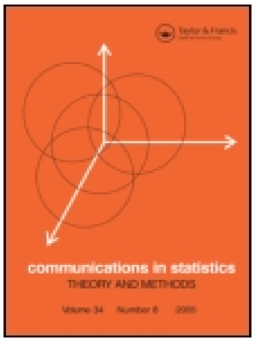

Communications in Statistics - Theory and Methods

ISSN: 0361-0926 (Print) 1532-415X (Online) Journal homepage: http://www.tandfonline.com/loi/lsta20

Minimax designs for 2^k factorial experiments for generalized linear models

Sofia Normark

To cite this article: Sofia Normark (2016) Minimax designs for 2^k factorial experiments for generalized linear models, Communications in Statistics - Theory and Methods, 45:16, 4788-4797, DOI: 10.1080/03610926.2014.927502

To link to this article: https://doi.org/10.1080/03610926.2014.927502

COMMUNICATIONS IN STATISTICS—THEORY AND METHODS
2016, VOL. 45, NO. 16, 4788–4797
http://dx.doi.org/10.1080/03610926.2014.927502

Taylor & Francis
Taylor & Francis Group

Minimax designs for 2^k factorial experiments for generalized linear models

Sofia Normark

Department of Statistics, Stockholm University, Stockholm, Sweden

ABSTRACT

Formulas for A- and C-optimal allocations for binary factorial experiments in the context of generalized linear models are derived. Since the optimal allocations depend on GLM weights, which often are unknown, a minimax strategy is considered. This is shown to be simple to apply to factorial experiments. Efficiency is used to evaluate the resulting design. In some cases, the minimax design equals the optimal design. For other cases no general conclusion can be drawn. An example of a two-factor logit model suggests that the minimax design performs well, and often better than a uniform allocation.

ARTICLE HISTORY
Received 24 July 2013
Accepted 13 May 2014

KEYWORDS
A-optimality; C-optimality; Factorial designs; Generalized linear models; Minimax designs.

MATHEMATICS SUBJECT CLASSIFICATION
62K05, 62K15

1. Introduction

Optimal design theory addresses the problem of how to design the, in some sense, "best" experiment, given a specific statistical model. In general, the model is defined by a response variable Y, with probability distribution $f_Y(y|x)$, where x is a vector of experimental variables. A specific value x is called a design point. Optimal design concerns choosing the number of design points, setting the values of the experimental variables x and deciding how many experimental subjects should be allocated to each design point, in order to obtain estimates of the model parameters that are as precise as possible.

A common type of experiment aims to investigate the effects of k binary factors on the response variable Y. To estimate the effect of all factors, and all interaction effects, 2^k designs points are needed. Constructing a design of such experiments are different from other design problems in the sense that the design points are made up by combinations of the k treatments and hence are fixed in advance. This leaves the task to find the optimum allocation of the experimental units into treatment groups. This paper uses the approach of continuous designs, which means that we focus on the proportion of experimental units assigned to each design point, rather than the exact number. For further discussions about continuous versus exact designs, see for example Atkinson et al. (2007) or Silvey (1980).

Many text books on design and analysis of experiments, such as Montgomery (2009), or Wu and Hamada (2009), routinely use a "balanced" designs, which means that int eh case of orthogonal designs, such as factorials, the experimental units are divided evenly between the groups (uniform allocation). Such designs have been shown to be D-optimal when the design points are restricted to a hypercube in R^k and when the observations are assumed to

CONTACT Sofia Normark sofia.normark@stat.su.se Dept. of Statistics, Stockholm University, SE-106 91 Stockholm, Sweden.

be normally distributed with common variance (Box and Draper, 1971). This result may not carry over to other distributions and/or if variance heterogeneity is present (Wong and Zhu, 2008).

Variance heterogeneity often occurs if the variance of Y is related to its expected value, as is the case in many distributions that may be modeled by a generalized linear model (GLM), see, e.g., McCullagh and Nelder (1989). Optimal allocation in experiments with two binary factors under such variance heterogeneity has been studied by Yang et al. (2012) (D-optimal allocations) and by Arnoldsson (1996) (A- and D-optimal allocations). This paper is partly a generalization of the results concerning A-optimality by Arnoldsson, to a higher number of factors, and explicit allocation rules for A- and C-optimality (generalized A-optimality) are given. The similarity between these and the formulas for Neyman allocation used in stratified sampling (see , e.g., Lohr, 2010, pp. 89–91) is pointed out.

Yet, it is common that the optimal design cannot be carried out since the optimal allocation depends on unknown parameters. There are a number of strategies available to resolve this issue, including simply guessing parameter values (locally optimal designs) to elaborate methods such as optimum-on-the-average designs. The latter approach has been studied for binary outcomes by Chaloner and Larntz (1989) and for GLM:s in general by Pettersson and Nyquist (2003).

However, the optimum-on-the-average approach has the drawback that it may perform poorly for some parameter values. To avoid this, the minimax strategy may be used. In the strategy of minimax design, the designer defines a region that is assumed to contain the true values of the parameters. Then a design is constructed that makes the worst possible performance of the experiment as good as possible. This approach to experiments with binary response was proposed by Sitter (1992) and expanded to various non linear models by Dette et al. (2006). Often, the minimax design is very time consuming to construct. Algorithms to compute designs that are minimax with regard to the D-criterion are given by King and Wong (2000) and Pronzato and Walter (1988). Fackle-Fornius and Nyquist (2015) show that the minimax strategy is particularly easy to apply to treatment group-type of experiments, under the assumption of independence between groups. In this paper, it is shown to be true for 2^k factorial experiments as well.

The paper is organized as follows. In Section 2, the setup for the experiments considered in this paper and criterion functions for A- and C-optimality are given. Formulas for the optimal allocation are given in Section 3. In Section 4, it is shown how to apply the minimax strategy to these results. Section 5 discusses efficiency of the minimax design by comparing it to locally optimal designs and to designs using uniform allocation. A summary and some concluding remarks are found in Section 6.

2. Preliminaries

2.1. Model and information

Consider a full factorial experiment that will yield observations on Y_{ij}, $i = 1, 2, \ldots, n$, $j = 1, \ldots, N_i$, where $n = 2^k$, N_i is the number of experimental units assigned to group i and $\sum_{i=1}^{2^k} N_i = N$. In each group the random variables Y_{ij} are identically distributed as a random variable Y_i having a distribution from the exponential family, i.e., with probability density that can be written as

$$f_{Y_i}\left(y_i|\theta_i\right) = exp\left\{a(y_i)b(\theta_i) + c(\theta_i) + d(y_i)\right\},$$

where a, b, c and d are known functions.

The expectation of Y_i, $E[Y_i]$ is denoted μ_i and the variance of Y_i, $Var(Y_i)$ is denoted σ_i^2. The model is a GLM if the following points apply

1. Response variables $Y_1, Y_2, ..., Y_n$ are assumed to come from the same exponential family of distributions
2. There is a vector of parameters β and explanatory variables

$$X = \begin{bmatrix} x_1^T \\ \vdots \\ x_n^T \end{bmatrix} = \begin{bmatrix} x_{11} & \dots & x_{1p} \\ \vdots & & \vdots \\ x_{n1} & & x_{np} \end{bmatrix}$$

3. There is a monotone link function $g(\mu_i)$ such that $g(\mu_i) = x_i^T\beta = \eta_i$.

In the case of the 2^k factorial experiment, the matrix of experimental variables X is a $2^k \times 2^k$ indicator matrix where x_{ij} indicates presence or absence of a treatment or a combination of treatments. The model for a corner point design of such an experiment is

$$\eta_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_k x_{2i} + \beta_{1,2} x_{1,2i} + \cdots + \beta_{k-1,k} x_{k-1,k,i} + ... + \beta_{1,2,...,k} x_{1,2,...k,i},$$

where β_0 corresponds to the outcome in the group receiving none of the treatments of the experiment (i.e., placebo or other reference treatment). $\beta_1, ...\beta_k$ are the main effects of treatment $1, \dots, k$, and $\beta_{1,2}, \beta_{1,3}, \dots \ \beta_{k-1,k}$ are the first order interaction effects, and so on, up to $\beta_{1,2,...,k}$, the k : th order interaction effect.

If $Y_1, Y_2, \dots, Y_{2^k}$ are independent random variables, the standardized information matrix for the parameters $\beta_0, \dots, \beta_{1,...,k}$, is given by

$$M = X^T W X,$$

where X is the matrix of explanatory variables as described above, and W is a diagonal matrix with elements $w_{ii} = \upsilon_i \omega_i$. $\upsilon = \{\upsilon_1, \dots, \upsilon_{2^k}\}$ are the GLM weights given by

$$\upsilon_i = \frac{1}{\sigma_i^2}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2$$

and $\omega = \{\omega_1, \dots, \omega_{2^k}\}$ are the design weights, i.e., $\omega_i = \frac{N_i}{N}$.

2.2. Design and optimality criteria

As described in Section 1, a design of an experiment consists of a set of experimental conditions x and the proportion of experimental units assigned to that condition, the design weight. A design can be described by the table

$$\xi = \begin{Bmatrix} x_1 & x_2 & \dots & x_n \\ \omega_1 & \omega_2 & \dots & \omega_n \end{Bmatrix},$$

where ω_i denotes the design weight of the i:th group.

An optimal design of an experiment is a design that minimizes some function of the variance of the parameter estimators. This is done by finding a design that minimizes the criterion function, which is some function of the information matrix. In general, the criterion function is written as $\Psi(\xi, \theta)$, since it depends on the design as well as on some parameter θ. An optimal design is denoted

$$\xi^* = \arg \min_{\xi} \Psi(\xi, \theta).$$

More specifically, an A-optimal design minimizes the sum (or average) of the variance of the parameter estimators. This is done by minimizing the trace of the inverted information matrix, which gives the criterion function

$$\Psi_A(\xi, \theta) = \mathrm{tr} M(\xi)^{-1}. \tag{1}$$

A-optimality can be viewed as a special case of C-optimality (generalized A-optimality). C-optimality minimizes the sum (or average) variance for some linear combinations of the parameters. The criterion function for C-optimality is

$$\Psi_C(\xi, \theta) = \mathrm{tr}[A^T M(\xi)^{-1} A], \tag{2}$$

where A is a $2^k \times m$ matrix defining the linear combinations. If A is the identity matrix, C-optimality equals A-optimality.

3. Locally A- and C-optimal allocations

As pointed out in Section 1, optimal design for factorial experiments is, in fact, an allocation problem. This means that in order to find an A- or C-optimal design for a 2^k factorial experiment, we need to find the design weights that minimize the expressions given in (1) and (2), respectively. To help with that, Lemma 1 gives an expression for the diagonal elements of $\left(XX^T\right)^{-1}$, and $X^{T^{-1}} A A^T X^{-1}$, respectively, which will come in handy when deriving expressions for the optimal allocations. The proofs of the theorems presented can be found in Appendix A.

Lemma 1. *Let s_i denote the i:th diagonal element in $X^{T^{-1}} A A^T X^{-1}$. Each s_i equals the sum of the squared elements in the corresponding column in the matrix $A^T X$.*

Theorem 2 gives the criterion for A- and C-optimality.

Theorem 2. *The criterion for A- or C-optimality is*

$$\sum_{i=1}^{2^k} \frac{s_i}{\upsilon_i \omega_i}. \tag{3}$$

From the results of Theorem 2, it is easy to find the optimal allocation.

Corollary 3. *The A- or C- optimal allocation weights ω_i^*, $i = 1, \ldots, 2^k$ are given by*

$$\omega_i^* = \sqrt{\frac{s_i}{\upsilon_i}} \left(\sum_{j=1}^{2^k} \sqrt{\frac{s_j}{\upsilon_j}} \right)^{-1}. \tag{4}$$

One may notice the similarity between (4) and the formula for Neyman allocation used in stratified sampling with υ corresponding to the inverse of the stratum variance and s to the size of the stratum.

Example 1. Suppose that we want to estimate the effect of two treatments and their interaction on a Bernoulli distributed variable, using a logit model. For a logit model the GLM weights are $\upsilon_i = \sigma_i^2 = \mu_i(1-\mu_i)$. If, for instance, we have $\upsilon = \{.15,\ .15,\ .25,\ .25\}$ and $X = \begin{bmatrix} 1&0&0&0 \\ 1&1&0&0 \\ 1&0&1&0 \\ 1&1&1&1 \end{bmatrix}$. From Corollary 4 yield the A-optimal allocation $\omega^* = \{0.3785,\ .2676,\ .2073,\ .1466\}$.

4. The minmax strategy

In the previous section, we found a formula for the optimal allocation to treatment groups. A problem for using this is that it depends on the GLM weighs υ_i that in turn depend on parameters that are unknown and/or are to be estimated by the experiment. To emphasize the dependence of υ_i on some unknown parameters, it may be denoted $\upsilon_i(\theta_i)$. An approach to solve this problem in a way that is robust against unfavorable values of $\upsilon_i(\theta_i)$ is to use a minimax strategy. This is done by defining a region $\Theta_0 \subset \Theta$ of possible parameter values and compute a design that is minimax over Θ_0. This is, finding a design that minimizes

$$\max_{\theta \in \Theta_0} \Psi(\xi, \theta). \tag{5}$$

Consider the case where the parameter θ_i (that gives the GLM weight $\upsilon(\theta_i)$) is believed to belong to the set Θ_{0i} and the set Θ_0 is the Cartesian product $\Theta_0 = \Theta_{01} \times \Theta_{02} \times ... \times \Theta_{02^k}$. The theorem below shows that maximization of the criterion functions for A- and C-optimality (1) and (2) over $\Theta_0 \subset \Theta$ is obtained by minimizing each of the GLM weights $\upsilon_i(\theta_i)$ over Θ_{0i}.

Theorem 4. *The criterion function Ψ given in (3) is maximized by replacement of each element in $\upsilon = \{\upsilon_1, \ldots, \upsilon_{2^k}\}$ by the corresponding element in $\delta = \{\delta_1, \ldots, \delta_{2^k}\}$ where*

$$\delta_i = \min_{\theta_i \in \Theta_{0i}} \upsilon_i(\theta_i),\ i = 1, \ldots 2^k.$$

From the results of Theorem 4, we may find the minimax allocations, either by Lagrange optimization or by simply replacing each υ_i, by δ_i in expression (4) and find

$$\omega_i^{\text{minimax}} = \sqrt{\frac{s_i}{\delta_i}} \left(\sum_{j=1}^{2^k} \sqrt{\frac{s_j}{\delta_j}} \right)^{-1}. \tag{6}$$

Note that for some GLM:s, e.g., the logit model, $\theta_1 ... \theta_{2^k}$ are functions of the parameters $\beta = \{\beta_0, \beta_1, ...\}$. This implies that one cannot freely choose $\theta_1 \in \Theta_{01},\ \theta_2 \in \Theta_{02}$, and so on . However, a set Θ_0 based on β may be very complicated to construct and this paper considers the simplification that Θ_0 is separable.

Example 2. Consider the experiment in Example 1 again, but assume that we do not know the true values of υ. First, recall that in a logit model $\upsilon_i = \mu_i(1-\mu_i) = \upsilon_i(\theta_i)$. To use the minimax strategy as described above, assume that the true values of μ belong to the following intervals: $\Theta_{01} = [.1; .3], \Theta_{02} = [.5; .9]$, $\Theta_{03} = [.6; .8]$, and $\Theta_{04} = [.6; .95]$, which gives

$\delta = \{.09, .09, .0.16, .0475\}$. We can now obtain the minimax design weights $\omega^{\text{minimax}} = \{.1798, .1272, .3016, .3914\}$.

5. Efficiency comparisons

To evaluate the performance of the minimax design, it is compared to a locally optimal design and to a design using a uniform allocation. The A- and the C-efficiency is measured as the ratio between the average variance of each design,

$$eff(\xi_1, \xi_2) = \frac{\Psi(\xi_1, \theta)}{\Psi(\xi_2, \theta)}. \tag{7}$$

A ratio smaller than 1 implies that ξ_1 is the better design. If, for instance, the ratio is 0.5 it means that the design ξ_2 requires twice as many observations in order to obtain estimates as precise as if ξ_1 had been used.

5.1. Efficiency when $\frac{\delta_i}{\upsilon_i} = r$ in all groups

Consider the case when there is a constant ratio between the true GLM weights υ_i and the GLM weights used to construct the minimax design δ_i, i.e., all $\frac{\delta_i}{\upsilon_i} = r$, $i = 1, \ldots, 2^k$. Theorem (5) shows that if this is the case, the optimal and minimax allocations are the same.

Theorem 5. *If the ratio $\frac{\delta_i}{\upsilon_i} = r$ $i = 1, \ldots, 2^k$ is constant over all groups, then $\omega_i^{minimax} = \omega_i^*$ for $i = 1, \ldots, 2^k$*

This may be used to compare the minimax design to the design using uniform allocation. From the result above follows that if r is constant across the groups, the minimax design performs at least as well than a design using an uniform allocation.

5.2. Efficiency when $\frac{\delta_i}{\upsilon_i}$ vary across groups

If the ratio $\frac{\delta_i}{\upsilon_i}$ vary across the treatment groups, no general conclusions can be drawn about the efficiency of the minimax design, as it will depend on how large these differences are and in which group or groups the largest difference occur. With the purpose to examine how the efficiency vary as the values of υ and δ are varied one must turn to empirical results. Below, efficiency of the minimax design for a 2^2 experiment where a Bernoulli distributed response variable is modeled by a logit link function is, examined.

Example 3. First, let the true values of the GLM weights υ_i be either 0.15 or 0.25 (both are plausible values considering the model). Then, the GLM weights are believed to be at most 0.5 or 0.2 of the respective value, i.e., $\delta_i = 0.5 * \upsilon_i$ or $0.2 * \upsilon_i$. These values were varied across the treatment groups and were used to construct minimax designs. The average variance were computed for each design and compared to the optimal design for the corresponding true GLM weights.

Naturally, the minmax design is always less efficient than a design that is locally optimal, but for the cases described above, the efficiency was at least 0.95. This can be compared to the uniform allocation, whose efficiency compared to the optimal allocation varied between 0.89 to 0.97. The full result of this example is viewed in Tables A1 and A2.

6. Summary and conclusions

As is seen in the previous sections, finding the optimal design of a 2^k factorial experiment is the same as finding the optimal allocation of the experimental units into treatment groups. If there is variance heterogeneity, such as in the case with many GLMs, the optimal allocation depends on the GLM weights. In general, these are not known prior to analyzing data from the experiment. One way to solve this issue is to use minimax designs which are shown to be easy to construct for factorial experiments. It is demonstrated that for some cases, this strategy can be used with very small loss of efficiency. Although one should be careful about drawing conclusions from single examples, this study indicates that the minimax design performs close to optimal even if the minimax GLM weights are quite far from the true values, at least in experiments involving a small number of treatments. In many cases, the efficiency loss is smaller than if the uniform allocation was used.

Further topics to study in this area may be designing for minimax loss of efficiency. The results of this paper may also be extended to experiments with two or more response variables.

Acknowledgment

The author would like to thank the referees for their thorough reading and valuable comments on earlier versions of this paper.

Appendix A

Proof of Lemma 1

Proof. First, note that X is a is a $2^k \times 2^k$ triangular matrix. The first row in X consist of a 1 in the first column, followed by zeros. The next $\binom{k}{1}$ rows consist of a 1 in the first column and exactly one 1 in one of the following $\binom{k}{1}$ columns, indicating main treatment. The following $\binom{k}{2} \dots \binom{k-1}{k}$ rows consist of a 1 in the first column followed by 1:s the columns indicating a treatment or a treatment combination. The last row consists entirely of 1:s.

Because X is triangular, its inverse X^{-1} is particularly easy to find by Gauss–Jordan elimination. The first row in X^{-1} will obviously be 1 followed by zeros. Then follows $\binom{k}{1}$ rows, with a 1 in the column where X had a 1, and a -1 in the first column. All other entries in these rows are 0. For $k \geq 2$, the next $\binom{k}{2}$, $\binom{k}{3}$, ..., $\binom{k}{k}$ rows will consist of a 1 in the column corresponding to that interaction effect, -1:s in the columns that correspond to the effects of the previous order involved, then 1:s in the columns that corresponds to the involved effects of the level below and so on. Again, other entries are 0. The last row will consist of entirely of 1:s and −1:s. Each column will have the same number of non zero entries as the number of times the corresponding treatment or combination of treatments occurs in the design. Matrix multiplication gives

$$(X^{-1})^T X^{-1} = \begin{pmatrix} \sum (x^{j1})^2 & \sum x^{j1}x^{j2} & \dots & \sum x^{j1}x^{j2^k} \\ \sum x^{j2}x^{j1} & \sum (x^{j2})^2 & \dots & \sum x^{j2}x^{j2^k} \\ \vdots & \vdots & & \vdots \\ \sum x^{j2^k}x^{j1} & \sum x^{j2^k}x^{j2} & & \sum (x^{j2^k})^2 \end{pmatrix},$$

where x^{ij} is the element ij in X^{-1}, and shows that the i:th diagonal element in the $(X^{-1})^T X^{-1}$ equals $\sum_j (x^{ji})^2$ This is the number of non zero entries in column i in X^{-1}, which is the same as the sum of column i in X, as stated in Lemma 1.

The proof for finding the diagonal elements in $X^{T-1}AA^T X^{-1}$ is similar. □

Proof of Theorem 2

Proof. The criterion function given in (2) may be rewritten as $tr(X^{T-1}AA^T X^{-1} W^{-1})$, due to the cyclic property of the trace operator. The diagonal elements of $X^{T-1}AA^T X^{-1}$ are given by Lemma 1. Since W^{-1} is a diagonal matrix, the diagonal elements of their product equal the products of their diagonal elements, which can be multiplied by the diagonal elements of $X^{T-1}AA^T X^{-1}$ to get the result stated in the theorem. □

Proof of Theorem 4

Proof. The criterion function for A- or C-optimality is $\Psi(\xi, \theta) = \sum_{i=1}^{2^k} \frac{s_i}{\upsilon_i \omega_i}$. It is easy to see that this expression is maximized when each term $\frac{s_i}{\upsilon_i \omega_i}$ is maximized, which in turn is maximized by the smallest possible value of each υ_i. Hence, $\Psi(\xi, \theta)$ is maximized when each υ_i is replaced by $\delta_i = \min_{\theta_i \subset \Theta_{0i}} \upsilon_i(\theta_i)$. □

Proof of Corollary 3

Proof. The optimum allocation is found by Lagrange optimization of 3 under the constraint $\sum_{i=1}^{2^k} \omega_i = 1$. This yields the formula for optimal allocation to treatment group i, denoted ω_i^*. □

Proof of Theorem 5

Proof. If all $\frac{\delta_i}{\upsilon_i} = r$, then all $\delta_i = r\upsilon_i$ and

$$\omega_i^{\text{minimax}} = \frac{\sqrt{\frac{s_i}{\delta_i}}}{\sum_{j=1}^{2^k} \sqrt{\frac{s_j}{\delta_j}}} = \frac{\frac{1}{\sqrt{r}}\sqrt{\frac{s_i}{\upsilon_i}}}{\frac{1}{\sqrt{r}} \sum_{j=1}^{2^k} \sqrt{\frac{s_j}{\upsilon_j}}} = \omega_i^*$$

for $i = 1, \ldots, 2^k$.

□

Full Results of Example 3

Table A1. Efficiency of minimax allocation, compared to optimal allocation.

Values of υ				Values of $r_1, r_2, r_3, r_4, (\delta_i = \upsilon_i r_i)$				
υ_1	υ_2	υ_3	υ_4	$\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{5}$	$\frac{1}{2}, \frac{1}{2}, \frac{1}{5}, \frac{1}{2}$	$\frac{1}{2}, \frac{1}{2}, \frac{1}{5}, \frac{1}{5}$	$\frac{1}{2}, \frac{1}{5}, \frac{1}{2}, \frac{1}{2}$	$\frac{1}{2}, \frac{1}{5}, \frac{1}{2}, \frac{1}{5}$
0.15	0.15	0.15	0.15	0.9705	0.9622	0.9507	0.9622	0.9507
			0.25	0.9752	0.9613	0.9516	0.9613	0.9516
		0.25	0.15	0.9692	0.9671	0.9521	0.9608	0.9500
			0.25	0.9740	0.9661	0.9534	0.9598	0.9507
	0.25	0.15	0.15	0.9692	0.9608	0.9500	0.9670	0.9521
			0.25	0.9740	0.9598	0.9507	0.9661	0.9534
0.25	0.15	0.25	0.15	0.9672	0.9650	0.9507	0.9588	0.9494
			0.25	0.9721	0.9639	0.9518	0.9577	0.9497
		0.15	0.15	0.9687	0.9603	0.9498	0.9603	0.9498
			0.25	0.9734	0.9592	0.9504	0.9592	0.9504
	0.25	0.25	0.15	0.9657	0.9634	0.9450	0.9634	0.9500
			0.25	0.9705	0.9622	0.9507	0.9622	0.9507
Values of υ				Values or r				
υ_0	υ_1	υ_2	υ_4	$\frac{1}{5}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}$	$\frac{1}{5}, \frac{1}{2}, \frac{1}{2}, \frac{1}{5}$	$\frac{1}{5}, \frac{1}{2}, \frac{1}{5}, \frac{1}{2}$	$\frac{1}{5}, \frac{1}{2}, \frac{1}{2}, \frac{1}{5}$	$\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{2}$
0.15	0.15	0.15	0.15	0.9507	0.9622	0.9541	0.9494	0.9705
			0.25	0.9516	0.9613	0.9533	0.9493	0.9752
		0.25	0.15	0.9500	0.9608	0.9529	0.9497	0.9692
			0.25	0.9507	0.9598	0.9522	0.9494	0.9740
	0.25	0.15	0.15	0.9521	0.9671	0.9529	0.9497	0.9692
			0.25	0.9534	0.9661	0.9522	0.9494	0.9740
0.25	0.15	0.25	0.15	0.9494	0.9588	0.9566	0.9493	0.9672
			0.25	0.9497	0.9577	0.9556	0.9494	0.9721
		0.15	0.15	0.9498	0.9603	0.9580	0.9494	0.9687
			0.25	0.9504	0.9592	0.9570	0.9498	0.9734
	0.25	0.25	0.15	0.9500	0.9634	0.9551	0.9496	0.9657
			0.25	0.9507	0.9622	0.9541	0.9494	0.9705

Table A2. Efficiency of uniform allocation, compared to optimal allocation.

υ_0	υ_1	υ_1	$\upsilon_{1,2}$	Efficency$(\xi_{\text{uniform}}, \xi^*)$
0.15	0.15	0.15	0.15	0.9436
0.15	0.15	0.15	0.25	0.9126
0.15	0.15	0.25	0.15	0.9255
0.15	0.15	0.25	0.25	0.8950
0.15	0.25	0.15	0.15	0.9255
0.15	0.25	0.15	0.25	0.8945
0.25	0.15	0.25	0.15	0.9694
0.25	0.15	0.25	0.25	0.9420
0.25	0.15	0.15	0.15	0.9770
0.25	0.15	0.15	0.25	0.9481
0.25	0.25	0.25	0.15	0.9685

References

Arnoldsson, G. (1996). Optimal allocation for linear combinations of linear predictors in generalized linear models. *J. Appl. Stat.* 23:493–505.

Atkinson, A., Donev, A., Tobias, R. (2007). *Optimum Experimental Designs, With SAS*. New York, NY: Oxford University Press Inc.

Box, M.J., Draper, N.R. (1971). Factorial designs, the X'X| criterion, and some related matters|. *Technometrics* 13(4):731–742.

Chaloner, K., Larntz, K. (1989). Optimal Bayesian design applied to logistic regression experiments. *J. Stat. Plann. Inference* 21:191–208.

Dette, H., Lopez, I.M., Ortiz Rodriguez, I.M., Pepelyshev, A. (2006). Maximin efficient design of experiment for exponential regression models. *J. Stat. Plann. Inference* 136(12):4397–4418.

Fackle-Fornius, E., Nyquist, H. (2015). Optimal allocation to treatment groups under variance heterogeneity. To appear in *Statistica Sinica* 25:537–549.

King, J., Wong, W.K. (2000). Minimax D-optimal designs for the logistic model. *Biometrics* 56:1263–1267.

Lohr, S.L. (2010). *Sampling: Design and Analysis* (2nd ed.). Boston: Brooks/Cole.

McCullagh, P., Nelder, J. (1989). *Generalized Linear Models* (2nd ed.). Boca Raton, FL: Chapman and Hall.

Montgomery, D.C. (2009). *Design and Analysis of Experiments*. Hoboken: John Wiley and Sons.

Pettersson, H., Nyquist, H. (2003). Computation of optimum in average designs for experiments with finite design space. *Commun. Stat.-Simul. Comput.* 32(1):205–221.

Pronzato, L., Walter, E. (1988). Robust experiment design via maximin optimization. *Math. Biosci.* 89(2):161–176.

Silvey, S. (1980). *Optimal Design*. London: Chapman and Hall.

Sitter, R.R. (1992). Robus designs for binary data. *Biometrics* 2:1145–1155.

Wong, W.K., Zhu, W. (2008). Optimum treatment allocation rules under a variance heterogenity model. *Stat. Med.* 27:4581–4595.

Wu, J.C.F., Hamada, M.S. (2009). *Experiments*. Hoboken: John Wiley and Sons.

Yang, J., Mandal, A., Majumdar, D. (2012). Optimal designs for two-level factorial experiments with binary response. *Stat. Sin.* 22:554–907.