Example

- Experiment: 2^{4-1} fractional factorial experiment
- Defining relation: I = 1234
- Size: 8 runs
- Design Matrix

Run	X_1	X_2	X_3	X_4
1	-	-	-	-
2	+	-	-	+
3	-	+	-	+
4	+	+	-	-
5	-	-	+	+
6	+	-	+	-
7	-	+	+	-
8	+	+	+	+

2.1.2. Comments on Orthogonal Designs

- Pros (Many desirable properties):
  - Easy to calculate
  - Easy to interpret
  - Usually the best choice in main-effects-only models
  - Maximum Precision (in some sense)
  - Tabled designs widely available

2.1.2. Comments on Orthogonal Designs

- Cons: Not applicable if:
  - Irregular designs space
  - Mixture experiments
  - Sample size not power of 2
  - Mixed qualitative and quantitative factors
  - Fixed covariates
  - Non-linear models

2.1.3. Optimal Designs

- Optimal Experimental Design: Standard alternative when classical designs not applicable
- Choice of a particular design: Depends on the experimenter's design criterion (Optimization Problem)
- $\mathbf{y} = \mathbf{X}\boldsymbol{\theta} + \epsilon$, $\mathbf{X}_{n\times p}$:model matrix, $\boldsymbol{\theta}_{p\times 1}$:unknown, σ^2:known; $\mathbf{X}_{n\times p} = [f(x_1), ..., f(x_n)]'$
- Design matrix: $\mathbf{X}_{n\times m} = [x_1, ..., x_n]'$
- Design ξ: Probability measure over a compact region χ with $\xi(x_i) = \xi_i$
- For an exact design ξ, $n\xi(\mathbf{x})$ is an integer for any $\mathbf{x}\epsilon\chi$
- Optimality criteria: ξ^* optimizes a general measure of imprecision $\psi(\mathbf{X}'\mathbf{X})$
- MEPI$_g$ = {models with all m main effects and any g two-factor interactions}

2.1.4. Some Useful Criteria

- D-Optimality: $\max \mid \mathbf{X}'\mathbf{X} \mid$
- A-Optimality: $\min\{trace(\mathbf{X}'\mathbf{X})^{-1}\}$
- G-Optimality: $\min\{\max d(\mathbf{x})\}$ where d($\mathbf{x}$)=$f(\mathbf{x})'(\mathbf{X}'\mathbf{X})^{-1}f(\mathbf{x})$
- V-Optimality: $\min\{\text{average } d(\mathbf{x})\}$

2.1.5. Algorithms for D-Optimal Designs

- Modified Fedorov algorithm (Cook and Nachtsheim, 1980)
- K-exchange algorithm (Johnson and Nachtsheim, 1983)
- K-L exchange algorithm (Donev and Atkinson, 1989)
- Coordinate exchange algorithm (Meyer and Nachtsheim, 1995)
- CP algorithm (Li and Wu , 1997)

2.1.5. Algorithms for D-Optimal Designs

- Exchange procedure: Delete x_j and add $x \in \chi$; χ: design space
- k-exchange: Drop k ”bad points”,i.e: k-opt / iteration
- Coordinate exchange algorithm:
  - Identify “k bad points”; g coordinate groups
  - Exchange based on coordinate-group
  - kp-opt / iteration
  - Eliminate the need for candidate points

2.2.1. Introduction

- General situation: Unknown true model of known set
- Objective: Select an optimal design from candidate designs, such that it is optimal over all models in F, with respect to a criterion
- Previous work:
  - Lauter (1974), Cook and Nachtsheim (1982)
  - Cheng, Steinberg and Sun (1999), Li and Nachtsheim (2000)
  - Montgomery et al. (2003)
- Framework: Three main elements
  - Model space: $F = \{f_1, f_2, ..., f_u\}$
  - Criterion (e.g., EC, IC)
  - Candidate designs (e.g., orthogonal designs)

2.2.2. Model Robust Linear Optimal Designs

- Cook and Nachtsheim (1982)
- $$y_k(x) = f_k'(x)\theta + \epsilon \tag{1}$$
- Model space: $\mathcal{F} = \{f_0, f_1, ..., f_k\}$
- Criterion: $\overline{L}$ optimality $\Longleftrightarrow$ Minimize

$$L(\xi) = \int_I E_i^{-1}(\xi) d\beta(i) \tag{2}$$

where $E_i = \frac{L(D_i(\xi_i))}{L(D_i(\xi))}$, ξ_i the is the optimal design for f_i, $L(D) = \int_\chi f^{'}(x) D(\xi) f(x) d\lambda(x)$ and β is a probability measure on I.

2.2.3. Model Robust Factorial Designs

- Project: Reduce the leakage of a clutch slave cylinder
- Significant factors: body inner diameter, body outer diameter, seal inner diameter, seal outer diameter
- Each factor: two levels; Budget: 8 runs
- Two two-factors interactions might be active
- Standard approach: 2^{4-1} resolution IV fractional factorial design with $I = 1234$

2.2.3. Model Robust Factorial Designs

- Standard resolution IV designs: Estimation of the four main effects as well as the confounded pairs 12=34, 13=24, 14=23
- Standard resolution IV designs: Only 12 out of the 15 candidate models (80%) are estimable
- Not a model robust design
- Question: Is there a better design?

2.2.3. Model Robust Factorial Designs

- The Approach:
  - $\mathcal{F}$: Set of models having g two-factor interactions
  - ξ^* is robust for $\mathcal{F}$ if $\xi^*(n) = arg\max \sum w_i e_i(\xi(n))$
  - For Estimation Capacity (EC_g-Optimality)
    $e_i = \begin{cases} 1 & \text{if } f_i \text{ is estimable,} \\ 0 & \text{otherwise.} \end{cases}$
  - For Information Capacity (IC_g-Optimality)
    $e_i = (\frac{D_i(\xi(n))}{D_i(\xi_i^*(n))})^{\frac{1}{p_i}}$
    where $\xi_i^*(n)$ is the D-optimal design for model f_i and
    $D_i(\xi(n)) = \mid \mathbf{X}_i'\mathbf{X}_i \mid$

2.3.1. Main Requirements

- Chaloner and others
- Prior distribution $p(\boldsymbol{\theta})$

2.3.1. Main Requirements

- Chaloner and others
- Prior distribution $p(\boldsymbol{\theta})$
- Distribution of data $p(\mathbf{y} \mid \boldsymbol{\theta})$

2.3.1. Main Requirements

- Chaloner and others
- Prior distribution $p(\boldsymbol{\theta})$
- Distribution of data $p(\mathbf{y} \mid \boldsymbol{\theta})$
- Utility function $U(d, \boldsymbol{\theta}, \xi, \mathbf{y})$

2.3.1. Main Requirements

- Chaloner and others
- Prior distribution $p(\boldsymbol{\theta})$
- Distribution of data $p(\mathbf{y} \mid \boldsymbol{\theta})$
- Utility function $U(d, \boldsymbol{\theta}, \xi, \mathbf{y})$
- Design space Ξ

2.3.1. Main Requirements

- Chaloner and others
- Prior distribution $p(\boldsymbol{\theta})$
- Distribution of data $p(\mathbf{y} \mid \boldsymbol{\theta})$
- Utility function $U(d, \boldsymbol{\theta}, \xi, \mathbf{y})$
- Design space Ξ

2.3.2. Criterion

- For any design ξ, the expected utility of the best decision is given by:
  $U(\xi) = \int_{Y} max_{d \epsilon D} \int_{\Theta} U(d, \boldsymbol{\theta}, \xi, \mathbf{y}) p(\mathbf{y}, \boldsymbol{\theta} \mid \xi) d\boldsymbol{\theta} d\mathbf{y}$

2.3.2. Criterion

- For any design ξ, the expected utility of the best decision is given by:
$U(\xi) = \int_Y max_{d \epsilon D} \int_\Theta U(d, \boldsymbol{\theta}, \xi, \mathbf{y}) p(\mathbf{y}, \boldsymbol{\theta} \mid \xi) d\boldsymbol{\theta} d\mathbf{y}$
- Bayesian solution to the Optimal design: ξ^* maximizing $U(\xi)$

2.3.3. Standard Bayesian Setup

- Prior Distribution of $\boldsymbol{\theta}$: $\boldsymbol{\theta} \sim N(\boldsymbol{\theta}_0, \sigma^2 \mathbf{R}^{-1})$
- Distribution of $\mathbf{y} \mid \boldsymbol{\theta}$: $\mathbf{y} \mid \boldsymbol{\theta} \sim N(\mathbf{X}\boldsymbol{\theta}, \sigma^2 \mathbf{I})$
- Shannon utility function:

$$U(\xi) = \int \int \log p(\boldsymbol{\theta} \mid \mathbf{y}, \xi) p(\mathbf{y}, \boldsymbol{\theta} \mid \xi) d\boldsymbol{\theta} d\mathbf{y} \tag{3}$$

- Ξ : Set of probability measures on χ
- We can show that:

$$U(\xi) = (-\frac{k}{2}) \log 2\pi + \frac{1}{2} \log \det(\sigma^{-2}(n\mathbf{M}(\xi) + \mathbf{R})) - \frac{k}{2} \tag{4}$$

- The Bayesian optimal design maximizes

$$\phi_B(\xi) = \det\{n\mathbf{M}(\xi) + \mathbf{R}\} = \mid \mathbf{X}'\mathbf{X} + \mathbf{R} \mid$$

3. Bayesian Model Robust Design

- Previous Work
- Bayesian Model Robust Criterion
- Examples
- Performance of BMRDs
- Evaluation of 16-runs orthogonal arrays

3.1.1. DuMouchel-Jones Approach

- Main idea:
  - Primary terms: Assumed to be present in the true model (p_1)
  - Potential terms: May/may not be present in the model (p_2)
- Prior distribution: variance of parameters $\mathbf{R} = \frac{\mathbf{K}}{\tau^2}$ where

$$\mathbf{K} = \begin{pmatrix} \mathbf{0}_{p_1 \times p_1} & \mathbf{0}_{p_1 \times p_2} \\ \mathbf{0}_{p_2 \times p_1} & \mathbf{I}_{p_2 \times p_2} \end{pmatrix}$$

- Criterion: Maximize

$$\phi(\xi) = \left| \mathbf{X}'_\xi \mathbf{X}_\xi + \frac{\mathbf{K}}{\tau^2} \right|$$

3.1.2. The Jones, Lin, Nachtsheim Approach (2006)

- Introduction of a new class of supersaturated designs using Bayesian D-optimality
- Application of the DuMouchel-Jones criterion directly to the supersaturated design problem
- Bayesian designs compared favorably to $E(s^2)$

3.2. Bayesian Model Robust Design Criterion

3.2.1. Setup

- X_i:design matrix for the model f_i with dimension k_i
- $p(i) = w_i$: prior probability for model f_i
- Prior distribution: $\boldsymbol{\theta} \,|i \sim N(\boldsymbol{\theta}_{0i},\, \sigma^2 \mathbf{R}_i^{-1})$
- Assumption: $p(i)$ independent of $p(\mathbf{y}, \boldsymbol{\theta} \mid i, \xi)$

$$\begin{aligned}
U(\xi) &= \int \int \int \log p(\boldsymbol{\theta} \mid \mathbf{y}, i, \xi)\} p(\mathbf{y}, \boldsymbol{\theta}, i \mid \xi) d\boldsymbol{\theta} d\mathbf{y} c(di) \\
&= \int \int \int \log p(\boldsymbol{\theta} \mid \mathbf{y}, i, \xi)\} p(\mathbf{y}, \boldsymbol{\theta} \mid i, \xi) p(i) d\boldsymbol{\theta} d\mathbf{y} c(di) \\
&= \sum_{i=1}^{d} \int \int \log p(\boldsymbol{\theta} \mid \mathbf{y}, i, \xi) p(\mathbf{y}, \boldsymbol{\theta} \mid i, \xi) p(i) d\boldsymbol{\theta} d\mathbf{y}
\end{aligned}$$

- The Bayesian Model Robust Optimal design maximizes:

$$\phi_{BR}(\xi) = \sum_{i=1}^{d} w_i \log \mid (\mathbf{X}'_{i,\xi} \mathbf{X}_{i,\xi} + \mathbf{R}_i) \mid$$

3.2.2. Prior Probabilities

- Uniform Prior
  - Uniform prior probabilities on each candidate model
  - prior probabilities based on expert knowledge about the models
- Hierarchical Prior
  - Advocated by Chipman, Hamada, and Wu (1997)
  - $Conditional\ independence$ and $Inheritance$ principle
  - A vector $\boldsymbol{\delta}$ of zeros and ones having captures the importance of all the models
  - For a model (A, B, C, AB, AC, BC); Prior on $\boldsymbol{\delta} = (\delta_A, ..., \delta_{BC})$: $P(\boldsymbol{\delta}) = P(\delta_A)P(\delta_B)P(\delta_C)P(\delta_{AB} \mid \delta_A, \delta_B)P(\delta_{AC} \mid \delta_A, \delta_C)P(\delta_{BC} \mid \delta_B, \delta_C)$

3.2.3. Example

- The Bayesian Model Robust Optimal design maximizes:

$$\phi_{BR}(\xi) = \sum_{i=1}^{d} w_i \log \mid (\mathbf{X}'_{i,\xi}\mathbf{X}_{i,\xi} + \mathbf{R}_i) \mid$$

- In our examples: R_i = $R = cI$, $\sigma^2 = 1$, where I is the identity matrix and c is a real number.

3.2. Bayesian Model Robust Design Criterion

Table 1: Bayesian model robust design for $n = 8$, $m = 4$, $g = 1$, $c = 100$

1	1	1	1	-1
2	-1	-1	1	-1
3	-1	-1	-1	1
4	1	-1	1	1
5	1	1	-1	1
6	-1	1	1	1
7	-1	1	-1	-1
8	1	-1	-1	-1

3.3. Performance of BMRDs

Table 2: Criteria values for models with $n = 8$, $m = 4$, $g = 1$

c	Designs	EC	IC	BMR
100	FFD	1.00	1.00	12.48
100	MRFD	1.00	1.00	12.48
100	BMRD	1.00	1.00	12.48
10	FFD	1.00	1.00	12.55
10	MRFD	1.00	1.00	12.55
10	BMRD	1.00	1.00	12.55
1	FFD	1.00	1.00	13.18
1	MRFD	1.00	1.00	13.18
1	BMRD	1.00	1.00	13.18

3.3. Performance of BMRDs

Table 3: Criteria values for models with $n = 8$, $m = 4$, $g = 2$

c	Designs	EC	IC	BMR
100	FFD	0.80	0.80	13.37
100	MRFD	1.00	0.76	12.63
100	BMRD	1.00	0.85	13.39
10	FFD	0.80	0.80	13.90
10	MRFD	1.00	0.76	12.79
10	BMRD	0.80	0.80	13.90
1	FFD	0.80	0.80	15.07
1	MRFD	1.00	0.76	14.08
1	BMRD	0.80	0.80	15.07

3.3. Performance of BMRDs

Table 4: Criteria values for models with $n = 12$, $m = 5$, $g = 1$

c	Designs	EC	IC	BMR
100	OD	1.00	0.94	17.00
100	MRFD	1.00	0.94	17.00
100	BMRD	1.00	0.96	17.10
10	OD	1.00	0.94	17.06
10	MRFD	1.00	0.94	17.06
10	BMRD	1.00	0.96	17.16
1	OD	1.00	0.94	17.62
1	MRFD	1.00	0.94	17.62
1	BMRD	1.00	0.96	17.70

3.3. Performance of BMRDs

Table 5: Criteria values for models with $n = 12$, $m = 5$, $g = 2$

c	Designs	EC	IC	BMR
100	OD	1.00	0.89	18.96
100	MRFD	1.00	0.92	19.22
100	BMRD	1.00	0.93	19.35
10	OD	1.00	0.89	19.04
10	MRFD	1.00	0.92	19.29
10	BMRD	1.00	0.93	19.42
1	OD	1.00	0.89	19.77
1	MRFD	1.00	0.92	19.96
1	BMRD	1.00	0.93	20.07

3.3. Performance of BMRDs

- Findings:
  - BMRDs: relatively insensitive to choice of c
  - In general BMRDs provide better criteria values than FFDs and MRFDs (mostly for larger m)
  - BMRDs: often unbalanced; Lack of balance seems to be associated with increases in all three criterion values

3.4. Evaluation of 16-run orthogonal arrays

- Use of BMR to examine the model-robustness properties of all 16×7 orthogonal array (OA)
- A catalog of the 55 16×7 OA designs provided by Sun, Li, and Ye (2002)
- For each of the 55 OA designs: Compare the GWLP provided in Li, Lin, and Ye (2003) and calculate BMR for combinations with $g = 1$ or 2 and $c = 1, 10$, or 100
- OA not same in terms of GWLP
- GWLP: Most popular criterion for selecting orthogonal array designs
- Generalized minimum aberration design: GWLP is sequentially minimized
- OA: Not same in terms of BMR

3.4. Evaluation of 16-run orthogonal arrays

Table 6: Rank correlations table

	wlp	$c1g1$	$c1g2$	$c10g1$	$c10g2$	$c100g1$	$c100g2$
wlp	1	0.982	0.979	0.977	0.977	0.963	0.960
$c1g1$	0.982	1	0.995	0.972	0.964	0.956	0.940
$c1g2$	0.979	0.995	1	0.963	0.963	0.947	0.942
$c10g1$	0.977	0.972	0.963	1	0.994	0.996	0.980
$c10g2$	0.977	0.964	0.963	0.994	1	0.989	0.994
$c100g1$	0.963	0.956	0.947	0.996	0.989	1	0.982
$c100g2$	0.960	0.940	0.942	0.980	0.994	0.982	1

3.4. Evaluation of 16-run orthogonal arrays

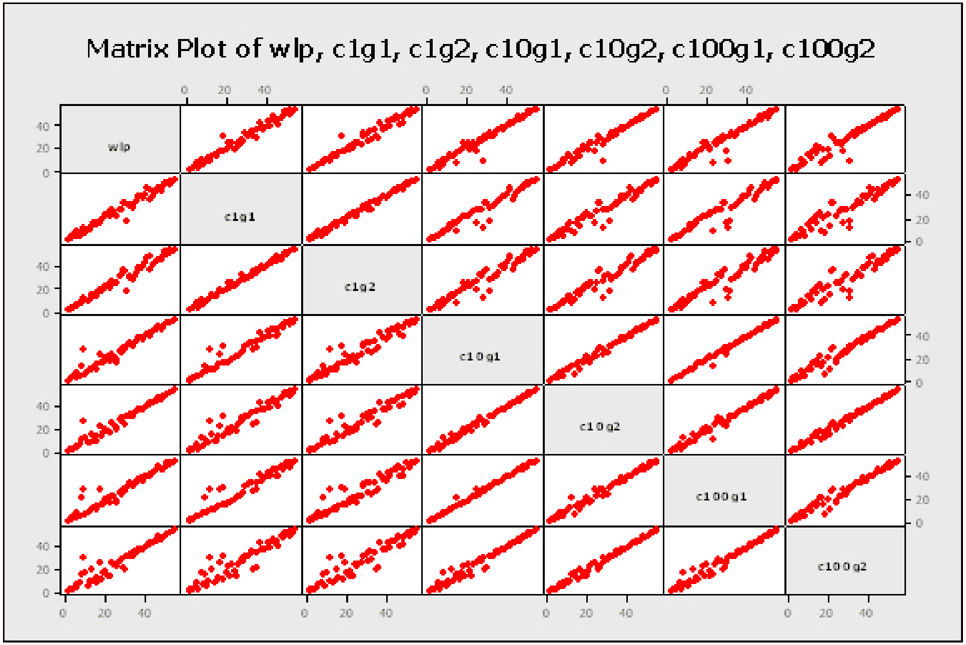

Figure 1: Scatterplot matrix of the rank of the 16-run orthogonal arrays between the GWLP and the BMRs

3.4. Evaluation of 16-run orthogonal arrays

- BMR strongly related to GWLP
- In general, the rankings by BMR and by GWLP were nearly the same
- The plots: Strong relationship between BMR and GWLP within this class of OA.
- In every case, BMRD corresponded to the generalized minimum aberration design.
- Criticism about GWLP: not a statistical criterion, simply a combinatorial index
- BMR: Clearly a statistical criterion that may capture the utility of GWLP
- BMR: Good alternative to GLWP

4. Future Research

- Useful to compare BMRDs to MFRDs on a more equal footing
- Construct supersaturated Bayesian designs with the proposed BRD criterion and compare these designs to existing supersaturated designs (DuMouchel and Jones, 1996)
- Investigate more complex cases of hierarchical Bayesian model robust designs
- Look into model discrimination issues

THANKS TO THE ORGANIZERS OF THE NEW RESEARCHER CONFERENCE. SPECIAL THANKS TO TRACY BERGERMAN.