International Journal of New Technology and Research (IJNTR)
ISSN:2454-4116, Volume-4, Issue-5, May 2018 Pages 31-36

Towards a better Estimation of the Parameters of Linear Regression Models: The Optimal Designed Experiment Approach

FASORANBAKU O.A., DARAMOLA G.O.

Abstract- **This study investigates and compares Optimal designed Experiment with classical design which is non-Optimal using statistical tools. The two designs were evaluated on the basis of six parameters; viz, Information matrix, Dispersion matrix, Prediction variance, A-Efficiency, D-Efficiency and G-efficiency. These six parameters help to determine the better design of the two experiment. Thus, it helps to establish efficient experiment suitable for better estimation of parameters of Linear Regression Model. The result obtained in this research work showed that the D-Optimal design increased the A-Efficiency, D-Efficiency and the G-efficiency of the Initial non optimal design. Furthermore the D-Optimal design maximized the determinant of the Information Matrix, Minimized the determinant of the Dispersion matrix and minimized the trace of the Dispersion matrix. It was therefore established in this research work that the D-Optimal Design Experiment has higher statistical efficiency than the initial non-optimal design. Moreso statistical analysis of the model parameters for both designs established the D-Optimal design experiment produce better models when used for estimating the parameters of Linear Regression models. It is therefore suggested that D-Optimal approach is suitable for fixing a poorly designed experiment. It is therefore recommended for use in estimating the parameters of Linear Regression Models.**

Index Terms— **Algorithm, Optimal Design, Model, Optimality Criterion, Flowchart, Optimization, Efficiency, Experimental Design.**

I. INTRODUCTION

Kiefer (1959) presented a paper to the Royal Statistical Society about his work on the theory of optimal design. During the presentation, he tried answering this major question "How do we find the best design?" This work initiated a whole new field of optimal design. According to Ramachandran and Chris (2009) optimal experimental design provide the technical tools for building experimental designs to attain well-defined objectives with efficiency and with minimum cost. The cost can be monetary, time, number of experimental runs, and so on. The theory of optimal experimental design as explained in the monographs of Fedorov (1972) clarified that given the total number of observations, the optimal design is determined by the design space (experimental region), the regression model and the optimality criterion. Searching for these optimal design yields challenging optimization problems. Experiments are therefore carried out in order to estimate parameters of regression models. Optimal experimental designs are therefore used to maximize the precision of the least squares estimator, given the total number of observations. Optimal designs are a class of experimental designs that are optimal with respect to some statistical criterion. When estimating statistical models, optimal designs allow parameters to be estimated without bias and with minimum variance.

According to Berger and Wong (2009), there are at least three ways to design an efficient experiment for better estimation of parameters of linear regression model. First, efficiency can be improved by measuring the dependent variable more accurately. A second way to improve efficiency is to increase the total sample size N, this however will lead to more costs for collecting the data and running the study. A third way to improve efficiency is to select the levels of the independent variable X in such a way that their information content will be as large as possible. i.e. | X'X| will be maximized.

Careful selection of the levels of the independent variable can be done by using algorithms to search for the optimal design of the experiment. Algorithms play important role in generating optimal designs of experiments because they help to reduce number of experimental runs required to estimate the parameters, and thereby reduce the costs of experimentation. Algorithms uses optimality criteria for the careful selection of the levels of independent variable. The D-Optimality criteria was used to achieve results in this research work because it has a number of advantages such as its invariance under linear transformations of the scale of the independent variable, especially when different scales of the independent variable are implemented in different studies.

Researchers have at sundry times developed algorithms through modifications to existing exchange algorithms and thereby generating optimal designs. For instance, Yang *et al.* (2013) used Fedorov's exchange algorithm, originally published in 1969 (see also Fedorov, 1972), to obtain optimal designs for generalized linear models. There are several versions of the exchange algorithms where the search begins with a single random design, and then each design point is considered for exchange with other points. The pair of points chosen for exchange is the pair which results in maximum gain of the optimality criterion.

Also, Jafari *et al.* (2014) used the idea of building locally optimal designs for logistic regression models without random effects. Locally D-optimal design model was computed for logistic regression model with three independent variables for several specific states.

Fasoranbaku, O.A. (Reader) Department of Statistics, Federal University of Technology, Akure, Nigeria.
Daramola, G.O. (Ph. D student) Department of Statistics, Federal University of Technology, Akure, Nigeria.

II. STATISTICAL MODELS AND ANALYSIS

The optimality of a design depends on the statistical model and is assessed with respect to a statistical criterion, which is related to the variance-matrix of the estimator. Specifying an appropriate model and specifying a suitable criterion function both require understanding of statistical theory and practical knowledge with designing experiments. According to (Eriksson et al. 2000), a model helps to transform the complexity of reality into an equation which is easy to handle. It is therefore an important step in Optimal Design of Experiments. Aside from the linear model, which is the most common, there are other Statistical models such as interactions model, quadratic model, Non-Linear model. etc.

A. Linear Models

In a linear model, each factor appears as a linear term. In this case, a linear term means a combination of a coefficient β_i and a factor x_i. A linear model with p number of factors has the following equation:

$$y=\beta_0+\beta_1 x_1+\ldots+\beta_p x_p+\varepsilon \quad (1)$$

Where $\beta_1, \beta_2, \ldots, \beta_p$ represents the regression coefficients and ε is the random part of the model which is assumed to be normally distributed with mean 0 and variance σ^2. Wu & Hamada (2000). The p factors x, $x_2, \ldots, x_p$ influence the response y.

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_g x_{ig} + \varepsilon_i, \quad i = 1, \ldots N \quad (2)$$

where y_i stands for the ith response with the factors $x_{i1}, x_{i2}, \ldots, x_{ip}$. The corresponding matrix notation is given as:

$$Y = X\beta + \varepsilon \quad (3)$$

Where the $N \times (p+1)$ model matrix X contains all factors for the responses and Y *and* € are N × 1 vectors. The regression coefficients β *are* the unknown parameter in the model (Wu & Hamada 2000).

B. Interaction Models

Interaction models are used to achieve the same experimental objectives as their linear counterpart. They are however more complex than their linear counterparts. They contain the same terms like the linear model but have additional interaction terms. An interaction term is the combination of two factors x_i and x_j with a conjoint coefficient β_{ij}. The following equation gives an example of an interaction model with three factors.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \varepsilon \quad (4)$$

C. Quadratic Models

Quadratic model extends the interaction model with additional quadratic terms for each factor. A quadratic term is the square of a factor x_i with its coefficient β_{ii}. Quadratic models are the most complex of the three basic model types and are used for optimization processes. A quadratic model with three factors is illustrated with the following equation.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{33} x_3^2 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \varepsilon \quad (5)$$

D. Non Linear Models

In Non Linear Models observational data are modeled by a function which is a nonlinear combination of the model parameters and depends on one or more independent variables. The data are fitted by a method of successive approximations.

The data consist of error-free independent variables (explanatory variables), x, and their associated observed dependent variables (response variables), y. Each y is modeled as a random variable with a mean given by a nonlinear function $f(x,\beta)$. Thus a function is nonlinear if it cannot be expressed as a linear combination of the two βs. Non-linear regression model is given by:

$$Y = f(X, \beta) + \varepsilon \quad (6)$$

Where X is a vector of p predictors, β is a vector of k parameters, f(.) is some known regression function, and ε is an error term whose distribution may or not be normal.

III. CRITERIA USED FOR SELECTION OF OPTIMAL DESIGN OF EXPERIMENTS

There are different criteria used for optimal design. They are D-Optimality Criterion, A-Optimality Criterion, V-Optimality Criterion and G-Optimality Criterion.

A. D-Optimality Criterion (Determinant)

The D-Optimality is the most common criterion which seeks to maximize $|X'X|$, the determinant of the information matrix $X'X$ of the design. Maximizing the determinant of the information matrix $(X'X)$ is equivalent to minimizing the determinant of the dispersion matrix $(X'X)^{-1}$.

$$|X^{*I}X^*| = \min_{\varepsilon_n} |(X'X)^{-1}| \quad (7)$$

where X* is optimal design matrix and ε_n is called matrix of candidate points having n rows.

B. A-Optimality Criterion (Trace)

Another Criterion for an optimal design is called the A-criterion. The design matrix is considered as A-Optimal when the trace of the dispersion matrix $(X'X)^{-1}$ is minimum. In this case, the trace of the square matrix is the sum of the elements on the main diagonal. Minimizing the trace of the matrix is similar to minimizing the average variance of the estimated coefficients. A-Optimal designs are rarely used because it is more computationally difficult to update during the selection process.

$$\text{Trace } (X^{*\prime}X^*)^{-1} = \min_{\varepsilon_n}(trace(|X'X|)^{-1}) \quad (8)$$

$$\text{Trace } (X^{*\prime}X^*)^{-1} = \sum_{i=1}^{p} C_{ii} \quad (9)$$

C. V-Optimality Criterion (Average Prediction Variance)

3.3 V-Optimality Criterion (Average Prediction Variance)

As de Aguiar et al (1995) describe the variance function or leverage is a measurement of the uncertainty in the predicted response'. This variance of prediction for a single candidate x_i can be calculated with equation (10). Where x_i equals a vector that describes a single experiment and x_i' represents the transpose of this vector. With the selection of a V-Optimal design, the chosen candidates have the lowest average variance of prediction.

$$d(x_i) = x'_i * (X^{*\prime}X^*)^{-1} * x_i \quad (10)$$

$$\frac{1}{n}\sum_{i=1}^{n} x'_i * (X^{*\prime} * X^*)^{-1} * x_i = \min_{\xi_n}\left(\frac{1}{n}\sum_{i=1}^{n} x'_i * (X' * X)^{-1} * x_i\right) \quad (11)$$

with the selection of a V-Optimal design, the chosen candidate have the lowest average variance of prediction as shown in equation (11) above.

D. G-Optimality Criterion (Maximum Prediction Variance)

The G-Optimal design deals with the variance of prediction of the candidate points. The selected optimal design matrix chosen to minimize the highest variance of prediction in the design. This is represented with the following equation.

$$\max(x'_i * (X^{*\prime}X^*)^{-1} * x_{i=} \min(\max(x'_i * (X^{*\prime}X^*)^{-1} * x_i)) \quad (12)$$

$$\min_{\xi_n}(\max(d(x_i) = x'_i * (X^{*\prime}X^*)^{-1} * x_i) \quad (13)$$

IV. PARAMETERS FOR EVALUATING EXPERIMENTAL DESIGNS

Any two designs can be evaluated on the basis of six parameters; viz, Information matrix, Dispersion matrix, Prediction variance, A-Efficiency, D-Efficiency and G-efficiency. These six parameters help to determine the better design of the two experiment.

A. Information and Dispersion Matrix

To use the later described criteria for the selection of the best design, we need to define two other types of matrices. The first one is the so-called information matrix ($X'X$). This matrix is the multiplication of the transpose of the design matrix X' and X itself. The dispersion matrix $(X'X)^{-1}$ is the inverse matrix of this calculation (de Aguiar et al. 1995).

B. Design Matrixx

The design matrix X is n × *p* matrix that depends on a model with p coefficients. The number of rows *n* can be chosen by the experimenter and represents the number of experiments in the design. With a given model and a candidate matrix, the construction of the design matrix is easy. Each column contains a combination of the G-factors from the candidate set, depending on the terms in the model.

The matrix can also be called model matrix, but in most cases the model matrix means a *N* × *p* matrix which contains the model-dependent rows for all candidates (de Aguiar et al. 1995). Subsets of ξ_N increases and the selection of the design matrix has to be done depending on a special criterion. 'The best combination of these points is called optimal and the corresponding design matrix is called optimal design matrix' X*(de Agui The f al. 1995).

C. Optimal Design Matrix

The optimal design matrix X* contains the n experiments which maximizes the determinant of the information matrix ($X'X$) or minimizes the determinant of the dispersion matrix $(X'X)^{-1}$ or in other words, the n runs 'span the largest volume possible in the experimental region' (Eriksson et al. 2000).

$$|X^{*\prime}X^*| = \max_{\xi n} |X'X| \quad (14)$$

Or

$$|X^{*\prime}X^*| = \min_{\xi n} |(X'X)^{-1}| \quad (15)$$

V. DESIGN EFFICIENCY

Efficiencies are measures of design goodness. Common measures of the efficiency of an (N_D x P) orthogonally coded design matrix X are based on the information matrix $X'X$. The variance-covariance matrix of the vector of parameter estimates β in a least squares analysis is proportional to $(X'X)$. The variance of $\hat{\beta}_i$ is proportional to the x_{ii} element of $(X'X)^{-1}$. An efficient design will have a "small" variance matrix, and the eigenvalues of $(X'X)^{-1}$ provide measures of its "size".

A. A-Efficiency

A-efficiency is a function of the arithmetic mean of the variances, i.e it is based on the average of the variances of the parameter estimates. And it is given is given by Trace $((X'X)^{-1})/p$. (The trace is the sum of the diagonal elements of $(X'X)^{-1}$, which is the sum of the variances and is also the sum of the eigenvalues, and it is given by the following equation.

$$\text{A-Efficiency} = 100 \text{ x } \frac{p}{N_D\, trace((X'X)^{-1}} \quad (16)$$

Where P stands for the number of factor effects in the design (columns in X), N_D is the number of requested runs and trace $(X'X)^{-1}$ stands for the trace of the dispersion matrix.

B. D-Efficiency

D-efficiency is a function of the geometric mean of the eigenvalues, which is given by $|(X'X)^{-1}|^{1/p}$. The determinant $|X'X)^{-1}|$ is the product of the eigenvalues of $(X'X)^{-1}$), and the p^{th} root of the determinant is given by the following equation.

$$\text{D-Efficiency} = 100 \text{ x } \frac{1}{N_D\,|(X'X)^{-1}|^{1/p}} \quad (17)$$

Where P stands for the number of factor effects in the design (columns in X), N_D is the number of requested runs and $|(X'X)^{-1}|$ stands for the determinant of the dispersion matrix.

C. G-Efficiency

G-efficiency is based the largest variance of prediction over the candidate set. G-efficiency is mostly applied to choose between several similar designs which were created with another criterion, like D-Optimality. It is defined as follows

$$\text{G-Efficiency} = 100 \text{ x } \frac{p}{n * d_{max}(X)} \quad (18)$$

Where p is the number of model terms or coefficients, n is the number of design runs and dmax (X) is the largest variance of prediction in the model matrix X.

VI. NUMERICAL ILLUSTRATION

The following is a numerical demonstration of the concept of Optimal Design of Experiments Consider, for example, the constrained mixture experiment for estimating the impact of three factors on the electric resistivity (*Y*) of a modified acrylonitrile powder described in Atkinson and Donev (1992) and explained by Goos and Leemans (2004). The components of the mixture under investigation are:

S/n	x_1	x_2	x_3
1	0.20	0.40	0.40
2	0.20	0.60	0.20
3	0.30	0.35	0.35
4	0.40	0.20	0.40
5	0.40	0.60	0.00
6	0.45	0.45	0.10
7	0.50	0.25	0.25
8	0.60	0.20	0.20
9	0.60	0.40	0.00

Table 1: Starting Design of the Components of the Mixture

x_1 : copper sulphate ($CuSO_4$),
x_2: sodium thiosulphate ($Na_2S_2O_3$),
x_3: glyoxal $(CHO)_2$

The electric resistivity of the powder did not depend on the total amount of the mixture but only on the relative proportions of the three components. Each component is therefore restricted to lie between 0 and 100%, i.e. $0 \leq x_i \leq 1$. In addition, the proportions in a mixture experiment have to add up to 100%, so that

$$x_1 + x_2 + x_3 = 1 \qquad (19)$$

It was also required that

$$0.2 \leq x_1 \leq 0.8,$$

$$0.2 \leq x_2 \leq 0.8,$$

$$0.0 \leq x_3 \leq 0.6.$$

Now, assume that the model is given by the first order Scheffè polynomial

$$Y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon_i, \qquad (20)$$

and that nine observations are available for estimating this model. For the first order Scheffè model, the design matrix and the extended design matrix are identical:

$$\mathbf{X} = \mathbf{F} = \begin{bmatrix} x_{11} & x_{21} & x_{31} \\ x_{12} & x_{22} & x_{32} \\ \vdots & \vdots & \vdots \\ x_{19} & x_{29} & x_{39} \end{bmatrix}.$$

The Starting Design Matrix X is given as follows:

$$\begin{pmatrix} 0.20 & 0.40 & 0.40 \\ 0.20 & 0.60 & 0.20 \\ 0.30 & 0.35 & 0.35 \\ 0.40 & 0.20 & 0.40 \\ 0.40 & 0.60 & 0.00 \\ 0.45 & 0.45 & 0.10 \\ 0.50 & 0.25 & 0.25 \\ 0.60 & 0.20 & 0.20 \\ 0.60 & 0.40 & 0.00 \end{pmatrix}$$

The Information matrix of the starting design is given as follows:

$$X^IX = \begin{pmatrix} 1.6625 & 1.3125 & 0.6750 \\ 1.3125 & 1.5075 & 0.6750 \\ 0.6750 & 0.6300 & 0.5950 \end{pmatrix}$$

The Determinant of the information matrix of the starting Design is 0.2358. i.e. $|X^IX| = 0.2358$

The Dispersion matrix of the starting design is given as follows

$$(X^IX)^{-1} = \begin{pmatrix} 2.1207 & -1.5084 & -0.8087 \\ -1.5084 & 2.2628 & -0.6846 \\ -0.8087 & -0.6846 & 3.3230 \end{pmatrix}$$

The Determinant of the Dispersion matrix of the starting Design is 2409 i.e. $|(X^IX)^{-1}| = 4.2409$

The Trace of the Dispersion matrix of the starting Design is 7.7065 i.e. TRACE $(X^IX)^{-1} = 7.7065$

Table 2: D-Optimal Design obtained by

S/n	x_1	x_2	x_3
1	0.20	0.20	0.60
2	0.20	0.20	0.60
3	0.20	0.20	0.60
4	0.20	0.80	0.00
5	0.20	0.80	0.00
6	0.20	0.80	0.00
7	0.80	0.20	0.00
8	0.80	0.20	0.00
9	0.80	0.20	0.00

Computer algorithm

$$X^* = \begin{pmatrix} 0.20 & 0.20 & 0.60 \\ 0.20 & 0.20 & 0.60 \\ 0.20 & 0.80 & 0.00 \\ 0.20 & 0.80 & 0.00 \\ 0.20 & 0.80 & 0.00 \\ 0.80 & 0.20 & 0.05 \\ 0.80 & 0.20 & 0.00 \\ 0.80 & 0.20 & 0.00 \end{pmatrix} =$$

Optimal Design Matrix X* is shown above
The information matrix of the starting design is given as follows

$$X^{*I}X^* = \begin{pmatrix} 2.1600 & 1.0800 & 0.3600 \\ 1.0800 & 2.1600 & 0.3600 \\ 0.3600 & 0.3600 & 1.0800 \end{pmatrix}$$

The Dispersion matrix of the optimal design is given as follows

$$(X^{*I}X^*)^{-1} = \begin{pmatrix} 2.1207 & -1.5084 & -0.8087 \\ -1.5084 & 2.2628 & -0.6846 \\ -0.8087 & -0.6846 & 3.3230 \end{pmatrix}$$

Table 3: Comparison of matrix information

S/n	Parameter	Starting Design Matrix (X)	Optimal Design Matrix (X*)
1	No of Rows of Model Matrix	9	9
2	Determinant of Information Matrix	0.2358	3.4992
3	No of Columns of the Dispersion Matrix	3	3
4	Determinant of the Dispersion Matrix	4.2409	0.2858
5	Trace of the Dispersion Matrix	7.7065	2.2593
6	Maximum Variance of Prediction	0.5514	0.3556

Table 4: Comparison of Efficiencies of Design

S/n	Parameter	Starting Design Matrix (X)	Optimal Design Matrix (X*)
1	A-Efficiency	4.3253	14.7538
2	D--Efficiency	6.8644	16.8683
3	G-Efficiency	60.4521	93.7382

VII. RESULTS AND COMPARATIVE STUDY

The determinant of the information matrix of the starting design as shown in table 3 is 0.2358 while the determinant of the information matrix of the D-Optimal design as shown in the same table is 3.4992. This reveals that the D-Optimal Design actually maximized the determinant of the information matrix. Meanwhile, the determinant of the dispersion matrix of the starting design as shown in table 3 is 4.2409 while the determinant of dispersion matrix of the D-Optimal design as shown in the same table is 0.2858. This reveals that the D-Optimal Design actually minimized the determinant of the dispersion matrix. The trace of the dispersion matrix of the starting design shown in table 3 is 7.7065 while the trace of the dispersion matrix of the D-Optimal design as shown in the same table is 2.2593. This reveals that the D-Optimal Design minimized the trace of the

dispersion matrix. Moreso, the maximum variance of prediction of the starting design as shown in table 3 is 0.5514 while the maximum variance of prediction of the D-Optimal design is 0.3556. This reveals that the D-Optimal Design matrix minimized the highest variance of prediction of the design Table 4 shows that the D-Optimal design increased the A-Efficiency of the initial design from 4.3253 to 14.7538, it also increased the D-Efficiency of the initial design from 6.8644 to 16.8683. The D-Optimal Design equally increased the G-Efficiency of the initial Design from 60.4521 to 93.7382.

Thus, in view of the above, the D-Optimal Design is the improved design with higher statistical efficiency and it is therefore the most efficient design for the objective of the study. It is therefore suitable for better estimation of the parameters of the linear regression model under study.

VIII. CONCLUSION AND RECOMMENDATION

The D-Optimal Design generated by Computer Algorithm is the best design for the problem under study, it is therefore recommended for use for the following reasons.

i. It gives more efficient solutions at minimum cost so as to ensure accuracy of the estimates of the model parameters.
ii. It helps Experimental design practitioners to reduce sample size and thereby reduce the cost of experimentation.
iii. It allows the computation of more efficient solutions for challenging practical problems.
iv It maximized the resulting knowledge i.e it ensure that the selected experiment is Maximally informative.
v. It helps search for the combination of factors which will give the best yield and lowest impurities at lowest possible cost, using lowest possible raw materials with minimum experimental runs.
vi The D-Optimal design is suitable for better estimation of the model parameters.
vii The D-Optimal Design is suitable for accuracy of the estimators of the model parameters in terms of the variance of the estimators.

REFERENCES

[1] Atkinson, A.C and Donev, A.N (1992) Optimum Experimental Designs. Oxford University Press, Oxford. Chernoff, H. (1953), "Locally Optimal Designs for Estimating Parameters," *Annals of Mathematical Statistics*, 24, 586–602.

[2] Berger, M. P. F. & Wong, W. K. (2009), An Introduction to Optimal Designs for Social and Biomedical Research, John Wiley & Sons, Ltd.

[3] de Aguiar, P. F., Bourguignon, B., Khots, M. S., Massart, D. L. & Phan-Than-Luu, R. (1995),'D-optimal Designs', *Chemometrics and Intelligent Laboratory Systems* 30(2), 199–210.

[4] Eriksson, L., Johansson, E., Kettaneh-Wold, N., WikstrÄom, C. & Wold, S. (2000), Design of Experiments: Principles and Applications, Learnways AB, Umeºa.

[5] Fedorov, V. V. (1972), 'Theory of Optimal Experiments (Review)', *Biometrika* 59(3), 697–698. Translated and edited by W. J. Studden and E. M. Klimko.

[6] Goos, P. and Leemans, H. (2004), "Teaching Optimal design of Experiments using a Spreadsheet", Journal of Statistics Education. Vol 12(3), 2004.

[7] Kiefer, J. (1959), "Optimum Experimental Designs," *Journal of the Royal Statistical Society*, Ser. B, 21, 272–319.

[8] Ramachandran, K.M and Chris P. Tsokos (2009) Mathematical Statistics with Applications.

[9] Wong, W. K.(1992) "A unified approach to the construction of mini-max Designs," *Biometrika*, vol. 79, pp. 611–620, 1992.

[10] Wu, C. F. J. & Hamada, M. (2000), Design of Experiments: Planing, Analysis and Parameter Design Optimization, John Wiley & Sons, INC., USA.

[11] Yang, J., Mandal A. and Mujumdar, D. (2013). "Optimal designs for $2^{k\,factorial}$ experiments with binary response," submitted for publication, available at http://arxiv.org/pdf/1109.5320v3