UIC
World-Class Education.
World-Class City.

Optimal Designs for 2^k Experiments with Binary Response

Dibyen Majumdar

Mathematics, Statistics, and Computer Science
College of Liberal Arts and Sciences
University of Illinois at Chicago

Joint research with Jie Yang and Abhyuday Mandal

UIC
World-Class Education.
World-Class City.

Outline

Outline

Optimal Designs for Binary Response

Example

- Hamada and Nelder (*Technometrics* 1997) reported an experiment performed at an IIT Thompson laboratory and was originally reported by Martin, Parker and Zenick (1987) about windshield molding.
- It was a 2_{III}^{4-1} experiment with N=8000 experimental units.
- There were four factors each at two levels:
  - (A) poly-film thickness,
  - (B) oil mixture ratio,
  - (C) material of gloves, and
  - (D) the condition of metal blanks.

UIC
World-Class Education. World-Class City.

Windshield Molding Slugging Experiment

Example

- A 2_{III}^{4-1} fractional factorial design with 1000 replicates for each design point was used for this experiment.

Run	A	B	C	D	y
1	$+$	$+$	$+$	$+$	338
2	$+$	$+$	$-$	$-$	826
3	$+$	$-$	$+$	$-$	350
4	$+$	$-$	$-$	$+$	647
5	$-$	$+$	$+$	$-$	917
6	$-$	$+$	$-$	$+$	977
7	$-$	$-$	$+$	$+$	953
8	$-$	$-$	$-$	$-$	972

- Can we do something better?

Can We Do Something Better?

- Although it is common practice to use the uniform design, that is, a design with equal number of replicates for each design point, it may not be the optimal one, even when the experimenter has only limited information about the process.
- Besides, if a fractional factorial design is preferred, the experimenter may also want to know which fraction should be recommended.
- Our research indicates that nonregular fractions might be more efficient.

UIC
World-Class Education.
World-Class City.

Moreover, Uniform Designs May Not Be Available

Example

- Seo, Goldschmidt-Clermont and West (*AOAS* 2007) discussed a 2^4 experiment with unequal replicates, where the response is whether or not a particular symptom is developed in mice.

Factor Settings in Mice Example

Code	Factor	$+$	$-$
A	Age	6 week	12 week
B	Genetic factor	Wild type	ApoE-/-
C	Dietary fat intake	Chow diet	Western diet
D	Gender	Female	Male

90 Runs for 2^4 Experiment

Example

Design Layout and Sample Sizes in the Mice Example (Total = 90)

Run	A	B	C	D	Replicates	Run	A	B	C	D	Replicates
1	+	+	+	+	5	9	−	+	+	+	5
2	+	+	+	−	5	10	−	+	+	−	5
3	+	+	−	+	5	11	−	+	−	+	6
4	+	+	−	−	5	12	−	+	−	−	5
5	+	−	+	+	5	13	−	−	+	+	3
6	+	−	+	−	6	14	−	−	+	−	3
7	+	−	−	+	8	15	−	−	−	+	9
8	+	−	−	−	4	16	−	−	−	−	11

And There Are Many More Examples

Example

Nair *et al.* (*American Journal of Public Health* 2008) reported a 2_{IV}^{6-2} fractional factorial experiment used in the study called Project Quit, a project on smoking cessation. The outcome of this experiment was a binary response, answer to the question "Did you smoke a tobacco cigarette in the past 7 days?" Six factors each with two levels were studied in this project.

- First I will present the results for the simplest case, that is, 2^2 design.
- Then I will present some results for the general 2^k designs.

Outline

Preliminary Setup

- Consider a 2^k experiment, i.e., an experiment with k explanatory variables at 2 fixed levels each, for instance qualitative variables.
- Suppose n_i units are allocated to the ith experimental condition such that $n_i \geqslant 0$, $i = 1, \ldots, 2^k$, and $n_1 + \cdots + n_{2^k} = n$.
- We suppose that n is fixed and consider the problem of determining the "optimal" n_i's.
- We write our optimality criterion in terms of the proportions:

$$p_i = \frac{n_i}{n}, \ i = 1, \ldots, 2^k$$

  and determine the "optimal" p_i's.

Preliminary Setup

- Suppose η is the linear predictor that consists of main effects and interactions that are assumed to be in the model.
- The maximum likelihood estimator of β has an asymptotic covariance matrix that is the inverse of $nX'WX$, where $W = diag\left(w_1p_1, \ldots, w_{2^k}p_{2^k}\right), \, w_i = \left(\frac{d\mu_i}{d\eta_i}\right)^2 / \left(\mu_i(1-\mu_i)\right) \geq 0$ and X is the "design matrix".
- The *D-optimality* criterion maximizes

$$\left|X'WX\right|^{1/p}$$

where p is the number of parameters in the model.

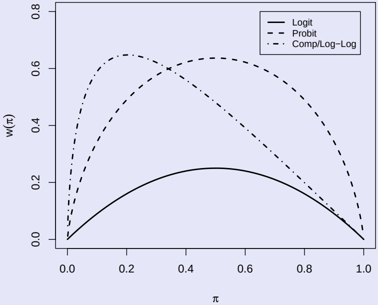
Preliminary Setup

For commonly used link functions,

$$w(\pi) = \begin{cases} \pi(1-\pi) & \text{for logit link} \\ \dfrac{\left[\phi\left(\Phi^{-1}(\pi)\right)\right]^2}{\pi(1-\pi)} & \text{for probit link} \\ \left(\frac{\pi}{1-\pi}\right)\left(\log(\pi)\right)^2 & \text{for log-log and complementary log-log link} \end{cases}$$

w for Different Link Functions

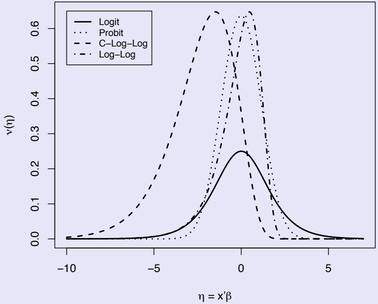
- For a binary response, we consider the commonly used link functions including logit, probit, log-log and complimentary log-log links.



plot of w versus π

w for Different Link Functions

- Given the link function g, let $\nu = \left((g^{-1})'\right)^2 / \left[g^{-1}(1-g^{-1})\right]$.
- Then $w_i = \nu(\eta_i) = \nu(\mathbf{x}_i{}'\beta)$, $i = 1, \ldots, 2^k$.



$w_i = \nu(\eta_i)$ for commonly used link functions

Preliminary Setup

- For 2^2 experiment with main-effect plan

$$X = \begin{pmatrix} 1 & -1 & -1 \\ 1 & -1 & +1 \\ 1 & +1 & -1 \\ 1 & +1 & +1 \end{pmatrix}, W = \begin{pmatrix} w_1p_1 & 0 & 0 & 0 \\ 0 & w_2p_2 & 0 & 0 \\ 0 & 0 & w_3p_3 & 0 \\ 0 & 0 & 0 & w_4p_4 \end{pmatrix}$$

- the optimization problem maximizing $|X'WX|^{1/p}$ reduces to maximizing $det(\mathbf{w}, \mathbf{p}) = 16w_1w_2w_3w_4L(\mathbf{p})$, where $v_i = 1/w_i$ and

$$L(\mathbf{p}) = v_4p_1p_2p_3 + v_3p_1p_2p_4 + v_2p_1p_3p_4 + v_1p_2p_3p_4 \qquad (1)$$

Outline

UIC
World-Class Education. World-Class City.

D-Optimal 2^2 Designs : Analytic Solution to Special Cases

Theorem

$L(\mathbf{p})$ in equation (1) has a unique maximum at $\mathbf{p} = (0, 1/3, 1/3, 1/3)$ if and only if $v_1 \geq v_2 + v_3 + v_4$.

- It corresponds to a saturated design.
- D-optimal design is saturated if and only if

$$\frac{2}{\min\{w_1, w_2, w_3, w_4\}} \geq \frac{1}{w_1} + \frac{1}{w_2} + \frac{1}{w_3} + \frac{1}{w_4}. \tag{2}$$

We call (2) the *saturation condition*.

Saturation Condition for the Logit Link

Theorem

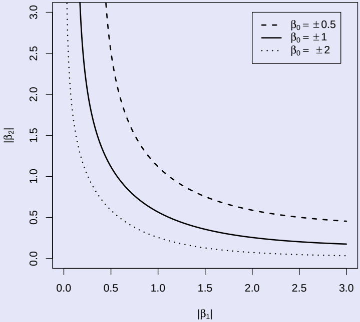
For the logit link, the saturation condition is true if and only if

$$\beta_0 \neq 0, \quad |\beta_1| > \frac{1}{2}\log\left(\frac{e^{2|\beta_0|}+1}{e^{2|\beta_0|}-1}\right), \quad \textit{and}$$

$$|\beta_2| \geq \log\left(\frac{2e^{|\beta_0|+|\beta_1|}+\sqrt{\left(e^{4|\beta_0|}-1\right)\left(e^{4|\beta_1|}-1\right)}}{\left(e^{2|\beta_0|}-1\right)\left(e^{2|\beta_1|}-1\right)-2}\right).$$

Saturation Condition for the Logit Link

For fixed β_0, a pair (β_1, β_2) satisfies the saturation condition if and only if the corresponding point is above the curve labelled by β_0.



Lower boundary of the region satisfying the saturation condition

UIC
World-Class Education. World-Class City.

Characterization of locally D-optimal designs

- We consider locally D-optimal designs for the full factorial 2^k experiment. The goal is to find an optimal $\mathbf{p} = (p_1, p_2, \ldots, p_{2^k})'$ which maximizes $f(\mathbf{p}) = |X'WX|$ for specified values of $w_i \geq 0, i = 1, \ldots, 2^k$.
- It is easy to see that there always exists a D-optimal allocation $\mathbf{p}$ since the set of all feasible allocations is bounded and closed.
- On the other hand, the uniqueness of D-optimal designs is usually not guaranteed, it depends on the design matrix X and $\mathbf{w} = (w_1, \ldots, w_{2^k})'$

UIC
World-Class Education. World-Class City.

Characterization of locally D-optimal designs

- Suppose the parameter of interest is $\beta = (\beta_0, \beta_1, \ldots, \beta_d)'$, where $d \geq k$. The following theorem expresses our objective function as an order-(d+1) homogeneous polynomial of $p_1, \ldots, p_{2^k}$, which is useful in deriving optimal p_i's.

Theorem

Let $X[i_1, i_2, \ldots, i_{d+1}]$ be the $(d+1) \times (d+1)$ sub-matrix consisting of the i_1th, i_2th, ..., i_{d+1}th rows of the design matrix X. Then

$$f(\mathbf{p}) = |X'WX| = \sum_{1 \leq i_1 < i_2 < \cdots < i_{d+1} \leq 2^k} |X[i_1, i_2, \ldots, i_{d+1}]|^2 \cdot p_{i_1} w_{i_1} p_{i_2} w_{i_2} \cdots p_{i_{d+1}} w_{i_{d+1}}.$$

2^3 main-effects plan

- For 2^3 experiment with main-effect plan

$$X = \begin{pmatrix} 1 & +1 & +1 & +1 \\ 1 & +1 & +1 & -1 \\ 1 & +1 & -1 & +1 \\ 1 & +1 & -1 & -1 \\ 1 & -1 & +1 & +1 \\ 1 & -1 & +1 & -1 \\ 1 & -1 & -1 & +1 \\ 1 & -1 & -1 & -1 \end{pmatrix}$$

Characterization of locally D-optimal designs

- From the previous theorem it is immediate that at least $(d+1)$ p_i's, as well as the corresponding w_i's, have to be positive for the determinant $f(\mathbf{p})$ to be nonzero. This implies if $\mathbf{p}$ is D-optimal, then $p_i < 1$ for each i.
- Our next theorem gives a sharper bound, $p_i \leq \frac{1}{d+1}$ for each $i = 1, \ldots, 2^k$, for the optimal allocation.
- To check this we define for each $i = 1, \ldots, 2^k$,

$$f_i(x) = f\left(\frac{1-x}{1-p_i}p_1, \ldots, \frac{1-x}{1-p_i}p_{i-1}, x, \frac{1-x}{1-p_i}p_{i+1}, \ldots, \frac{1-x}{1-p_i}p_{2^k}\right), \quad 0 \leq x \leq 1. \tag{3}$$

Note that $f_i(x)$ is well defined for all $\mathbf{p}$ of interest.

Characterization of locally D-optimal designs

The next theorem characterizes all locally D-optimal allocations.

Theorem

Suppose $f(\mathbf{p}) > 0$. Then $\mathbf{p}$ is D-optimal if and only if for each $i = 1, \ldots, 2^k$, one of the two conditions below is satisfied:

(i) $p_i = 0$ *and* $f_i\left(\frac{1}{2}\right) \le \frac{d+2}{2^{d+1}} f(\mathbf{p})$;

(ii) $0 < p_i \le \frac{1}{d+1}$ *and* $f_i(0) = \frac{1-p_i(d+1)}{(1-p_i)^{d+1}} f(\mathbf{p})$.

Saturated Designs Revisited

Theorem

Assume $w_i > 0$, $i = 1, \ldots, 2^k$. *Let* $\mathbf{I} = \{i_1, \ldots, i_{d+1}\} \subset \{1, \ldots, 2^k\}$ *be an index set satisfying* $|X[i_1, \ldots, i_{d+1}]| \neq 0$. *Then the saturated design satisfying* $p_{i_1} = p_{i_2} = \cdots = p_{i_{d+1}} = \frac{1}{d+1}$ *is D-optimal if and only if for each* $i \notin \mathbf{I}$,

$$\sum_{j \in \mathbf{I}} \frac{|X[\{i\} \cup \mathbf{I} \setminus \{j\}]|^2}{w_j} \leq \frac{|X[i_1, i_2, \ldots, i_{d+1}]|^2}{w_i}.$$

- Under the 2^2 main-effects model, $p_1 = p_2 = p_3 = 1/3$ is D-optimal if and only if $v_1 + v_2 + v_3 \leq v_4$, where $v_i = 1/w_i$, $i = 1, 2, 3, 4$.
- For the 2^3 main-effects model, $p_1 = p_4 = p_6 = p_7 = 1/4$ is D-optimal if and only if $v_1 + v_4 + v_6 + v_7 \leq 4\min\{v_2, v_3, v_5, v_8\}$, and $p_2 = p_3 = p_5 = p_8 = 1/4$ is D-optimal if and only if $v_2 + v_3 + v_5 + v_8 \leq 4\min\{v_1, v_4, v_6, v_7\}$.

EW D-optimal designs

- In order to use the D-optimal designs, we need to specify the β_i's, which gives us the w_i's.
- In situations where information from previous experiments and/or the nature of the experiments can only give a range or an approximate distribution of the w_i's, a natural option may be to use the expectation of the w_i's as input in the D-criterion. This we call the *EW D-optimality*.
- An alternative to local optimality is Bayes optimality. We will show that EW-optimal designs are often approximately as efficient as Bayes-optimal ones, while realizing considerable savings in computational time.
- In general, EW designs are robust.

EW D-optimal designs

The following theorem characterizes the EW D-optimal designs in terms of the regression coefficients.

Theorem

For any given link function, if the regression coefficients $\beta_0, \beta_1, \ldots, \beta_d$ *are independent with finite expectation, and* $\beta_1, \ldots, \beta_d$ *all have a symmetric distribution about* 0 *(not necessarily the same distribution), then the uniform design is an EW D-optimal design.*

A Bayes D-optimal design maximizes $E(\log|X'WX|)$ where the expectation is taken over the prior distribution of β_i's. Note that, by Jenson's inequality,

$$E\left(\log|X'WX|\right) \leq \log|X'E(W)X|$$

since $\log|X'WX|$ is concave in $\mathbf{w}$. Thus an EW D-optimal design maximizes an upper bound to the Bayesian D-optimality criterion.

EW D-optimal designs

Example

2^2 **main-effects model:** Suppose $\beta_0, \beta_1, \beta_2$ are independent, $\beta_0 \sim U(-1,1)$, and $\beta_1, \beta_2 \sim U[0,1)$.

- Under the logit link, the EW D-optimal design is $\mathbf{p}_e = (0.239, 0.261, 0.261, 0.239)'$.
- The Bayes optimal design, which maximizes $\phi(\mathbf{p}) = E \log |X'WX|$ is $\mathbf{p}_o = (0.235, 0.265, 0.265, 0.235)'$. The relative efficiency of $\mathbf{p}_e$ with respect to $\mathbf{p}_o$ is

$$\exp\left\{\frac{\phi(\mathbf{p}_e) - \phi(\mathbf{p}_o)}{k+1}\right\} \times 100\% = \exp\left\{\frac{-4.80665 - (-4.80642)}{2+1}\right\} \times 100\% = 99.99\%.$$

- The time cost for EW is 0.11 sec, while it is 5.45 secs for maximizing $\phi(\mathbf{p})$.

EW D-optimal designs

Example

2^3 **main-effects model:** Suppose $\beta_0, \beta_1, \beta_2, \beta_3$ are independent, and the experimenter has the following prior information for the parameters: $\beta_0 \sim U(-3, 3)$, and $\beta_1, \beta_2, \beta_3 \sim U[0, 3)$.

- For the logit link the uniform design is not EW D-optimal.
- In this case, EW solution may not be unique, so we run the program 1000 times starting with random initial allocations and check its relative efficiency with respect to $\mathbf{p}_o$ which maximizes $\phi(\mathbf{p})$.
- The minimum relative efficiency is 99.98%.
- On the other hand, the relative efficiency of the uniform design with respect to $\mathbf{p}_o$ is 94.39%.
- It takes about 2.39 seconds to find an EW solution while it takes 121.73 seconds to find $\mathbf{p}_o$. The difference in computational time is even more prominent for 2^4 case (24 seconds versus 3147 seconds).

UIC
World-Class Education. World-Class City.

Algorithm to search for locally D-optimal allocation

- We develop efficient algorithms to search for locally D-optimal solution with given w_i's.
- The same algorithms can be used for finding EW D-optimal designs.
- Our algorithms are much faster than commonly used optimization algorithms including Nelder-Mead, quasi-Newton, conjugate-gradient, and simulated annealing algorithms (for a comprehensive reference, see Nocedal and Wright (1999)).
  - For example, for a 2^3 main-effects case, it takes about 0.011 second using our lift-one algorithm to find an optimal design. This is more than 1000 times faster than the other competing algorithms mentioned above.
- Our algorithm finds the optimal solution with more accuracy as well. The improvement is even more prominent for larger values of k.

UIC
World-Class Education. World-Class City.

Lift-one algorithm for maximizing $f(\mathbf{p}) = |X'WX|$

The basic idea is that, for randomly chosen $i \in \{1, \ldots, 2^k\}$, we update p_i to p_i^* and all the other p_j's to $p_j^* = p_j \cdot \frac{1-p_i^*}{1-p_i}$.

1° Start with arbitrary $\mathbf{p}_0 = (p_1, \ldots, p_{2^k})'$ satisfying $0 < p_i < 1$, $i = 1, \ldots, 2^k$ and compute $f(\mathbf{p}_0)$.

2° Set up a random order of i going through $\{1, 2, \ldots, 2^k\}$.

3° For each i, calculate $f_i(x)$. In this step, either $f_i(0)$ or $f_i\left(\frac{1}{2}\right)$ needs to be calculated according to equation (3).

4° Define $\mathbf{p}_*^{(i)} = \left(\frac{1-x_*}{1-p_i}p_1, \ldots, \frac{1-x_*}{1-p_i}p_{i-1}, x_*, \frac{1-x_*}{1-p_i}p_{i+1}, \ldots, \frac{1-x_*}{1-p_i}p_{2^k}\right)'$, where x_* maximizes $f_i(x)$ with $0 \leq x \leq 1$. Note that $f(\mathbf{p}_*^{(i)}) = f_i(x_*) \geq f(\mathbf{p}_0) > 0$.

5° Replace $\mathbf{p}_0$ with $\mathbf{p}_*^{(i)}$, $f(\mathbf{p}_0)$ with $f\left(\mathbf{p}_*^{(i)}\right)$.

6° Repeat $2° \sim 5°$ until convergence, that is, $f(\mathbf{p}_0) = f(\mathbf{p}_*^{(i)})$ for each i.

Lift-one algorithm for maximizing $f(\mathbf{p}) = |X'WX|$

- While in all examples that we studied, the lift-one algorithm converges very fast, we do not have a proof of convergence. There is a modified version, which is only slightly slower, that can be shown to converge.
- For the $10m$th iteration we repeat steps $2^\circ \sim 5^\circ$, $m = 1, 2, \ldots$, if $\mathbf{p}_*^{(i)}$ is a better allocation found by the lift-one algorithm than the allocation $\mathbf{p}_0$, instead of updating $\mathbf{p}_0$ to $\mathbf{p}_*^{(i)}$ immediately, we obtain $\mathbf{p}_*^{(i)}$ for each i, and replace $\mathbf{p}_0$ with the best one among $\left\{\mathbf{p}_*^{(i)}, i = 1, \ldots, 2^k\right\}$. For iterations other than the $10m$th, we follow the original lift-one algorithm update.

Theorem

When the lift-one algorithm or the modified lift-one algorithm converges, the converged allocation $\mathbf{p}$ maximizes $|X'WX|$ on the set of feasible allocations. Furthermore, the modified lift-one algorithm is guaranteed to converge.

Algorithm for maximizing $|X'WX|$ with integer solutions

- To maximize $|X'WX|$, an alternative algorithm, called *exchange algorithm*, is to adjust p_i and p_j for randomly chosen index pair (i,j). The optimal adjusted (p_i^*, p_j^*) can be obtained easily by maximizing a quadratic function.
- Unlike the lift-one algorithm, the exchange algorithm can be applied to search for integer-valued optimal allocation $\mathbf{n} = (n_1, \ldots, n_{2^k})'$, where $\sum_i n_i = n$.
- As expected, the integer-valued optimal allocation $(n_1, \ldots, n_{2^k})'$ is consistent with the proportion-valued allocation $(p_1, \ldots, p_{2^k})'$ for large n.
- For small n, the result may be used for fractional design problem in the next Section.
- It should be noted that this exchange algorithm with slight modification is guaranteed to converge to the optimal allocation when searching for proportions but not for integer-valued solutions, especially when n is quite small compared to 2^k.

Outline

Fractional Factorial Designs

- Unless the number of factors k is very small, the total number of experimental conditions 2^k is large, and it could be impossible to take observations at all 2^k level combinations. So when k is large, fractions are an attractive option.
- For linear models, the accepted practice is to use regular fractions due to the many desirable properties like minimum aberration and optimality.
- In our setup the regular fractions are not always optimal.
- First we examine the situations when they are optimal.
- We use the 2^3 design with main-effects model for illustration.

UIC
World-Class Education. World-Class City.

2^3 Design with Main-effects Model

Theorem

For the 2^3 main-effects model, suppose $\beta_1 = 0$. This implies $w_1 = w_5$, $w_2 = w_6$, $w_3 = w_7$, and $w_4 = w_8$. The regular fractions $\{1, 4, 6, 7\}$ and $\{2, 3, 5, 8\}$ are D-optimal half-fractions if and only if

$$4\ \min\{w_1, w_2, w_3, w_4\} \geq \max\{w_1, w_2, w_3, w_4\}.$$

Further suppose $\beta_2 = 0$. Then $w_1 = w_3 = w_5 = w_7$ and $w_2 = w_4 = w_6 = w_8$. The two regular half-fractions mentioned above are D-optimal half-fractions if and only if $4 \min\{w_1, w_2\} \geq \max\{w_1, w_2\}$.

UIC
World-Class Education. World-Class City.

2^3 Main-effects Model

Example

- Under logit link, consider the 2^3 main-effects model with $\beta_1 = \beta_2 = 0$. The regular half-fractions $\{1,4,6,7\}$ and $\{2,3,5,8\}$ are D-optimal half-fractions if and only if one of the following happens:

$$\begin{aligned} &(i) \quad |\beta_3| \le \log 2 \\ &(ii) \quad |\beta_3| > \log 2 \quad \text{and} \quad |\beta_0| \le \log\left(\frac{2e^{|\beta_3|}-1}{e^{|\beta_3|}-2}\right). \end{aligned} \tag{4}$$

- When the regular half-fractions are not optimal, the goal is to find $\{i_1, i_2, i_3, i_4\}$ that maximizes

$$|X[i_1,i_2,i_3,i_4]|^2 w_{i_1} w_{i_2} w_{i_3} w_{i_4}.$$

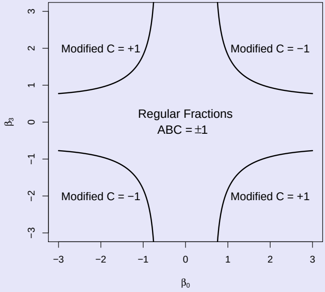
UIC
World-Class Education.
World-Class City.

2^3 Main-effects Model

Example

- Recall that in this case there are only two distinct w_i's. If $\beta_0\beta_3 > 0$, w_i's corresponding to $\{2,4,6,8\}$ are larger than others, so this fraction given by $C = -1$ will maximize $w_{i_1}w_{i_2}w_{i_3}w_{i_4}$ but this leads to a singular design matrix.
- It is not surprising that the D-optimal half-fractions are "close" to the design $\{2,4,6,8\}$, and are in fact given by the 16 designs each consisting of three elements from $\{2,4,6,8\}$ and one from $\{1,3,5,7\}$. For $\beta_0\beta_3 < 0$, D-optimal half-fractions are similarly obtained from the fraction $C = +1$.
- The message is that if β_1,β_2 and β_3 are small then the regular fractions are preferred. However, when at least one $|\beta_i|$ is large, the information is not uniformly distributed over the design points and the regular fractions may not be optimal.

Partitioning of the parameter space



$$\beta_1 = \beta_2 = 0$$

The parameters in the middle region would make the regular fractions D-optimal whereas the top-right and bottom-left regions correspond to the case $\beta_0\beta_3 > 0$. Similarly the other two regions correspond to the case $\beta_0\beta_3 < 0$.

UIC
World-Class Education.
World-Class City.

Distribution of D-optimal half-fractions under 2^3 main-effects model

In general, when all the β_i's are nonzero, the regular fractions given by the rows $\{1,4,6,7\}$ or $\{2,3,5,8\}$ are not necessarily the optimal ones. We illustrate this via simulation.

Rows	Percentages (Efficiencies)					
Simulation Setup	$\beta_0 \sim$	$U(-10,10)$				$N(0,5)$
	$\beta_1 \sim$	$U(-.3,.3)$	$U(-3,3)$	$U(-3,0)$	$U(0,1)$	$N(1,1)$
	$\beta_2 \sim$	$U(-.3,.3)$	$U(0,3)$	$U(0,3)$	$U(0,3)$	$N(2,1)$
	$\beta_3 \sim$	$U(-.3,.3)$	$U(1,5)$	$U(-2,2)$	$U(0,5)$	$N(3,1)$
1467		42.65 (99.5)	0.07 (0.22)	0.86 (2.18)	0.95 (2.61)	0.04 (2.89)
2358		42.02 (99.5)	0.04 (0.23)	0.68 (2.12)	1.04 (2.56)	0.08 (2.86)
1235			16.78		35.62	21.50
1347				19.98		
1567			17.45	19.21		
2348			17.54	19.11		
2568				20.01		
4678			16.12		35.41	21.65

Strategies for Finding Optimal Experimental Settings for General Cases

- The number of experimental settings (m, say) is fixed.
- One strategy would be to choose the m largest w_i's for identifying the rows as the w_i's denote the information at that point.
- Another one could be to use our algorithms for finding an optimal allocation for the full factorial designs first. Then choose the m largest p_i's and scale them appropriately.
  - One has to be careful in order to avoid designs which would not allow the estimation of the model parameters. In this case, the exchange algorithm may be used to choose the fraction with given m experimental units.
- Our simulations show that both of these methods perform satisfactorily with the second method giving designs which are generally more than 95% efficient for four factors with main-effects model.

Outline

Robustness for 2^2 Designs with Main Effects Model

- Since locally optimal designs depend on the assumed values of the parameters, it is important to study the robustness of the designs to these values.
- For experiments where there is no basis for making an informed choice of the assumed values, the natural design choice is the uniform design.

Robustness for misspecification of w

- Let us denote the D-criterion value as $\psi(\mathbf{p},\mathbf{w}) = |X'WX|$ for given $\mathbf{w} = (w_1,\ldots,w_{2^k})'$ and $\mathbf{p} = (p_1,\ldots,p_{2^k})'$. Suppose $\mathbf{p}_w$ is a D-optimal allocation with respect to $\mathbf{w}$. Then the relative loss of efficiency of $\mathbf{p}$ with respect to $\mathbf{w}$ can be defined as

$$R(\mathbf{p},\mathbf{w}) \quad = \quad 1-\left(\frac{\psi(\mathbf{p},\mathbf{w})}{\psi(\mathbf{p}_w,\mathbf{w})}\right)^{\frac{1}{d+1}}.$$

Let us define the maximum relative loss of efficiency of a given design $\mathbf{p}$ with respect to a specified region $\mathcal{W}$ of $\mathbf{w}$ by

$$R_{\max}(\mathbf{p}) = \max_{\mathbf{w}\in\mathcal{W}} R(\mathbf{p},\mathbf{w}).$$

This maximum will correspond to the worst case scenario. This tells us how bad the design can perform if we do not choose the $\mathbf{w}$ correctly.

Robustness of 2^2 Uniform Design with Main Effects Model

Theorem

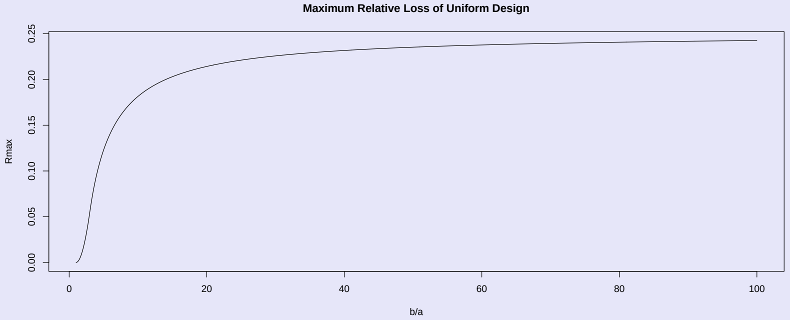
Suppose $0 < a \le v_{ti} \le b$, $i = 1,2,3,4$. *Let* $\theta = b/a \ge 1$. *Then*

$$R_{\max}^{(u)} = \begin{cases} 1 - \frac{3}{4}\left(1 + \frac{3}{\theta}\right)^{1/3}, & \text{if } \theta \ge 3 \\ 1 - \frac{1}{8}\left(2(\theta+3)(9-\theta)^2\right)^{1/3}, & \text{if } \theta_* \le \theta < 3 \\ 1 - \frac{3}{2}\left(\frac{(\theta+1)(\theta-1)^2}{(2\theta-1-\rho)(\theta-2+\rho)(\theta+1+\rho)}\right)^{1/3}, & \text{if } 1 < \theta < \theta_* \\ 0 & \text{if } \theta = 1 \end{cases}$$

where $\rho = \sqrt{\theta^2 - \theta + 1}$, *and* $\theta_* \approx 1.32$ *is the 3rd root of the equation*

$$3456 - 5184\theta + 3561\theta^2 + 596\theta^3 - 1506\theta^4 + 100\theta^5 + \theta^6 = 0.$$

Maximum Relative Loss of Uniform Design



Plot of $R_{\max}^{(u)}$ versus θ

Robustness of EW D-optimal Designs

- For most practical applications, the experimenter will have some idea about the direction of effects of factors, which determines the signs of the regression coefficients. In addition, suppose the experimenter has some idea about the distribution of the parameters within a certain range. In these cases, the EW D-optimal designs are very robust.

Robustness of Uniform Designs Restricted to Regular Fractions

- Simulations based on half fractions of the 2^4 main-effects model show
  - Uniform designs on regular fractions are among the most robust ones if the regression parameters are supported symmetrically around zero.
  - But they may not perform well if some of the β_i's are positive (or negative).

Outline

Recommendation - Windshield Experiment Revisited

- Consider the Windshield molding experiment.
- By analyzing the data presented in Hamada and Nelder (1997), we get an estimate of the unknown parameter as $\hat{\beta} = (1.77, -1.57, 0.13, -0.80, -0.14)'$ under logit link.
- In future, if someone wants to conduct a follow-up experiment on half-fractions, then it is sensible to use the knowledge obtained by analyzing the data.
- In this case, the efficiency of the original 2_{III}^{4-1} design is 78% of the locally D-optimal design if $\hat{\beta}$ were the true value of the unknown parameter β.

Recommendation - Windshield Experiment Revisited

- It is reasonable to assume the true value of the unknown parameter is not far from $\hat{\beta}$ and one might want to use that information to improve the efficiency of the follow-up experiment.
- If the efficiency of the regular 2_{III}^{4-1} design were 90% or higher, then we would have recommended using it for the follow-up experiment as well, because of their robustness.
- However, in this case, our recommendation would be to consider a few initial guesses around $\hat{\beta}$ and choose a design which has higher efficiency and is also preferred by the experimenter.

Recommendation - Windshield Experiment Revisited

- For example, it might be reasonable to consider an initial guess of $\beta = (2, -1.5, 0.1, -1, -0.1)'$. This will lead to the locally D-optimal half-fractional design $\mathbf{p}_a$ given below.

Locally D-optimal design for
$\beta = (2, -1.5, 0.1, -1, -0.1)'$

Row	A	B	C	D	$\mathbf{p}_a$
1	$+1$	$+1$	$+1$	$+1$	0.178
2	$+1$	$+1$	$+1$	-1	0.059
3	$+1$	$+1$	-1	-1	0.147
4	$+1$	-1	$+1$	$+1$	0.044
5	$+1$	-1	$+1$	-1	0.178
6	$+1$	-1	-1	$+1$	0.163
7	-1	$+1$	$+1$	-1	0.074
8	-1	-1	$+1$	$+1$	0.158

Recommendation - Windshield Experiment Revisited

- If we assume that the regression coefficients are uniform over the ranges: $(1,3)$ for β_0, $(-3,-1)$ for β_1, $(-0.5,0.5)$ for β_2,β_4, and $(-1,0)$ for β_3. For this choice of range for the parameter values, the EW D-optimal half-fractional design $\mathbf{p}_e$ with uniformity assumption on the ranges is given below.

EW D-optimal design

Row	A	B	C	D	$\mathbf{p}_e$
1	$+1$	$+1$	$+1$	$+1$	0.011
2	$+1$	$+1$	$+1$	-1	0.184
3	$+1$	$+1$	-1	$+1$	0.195
4	$+1$	-1	$+1$	$+1$	0.184
5	$+1$	-1	$+1$	-1	0.011
6	$+1$	-1	-1	-1	0.195
7	-1	$+1$	$+1$	-1	0.110
8	-1	-1	$+1$	$+1$	0.111

- Note that if we regard $\hat{\beta}$ as the true value of the parameter, the relative efficiency of $\mathbf{p}_a$ increases to 99%, whereas that of $\mathbf{p}_e$ increases to 98%.

Outline

Conclusions

- We consider the problem of obtaining locally D-optimal designs for factorial experiments with binary response and k qualitative factors at two levels each.
- We obtain a characterization for a design to be locally D-optimal.
- Based on this characterization, we develop efficient numerical techniques to search for locally D-optimal designs.
- We suggest the use of EW D-optimal designs. These are much easier to compute and still highly efficient compared with Bayesian D-optimal designs.
- We investigate the properties of fractional factorial designs.
- We also study the robustness of the D-optimal designs.

UIC
World-Class Education.
World-Class City.

THANK YOU!!!

References:

1. Yang, J., Mandal, A. and Majumdar, D. (2011), Optimal designs for two-level factorial experiments with binary response. *Statistica Sinica*, to appear.
2. Yang, J., Mandal, A. and Majumdar, D. (2011), Optimal designs for 2^k factorial experiments with binary response. `http://www.stat.uga.edu/~amandal/MYD2.pdf`.