D-optimal Designs for Factorial Experiments under Generalized Linear Models

Jie Yang

Department of Mathematics, Statistics, and Computer Science
University of Illinois at Chicago

Joint research with Abhyuday Mandal and Dibyen Majumdar

October 20, 2012

A motivating example: Hamada and Nelder (1997)

- Two-level factors: (A) poly-film thickness, (B) oil mixture ratio, (C) material of gloves, and (D) condition of metal blanks.
- Response: the windshield molding was good or not.

Row	A	B	C	D	Replicates	good molding
1	+	+	+	+	1000	338
2	+	+	−	−	1000	826
3	+	−	+	−	1000	350
4	+	−	−	+	1000	647
5	−	+	+	−	1000	917
6	−	+	−	+	1000	977
7	−	−	+	+	1000	953
8	−	−	−	−	1000	972

- **Question:** Can we do something better?

Preliminary setup

- Consider an experiment with m fixed and distinct design points:

$$X = \begin{pmatrix} \mathbf{x}_1' \\ \mathbf{x}_2' \\ \vdots \\ \mathbf{x}_m' \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1d} \\ x_{21} & x_{22} & \cdots & x_{2d} \\ \cdots & \cdots & \cdots & \cdots \\ x_{m1} & x_{m2} & \cdots & x_{md} \end{pmatrix}$$

  For example, a 2^k factorial experiment with main-effects model implies $m = 2^k$, $d = k + 1$, $x_{i1} = 1$, and $x_{i2}, \ldots, x_{id} \in \{-1, 1\}$.

- *Exact design problem:* Suppose n is given.
  Consider "optimal" n_i's such that $n_i \geq 0$ and $\sum_{i=1}^m n_i = n$.

- *Approximate design problem:* Let $p_i = n_i/n$.
  Consider "optimal" p_i's such that $p_i \geq 0$ and $\sum_{i=1}^m p_i = 1$.

Generalized linear model: single parameter

Consider independent univariate responses $Y_1, \ldots, Y_n$:

$$Y_i \sim f(y; \theta_i) = \exp\{yb(\theta_i) + c(\theta_i) + d(y)\}$$

For example,

$$\begin{array}{ll}
\exp\left\{y \log \frac{\theta}{1-\theta} + \log(1-\theta)\right\}, & \text{Bernoulli}(\theta) \\
\exp\{y \log \theta - \theta - \log y!\}, & \text{Poisson}(\theta) \\
\exp\left\{y \frac{-1}{\theta} - k \log \theta + \log \frac{y^{k-1}}{\Gamma(k)}\right\}, & \text{Gamma}(k, \theta),\ \text{fixed } k > 0 \\
\exp\left\{y \frac{\theta}{\sigma^2} - \frac{\theta^2}{2\sigma^2} - \frac{y^2}{2\sigma^2} - \frac{1}{2}\log(2\pi\sigma^2)\right\}, & N(\theta, \sigma^2),\ \text{fixed } \sigma^2 > 0
\end{array}$$

Generalized linear model (McCullagh and Nelder 1989, Dobson 2008): $\exists$ link function g and parameters of interest $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_d)'$, such that

$$E(Y_i) = \mu_i \text{ and } \eta_i = g(\mu_i) = \mathbf{x}_i' \boldsymbol{\beta}.$$

Information matrix and D-optimal design

Recall that there are m distinct predictor combinations $\mathbf{x}_1, \ldots, \mathbf{x}_m$ with numbers of replicates $n_1, \ldots, n_m$, respectively.

The maximum likelihood estimator of β has an asymptotic covariance matrix that is the inverse of the *information matrix*

$$\mathbf{I} = nX'WX$$

where $X = (\mathbf{x}_1, \ldots, \mathbf{x}_m)'$ is an $m \times d$ matrix, and $W = \mathrm{diag}(p_1 w_1, \ldots, p_m w_m)$ with $w_i = \frac{1}{\mathrm{var}(Y_i)} \left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2$.

A *D-optimal design* is the $\mathbf{p} = (p_1, \ldots, p_m)'$ which maximizes

$$|X'WX|$$

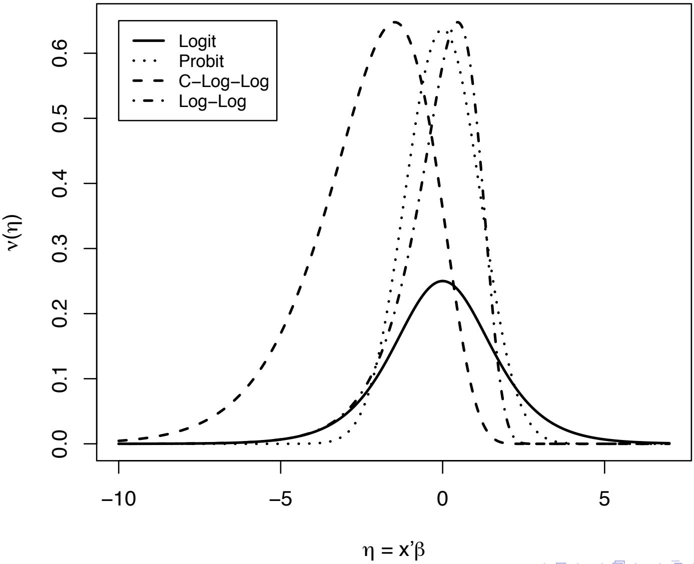
with given $w_1, \ldots, w_m$.

w as a function of β

Suppose the link function g is one-to-one and differentiable.
Suppose further μ_i itself determines $\mathrm{var}(Y_i)$.
Then $w_i = \nu(\eta_i) = \nu\left(\mathbf{x}_i{}'\beta\right)$ for some function ν.

- *Binary response, logit link:* $\nu(\eta) = \frac{1}{2+e^{\eta}+e^{-\eta}}$.
- *Poisson count, log link:* $w = \nu(\eta) = \exp\{\eta\}$.
- *Gamma response, reciprocal link:* $w = \nu(\eta) = k/\eta^2$.
- *Normal response, identity link:* $w = \nu(\eta) \equiv 1/\sigma^2$.

$w_i = \nu(\eta_i) = \nu(\mathbf{x}_i'\beta)$ for binary response



Example: 2^2 experiment with main-effects model

$$X = \begin{pmatrix} 1 & -1 & -1 \\ 1 & -1 & +1 \\ 1 & +1 & -1 \\ 1 & +1 & +1 \end{pmatrix}, W = \begin{pmatrix} w_1 p_1 & 0 & 0 & 0 \\ 0 & w_2 p_2 & 0 & 0 \\ 0 & 0 & w_3 p_3 & 0 \\ 0 & 0 & 0 & w_4 p_4 \end{pmatrix}$$

The optimization problem maximizing

$$|X'WX| = 16 w_1 w_2 w_3 w_4 L(\mathbf{p}),$$

where $v_i = 1/w_i$ and

$$L(\mathbf{p}) = v_4 p_1 p_2 p_3 + v_3 p_1 p_2 p_4 + v_2 p_1 p_3 p_4 + v_1 p_2 p_3 p_4$$

General case: order-d polynomial

For general case, X is an $m \times d$ matrix with distinct rows, $W = \mathrm{diag}(p_1 w_1, \ldots, p_m w_m)$.

Based on González-Dávila, Dorta-Guerra and Ginebra (2007) and Yang, Mandal and Majumdar (2012b), we have

Lemma

Let $X[i_1, i_2, \ldots, i_d]$ be the $d \times d$ sub-matrix consisting of the i_1th, ..., i_dth rows of the design matrix X. Then

$$f(\mathbf{p}) = |X'WX| = \sum_{1 \le i_1 < \cdots < i_d \le m} |X[i_1, \ldots, i_d]|^2 \cdot p_{i_1} w_{i_1} \cdots p_{i_d} w_{i_d}.$$

Characterization of locally D-optimal designs

For each $i = 1, \ldots, m$, $0 \le x < 1$, we define

$$\begin{aligned} f_i(x) &= f\left(\frac{1-x}{1-p_i}p_1, \ldots, \frac{1-x}{1-p_i}p_{i-1}, x, \frac{1-x}{1-p_i}p_{i+1}, \ldots, \frac{1-x}{1-p_i}p_m\right) \\ &= ax(1-x)^{d-1} + b(1-x)^d \end{aligned}$$

If $p_i > 0$, $b = f_i(0)$, $a = \frac{f(\mathbf{p}) - b(1-p_i)^d}{p_i(1-p_i)^{d-1}}$; otherwise, $b = f(\mathbf{p})$, $a = f_i\left(\frac{1}{2}\right) \cdot 2^d - b$.

Theorem

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

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

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

Saturated designs

Theorem

Let $\mathbf{I} = \{i_1, \ldots, i_d\} \subset \{1, \ldots, m\}$ *satisfying* $|X[i_1, \ldots, i_d]| \neq 0$. *Then* $p_{i_1} = p_{i_2} = \cdots = p_{i_d} = \frac{1}{d}$ *is D-optimal if and only if for each* $i \notin \mathbf{I}$,

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

- 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 $v_i = 1/w_i$.
- 2^3 *main-effects model:* $p_1 = p_4 = p_6 = p_7 = 1/4$ is D-optimal if and only if $v_1 + v_4 + v_6 + v_7 \leq 4\min\{v_2, v_3, v_5, v_8\}$.

Saturated designs: more example

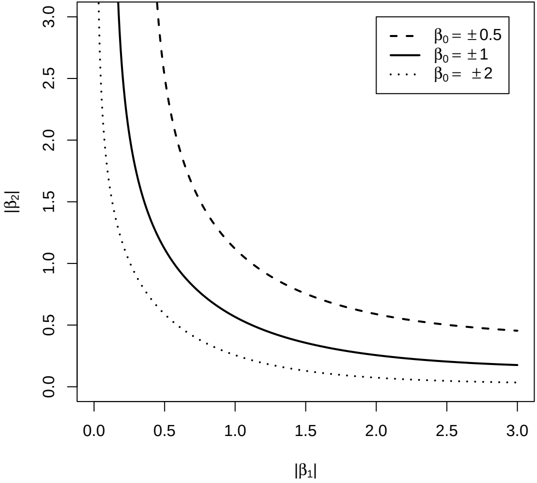
2×3 *factorial design:* Suppose the design matrix

$$X = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & -2 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & -1 & 0 & -2 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

- $p_1 = p_2 = p_3 = p_4 = 1/4$ is D-optimal if and only if $v_1 + v_2 + v_4 \le v_5$ and $v_1 + v_3 + v_4 \le v_6$.
- $p_2 = p_3 = p_4 = p_5 = 1/4$ is D-optimal if and only if $v_2 + v_4 + v_5 \le v_1$ and $v_2 + v_3 + v_5 \le v_6$.
- $p_3 = p_4 = p_5 = p_6 = 1/4$ is D-optimal if and only if $v_3 + v_4 + v_6 \le v_1$ and $v_3 + v_5 + v_6 \le v_2$.

Saturation condition in terms of β: logit link, 2^2 design

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.



Lift-one algorithm (Yang, Mandal and Majumdar, 2012b)

Given X and $w_1, \ldots, w_m$, search for $\mathbf{p} = (p_1, \ldots, p_m)'$ maximizing $f(\mathbf{p}) = |X'WX|$:

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

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

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

4° Define
$$\mathbf{p}_*^{(i)} = \left(\frac{1-z_*}{1-p_i}p_1, \ldots, \frac{1-z_*}{1-p_i}p_{i-1}, z_*, \frac{1-z_*}{1-p_i}p_{i+1}, \ldots, \frac{1-z_*}{1-p_i}p_{2^k}\right)',$$
where z_* maximizes $f_i(z)$, $0 \le z \le 1$. Here $f(\mathbf{p}_*^{(i)}) = f_i(z_*)$.

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

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

Comments on lift-one algorithm

- The basic idea of the lift-one algorithm is that, for randomly chosen $i \in \{1, \ldots, m\}$, we update p_i to p_i^* and all the other p_j's to $p_j^* = p_j \cdot \frac{1-p_i^*}{1-p_i}$.
- The major advantage of the lift-one algorithm is that in order to determine an optimal p_i^*, we need to calculate $|X'WX|$ only once.
- This algorithm is motivated by the *coordinate descent algorithm* (Zangwill, 1969).
- It is also in spirit similar to the idea of one-point correction in the literature (Wynn, 1970; Fedorov, 1972; Müller, 2007), where design points are added/adjusted one by one.

Convergence of lift-one algorithm

Modified lift-one algorithm:

(1) For the $10m$th iteration and a fixed order of $i = 1, \ldots, m$ we repeat steps $3^\circ \sim 5^\circ$. 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 $\mathbf{p}_*^{(i)}$ only.

(2) 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.*

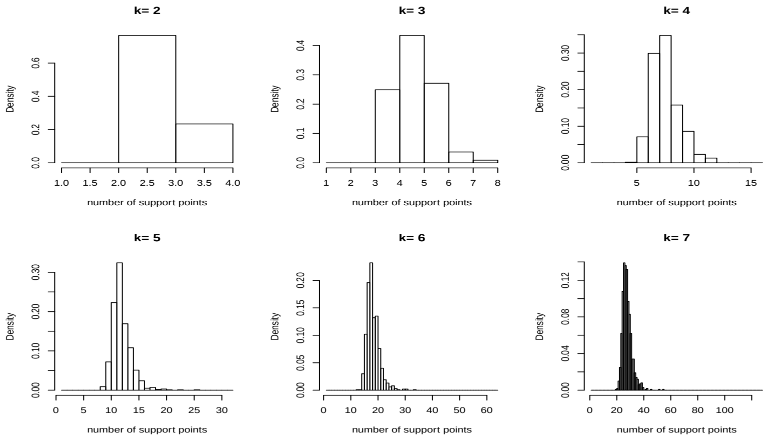
Performance of lift-one algorithm: comparison

Table: CPU time in seconds for 100 simulated β

	Algorithms				
	Nelder-Mead	quasi-Newton	conjugate gradient	simulated annealing	lift-one
2^2 Designs	0.77	0.41	4.68	46.53	0.17
2^3 Designs	46.75	63.18	925.5	1495	0.46
2^4 Designs	125.9	NA	NA	NA	1.45

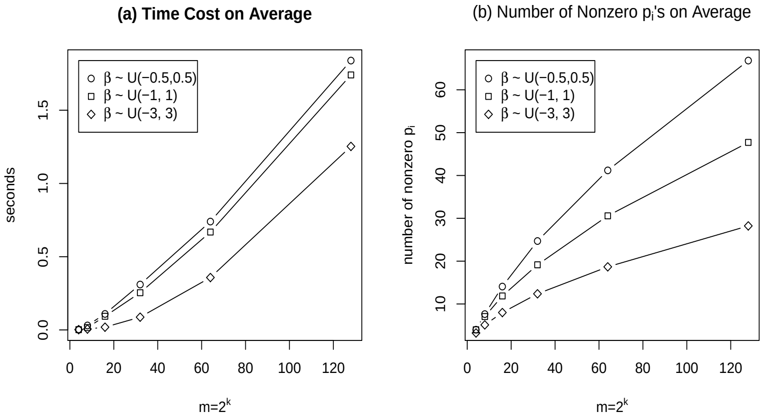
Performance of lift-one algorithm: number of nonzero p_i's

Under 2^k main-effects model, binary response, logit link, β_i's iid from Uniform(-3,3):



Performance of lift-one algorithm: time cost

Under 2^k main-effects model, binary response, logit link, we simulate β_i's iid from a uniform distribution and check the time cost in seconds by lift-one algorithm:



EW D-optimal designs

- For experiments under generalized linear models, we may need to specify the β_i's, which gives us the w_i's, to get D-optimal designs, known as *local D-optimality*.
- An *EW D-optimal* design is an optimal allocation of **p** that maximizes $|X'E(W)X|$. It is one of several alternatives suggested by Atkinson, Donev and Tobias (2007).
- EW D-optimal designs are often approximately as efficient as Bayesian D-optimal designs.
- EW D-optimal designs can be obtained easily using the lift-one algorithm.
- In general, EW D-optimal designs are robust.

EW D-optimal design and Bayesian D-optimal design

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 Jensen'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.

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

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

  for logit link, or 99.94% for probit link, 99.77% for log-log link, and 100.00% for complementary log-log link.
- The time cost for EW is 0.11 sec, while it is 5.45 secs for maximizing $\phi(\mathbf{p})$.
- It should also be noted that the relative efficiency of the uniform design $\mathbf{p}_u = (1/4, 1/4, 1/4, 1/4)'$ with respect to $\mathbf{p}_o$ is 99.88% for logit link, and is 89.6% for complementary log-log link.

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 is $\mathbf{p}_e = (0, 1/6, 1/6, 1/6, 1/6, 1/6, 1/6, 0)'$, while

  $$\mathbf{p}_o = (0.004, 0.165, 0.166, 0.165, 0.165, 0.166, 0.165, 0.004)'$$

  which maximizes $\phi(\mathbf{p})$.
- The relative efficiency of $\mathbf{p}_e$ with respect to $\mathbf{p}_o$ 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).

Robustness for misspecification of w

- Denote the D-criterion value as

$$\psi(\mathbf{p}, \mathbf{w}) = |X'WX|$$

for given $\mathbf{w} = (w_1, \ldots, w_m)'$ and $\mathbf{p} = (p_1, \ldots, p_m)'$.

- Define the relative loss of efficiency of $\mathbf{p}$ with respect to $\mathbf{w}$ as

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

where $\mathbf{p}_w$ is a D-optimal allocation with respect to $\mathbf{w}$.

- 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

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

Robustness of uniform design: 2^2 main-effects model

- Yang, Mandal and Majumdar (2012a) showed that under 2^2 experiment with main-effects model, if $w_i \in [a, b]$, $i = 1, 2, 3, 4$, $0 < a \leq b$, then the uniform design $\mathbf{p}_u = (1/4, 1/4, 1/4, 1/4)'$ is the most robust one in terms of the maximum of relative loss of efficiency.

- On the other hand, if the experimenter has some prior knowledge about the model parameters, for example, if one believes that $\beta_0 \sim \text{Uniform}(-1, 1)$, $\beta_1, \beta_2 \sim \text{Uniform}[0, 1)$ and $\beta_0, \beta_1, \beta_2$ are independent, then the theoretical $R_{\max}$ of uniform design is 0.134, while the theoretical $R_{\max}$ of design given by $\mathbf{p} = (0.19, 0.31, 0.31, 0.19)'$ is 0.116. That is, uniform design may not be the most robust one.

Robustness of uniform design: 2^4 main-effects model

Simulate $\beta_0, \ldots, \beta_4$ for 1000 times and calculate the corresponding $\mathbf{w}$'s. For each $\mathbf{w}_s$, we obtain a D-optimal allocation $\mathbf{p}_s$.

	Percentages											
	$\beta_0 \sim U(-3,3)$			$U(-1,1)$			$U(-3,0)$			$N(0,5)$		
	$\beta_1 \sim U(-1,1)$			$U(0,1)$			$U(1,3)$			$N(0,1)$		
	$\beta_2 \sim U(-1,1)$			$U(0,1)$			$U(1,3)$			$N(2,1)$		
	$\beta_3 \sim U(-1,1)$			$U(0,1)$			$U(-3,-1)$			$N(-.5,2)$		
	$\beta_4 \sim U(-1,1)$			$U(0,1)$			$U(-3,-1)$			$N(-.5,2)$		
	(I)	(II)	(III)	(I)	(II)	(III)	(I)	(II)	(III)	(I)	(II)	(III)
R_{99}	.35	.35	.35	.15	.11	.11	.50	.27	.30	.65	.86	.73
R_{95}	.30	.30	.30	.13	.09	.09	.50	.25	.26	.62	.79	.67
R_{90}	.27	.27	.27	.12	.08	.09	.49	.24	.23	.59	.74	.63

Note: (I) $= R_{100\alpha}(\mathbf{p}_u)$, uniform design;

(II) $= \min_{1\le s\le 1000} R_{100\alpha}(\mathbf{p}_s)$, best among 1000;

(III) $= R_{100\alpha}(\mathbf{p}_e)$, EW design.

Windshield experiment: revisited (Yang, Mandal and Majumdar, 2012b)

- Based on the data presented in Hamada and Nelder (1997), $\hat{\beta} = (1.77, -1.57, 0.13, -0.80, -0.14)'$ under logit link.
- The efficiency of the original 2_{III}^{4-1} design $\mathbf{p}_{HN}$ is 78% of the locally D-optimal design if $\hat{\beta}$ were the true value.
- 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 $\mathbf{p}_a$ with relative efficiency 99%.
- Another reasonable option is to consider a range, for example, $\beta_0 \sim \text{Unif}(1,3)$, $\beta_1 \sim \text{Unif}(-3,-1)$, $\beta_2, \beta_4 \sim \text{Unif}(-0.5, 0.5)$, and $\beta_3 \sim \text{Unif}(-1, 0)$, the relative efficiency of the EW D-optimal half-fractional design $\mathbf{p}_e$ is 98%.

Windshield experiment: optimal half-fractional design

Row	A	B	C	D	η	π	$\mathbf{p}_{HN}$	$\mathbf{p}_a$	$\mathbf{p}_e$
5	+1	−1	+1	+1	-0.87	0.295		0.044	0.184
1	+1	+1	+1	+1	-0.61	0.352	0.125	0.178	0.011
6	+1	−1	+1	−1	-0.59	0.357	0.125	0.178	0.011
2	+1	+1	+1	−1	-0.33	0.418		0.059	0.184
7	+1	−1	−1	+1	0.73	0.675	0.125	0.163	
3	+1	+1	−1	+1	0.99	0.729			0.195
8	+1	−1	−1	−1	1.01	0.733			0.195
4	+1	+1	−1	−1	1.27	0.781	0.125	0.147	
13	−1	−1	+1	+1	2.27	0.906	0.125	0.158	0.111
9	−1	+1	+1	+1	2.53	0.926			
14	−1	−1	+1	−1	2.55	0.928			
10	−1	+1	+1	−1	2.81	0.943	0.125	0.074	0.110
15	−1	−1	−1	+1	3.87	0.980			
11	−1	+1	−1	+1	4.13	0.984	0.125		
16	−1	−1	−1	−1	4.15	0.984	0.125		
12	−1	+1	−1	−1	4.41	0.988			

Conclusions

- We consider the problem of obtaining locally D-optimal designs for experiments with fixed finite set of design points under generalized linear models.
- 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 (not presented here, see Yang, Mandal and Majumdar, 2012b).
- We also study the robustness of the D-optimal designs.

References

1. Hamada, M. and Nelder, J. A. (1997). Generalized linear models for quality-improvement experiments, *Journal of Quality Technology*, **29**, 292−304.
2. McCullagh, P. and Nelder, J. (1989). *Generalized Linear Models*, Second Edition.
3. Dobson, A.J. (2008). *An Introduction to Generalized Linear Models*, 3rd edition, Chapman and Hall/CRC.
4. González-Dávila, E., Dorta-Guerra, R. and Ginebra, J. (2007). On the information in two-level experiments, *Model Assisted Statistics and Applications*, **2**, 173−187.
5. J. Yang, A. Mandal, and D. Majumdar (2012a). Optimal designs for two-level factorial experiments with binary response, *Statistica Sinica*, Vol. 22, No. 2, 885−907.
6. J. Yang, A. Mandal, and D. Majumdar (2012b). Optimal designs for 2^k factorial experiments with binary response, submitted for publication. Available at `http://arxiv.org/pdf/1109.5320v3.pdf`

More references

1. Zangwill, W. (1969), *Nonlinear Programming: A Unified Approach*, Prentice-Hall, New Jersey.
2. Wynn, H.P. (1970). The sequential generation of D-optimum experimental designs, *Annals of Mathematical Statistics*, **41**, 1655–1664.
3. Fedorov, V.V. (1972). *Theory of Optimal Experiments*, Academic Press, New York.
4. Müller, W. G. (2007). *Collecting Spatial Data: Optimum Design of Experiments for Random Fields*, 3rd Edition, Spinger.
5. Atkinson, A. C., Donev, A. N. and Tobias, R. D. (2007). *Optimum Experimental Designs, with SAS*, Oxford University Press.