Journal of the Nigerian Association of Mathematical Physics
Volume **18 (May, 2011), pp 367 – 372**

On The Choice of a Completely Randomized Design with Emphasis On The D-Optimality Criterion

[1]Iruegbukpe, D. O. and [2]Mbegbu, J. I.*

[1]Ministry of Economic Planning, Budget Department,
Asaba, Delta State
[2]Department of Mathematics, University of Benin,
Benin City, Nigeria

Abstract

Completely randomized designs (CRDs) are designs in which treatments are allotted to the experimental units by chance. The designs include the unconstrained CRD, the zero-sum constraint CRD and the baseline constraint CRD. Since there are several CRDs, we are faced with choosing a design that will yield the optimal result. In this paper, we present the use of the D-optimality criterion as an objective tool in making such a choice. The study shows that the baseline constraint CRD is the most suitable choice of designs for an experiment with five treatments and six replicates.

Keywords: Experimental design, CRD, D-optimality criterion, treatment, replicate, model, regression, information matrix, measure, eigenvalue, determinant.

1. Introduction:

Experimentation is one of the most common activities used by an investigator in order to understand and/or improve a system. The act of experimentation has received considerable attention in applied statistics and this has led to the increasing interest in experimental design. Experimental design is a body of knowledge and techniques that enable an investigator to conduct better experiments, analyze data efficiently, and make connections between the conclusions from the analysis and the original objectives of the investigation [9]. The basic principles, planning and implementation of an experiment have been documented by [9]. One of the methods to design an experiment is the use of completely randomized design (CRD). In CRD the treatments are assigned randomly to the experimental units. The reason for doing this is because of the assumption that the experimental material is relatively homogeneous and that the differences in yield from units receiving different treatments arise due to the differences between the treatments. Joshi [6] reported that the CRD is useful in laboratory experiments, in cookery, and in certain greenhouse studies. The linear model for CRD is of the form

$$y_{ij} = \eta + \tau_i + \in_{ij}, \qquad i = 1, \ldots, k; \;\; j = 1, \ldots, n_i, \tag{1}$$

where y_{ij} is the j^{th} observation with treatment i, τ_i is the i^{th} treatment effect, the errors $\in_{ij}$ are independent $N(0, \sigma^2)$ with mean 0 and variance σ^2, k is the number of treatments, and n_i is the number of observations with treatment i.

Our primary focus in this study is the use of D-optimality criterion to determine the CRD with the maximum information. The CRD considered include the unconstrained linear CRD model and the linear CRD model with the following constraints:

i. the zero-sum constraint i.e. $\sum_{i=1}^{k} \tau_i = 0$ and

ii. the baseline constraint where treatment 1 is a standard treatment and the other $(k-1)$ treatments are new treatments.

The implementation of the D-optimality criterion in doing this, is discussed.

1.1. Experimental designs in linear models

In many practical and theoretical problems the CRD model is transformed to

$$y_i = \eta(x_i, \beta) + e_i, \;\; x_i \in \aleph \subset R^k, \tag{2}$$

Corresponding authors: Mbegbu, J. I. Email: julianmbegbu@yahoo.com. Tel: +2348020740989

where η is a regression function, e_i is an independent identically distributed mean error term, β is a parameter vector with m elements, $\beta^T = (\beta_1, \ldots, \beta_m) \in \Omega \subset R^k$, and x_i is the i^{th} spectrum of the design. The choice of the function η is central to the model building process. The error term is added to the model so as to account for distortions such as random errors resulting from inaccuracies in the measuring devices and systematic errors that are due to inappropriateness of the model function. The collection of variables $x_1, x_2, \ldots, x_n, n_1, n_2, \ldots, n_n, \sum_{i=1}^{n} n_i = N$, is called the design of an experiment $\xi(N)$, which is a probability measure ξ on $\aleph$ supported by a finite set. For normalized design, $\xi(N)$ is the collection of variables $p_1, p_2, \ldots, p_n$, $x_1, x_2, \ldots, x_n$, where $\sum_{i=1}^{n} p_i = 1$ and $p_i = \frac{n_i}{N}$. Also, for a discrete design, $N(\xi) = \begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ p_1 & p_2 & \ldots & p_n \end{pmatrix}$, and for an exact design, $\xi(N) = \begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ n_1 & n_2 & \ldots & n_n \end{pmatrix}$. If the number of replications is restricted to one (i.e. a replication-free design) then the observations would have a minimal distance between each other. In this case, the design space $\aleph$ is considered as a gird of Q candidate points or possible measurement points (observations) (see [3]). So, the design problem becomes a combinatorial problem of selecting N observations from Q candidate points, and the solution of the problem will give an exact design with only one replication at each support point ($N = n$).

Consider the linear model of the form $\eta(x, \beta) = \beta^T f(x)$ with the regression assumptions of independent errors and constant variance. The information matrix for the discrete design is given as $M(\xi) = \sum_{i=1}^{n} p_i f(x_i) f^T(x_i)$ and the associated standardized variance function is $d(x, \xi) = \text{var}\left(\hat{\eta}(x, \beta)\right) = f^T(x) M^{-1}(\xi) f(x)$, where $\hat{\eta}(x, \beta)$ is the function obtained by substituting $\hat{\beta}$ into $\eta(x, \beta)$. $\hat{\beta}$ is the result of minimizing the sum of squares $S(\beta) = \left(y - f^T(x)\beta\right)^T \left(y - f^T(x)\beta\right)$.

2.0 Materials and methods

2.1 The D-optimality criterion for experimental designs. A very popular theorem in optimal experimental design literature is the General Equivalence Theorem (see [1, 5, and 7]). The use of optimality criteria in experimental design literature has gained prominence in recent times as a means of obtaining a better model for an experiment, when several experiments are being compared [4]. By the definition of information matrix and the associated variance function, [5] presented the following approaches for comparing experiments.

i. Experiment ξ_1 is preferred to experiment ξ_2 if the difference $D(\xi_2) - D(\xi_2)$ is a positive-definite matrix. That is, $\xi_1 > \xi_2$ if $D(\xi_1) < D(\xi_2)$, where $D(\bullet)$ is the dispersion matrix given as $D(\bullet) = M^{-1}(\bullet)$.

ii. $\xi_1 > \xi_2$ if $\left|D(\xi_1)\right| < \left|D(\xi_2)\right|$.

iii. $\xi_1 > \xi_2$ if $TrD(\xi_1) < TrD(\xi_2)$.

iv. $\xi_1 > \xi_2$ if $\max_{\alpha} D_{\alpha\alpha}(\xi_1) < \max_{\alpha} D_{\alpha\alpha}(\xi_2)$.

v. $\xi_1 > \xi_2$ if $D(\varphi, \xi_1) < D(\varphi, \xi_2)$, where $\varphi = L\beta$ and $D(\varphi, \xi) = LD(\xi)L^T$. L is a matrix of coefficients.

vi. $\xi_1 > \xi_2$ if for the dispersion of the respective estimators $\hat{\eta}(x, \xi_1)$ and $\hat{\eta}(x, \xi_2)$, the following inequality is satisfied: $\max_{x \in Z} d(x, \xi_1) < \max_{x \in Z} d(x, \xi_2)$, where $d(x, \xi)$ is the dispersion of the estimates $\hat{\eta}(x, \beta) = f^T(x)\hat{\beta}$ given as $d(x, \xi) = f^T(x) D(\xi) f(x)$.

vii. $\xi_1 > \xi_2$ if $\int_Z d(x,\xi_1)\,dx < \int_Z d(x,\xi_2)\,dx$.

The region Z may coincide with the region $\aleph$ of possible measurements, may be a subset of $\aleph$, or has no point in common with $\aleph$ (i.e. an extrapolation).

[1] define D-optimality criterion design of experiment as a design that maximizes the determinant of the information matrix, $M(\xi)$, or that minimizes $\det[M^{-1}(\xi)]$. For this reason, if $\lambda_1,\ldots,\lambda_p$, are eigenvalues of $M(\xi)$, then the eigenvalues of $M^{-1}(\xi)$ are $\frac{1}{\lambda_1},\ldots,\frac{1}{\lambda_p}$, so that the design criterion can be stated as $\min \prod_{i=1}^{p} \frac{1}{\lambda_i}$ and the general measure of impression $\Psi\{M(\xi)\}$ becomes $\Psi\{M(\xi)\} = \lim_{k\to 0}\left(\frac{1}{p}\sum_{i=1}^{p}\lambda_i^{-k}\right)^{\frac{1}{k}}$. The D-optimality criterion is invariant to non-degenerate linear transformation of the model. That is, a design which is D-optimum for the model $\eta(x,\beta) = f^T(x)\beta$ is also D-optimum for the model $\eta(x,\gamma) = g^T(x)\gamma$ if $g(x) = Af(x)$ and $|A| \neq 0$ for the unknown $p\times 1$ vectors of parameters β and γ respectively. D-optimality has the advantage of invariance as linear transformations leave the D-optimum design unchanged and that the optimum designs for quantitative factors do not depend upon the scale of the variables. For this reason we give preference to the D-optimality criterion amongst other criteria. [2] proposed a numerical construction of parameter maximum D-optimal designs in which the specification of a certain range for the unknown parameters is required. Under D-optimal designs, the significance of the treatments can be tested, in general, as follows [8]:

$$H_0 : \tau_1(\xi^*) = \tau_2(\xi^*) = \ldots = \tau_k(\xi^*) = 0$$

$$\text{vs } H_1 : \text{At least one } \tau_i(\xi^*) \neq 0, \quad i = 1,2,\ldots,k,$$

at α level of significance, where ξ^* is the D-optimal design. The ANOVA table for this test is shown in Table 1.

Table 1: The ANOVA table in respect of D-optimal designs.

Source of variation	Degree of freedom	Sum of squares	Mean squares
Regression	k	$\hat{\beta}^T(\xi^*)X^T(\xi^*)X(\xi^*)\hat{\beta}(\xi^*) - N\bar{y}^2$	$\frac{SS\text{Regression}}{k}$
Residual	$N-k-1$	$\left(y - X(\xi^*)\hat{\beta}(\xi^*)\right)^T\left(y - X(\xi^*)\hat{\beta}(\xi^*)\right)$	$\frac{SS\text{Residual}}{N-k-1}$
Total	$N-1$	$y^Ty - N\bar{y}^2$	

The F statistic is computed as

$$F_{cal} = \frac{\left(\hat{\beta}^T(\xi^*)X^T(\xi^*)X(\xi^*)\hat{\beta}(\xi^*) - N\bar{y}^2\right)\Big/k}{\left(y - X(\xi^*)\hat{\beta}(\xi^*)\right)^T\left(y - X(\xi^*)\hat{\beta}(\xi^*)\right)\Big/(N-k-1)}.$$

If $F_{cal} > F_{k,(N-k-1)}(\alpha)$, H_0 is rejected; otherwise, we do not reject H_0.

2.2. Selection of a CRD model using D-optimality criterion. In this sub-section, we illustrate the use of the D-optimality criterion by considering the CRD model in equation (1) for $k = 4$ and $n_i = n = 5$. The model in equation (1) is expressed in the form of equation (2), where the regression function $\eta(x,\beta)$ takes the linear form $\eta(x,\beta) = X\beta$. Here, $\beta = (\eta,\tau_1,\tau_2,\tau_3,\tau_4)^T$ and the model matrix X is systematically arranged treatment by treatment to have 20×5 elements. The least squares estimate of β is given as $\hat{\beta} = \left(X^TX\right)^{-1}X^Ty$, where X^TX is the information matrix. By the D-optimality criterion, the larger the value

of the determinant of $X^T X$, the greater is the information contained in the design. Since different CRD models are compared, we indicate the different models by writing X as $X(\xi_\nu)$, where ξ_ν is the ν^{th} CRD. Similarly, we write β as $\beta(\xi_\nu)$. Suppose ξ^* is the design which satisfies $\min_{\forall \nu}\left(\Psi\{M(\xi_\nu)\}\right)$, $\nu = 1,\ldots,\kappa$, where $\Psi\{M(\xi_\nu)\} = \lim_{k\to 0}\left(\frac{1}{p}\sum_{i=1}^{p}\lambda_i^{-k}\right)^{\frac{1}{k}}$, then ξ^* is the D-optimum design.

We implement the following designs on Matlab:

i. the unconstrained linear CRD model;

ii. the linear CRD model with the zero-sum constraint i.e. $\sum_{i=1}^{k}\tau_i = 0$ and

iii. the linear CRD model with the baseline constraint where treatment 1 is a standard treatment and the other $(k-1)$ treatments are new treatments; for $k = 5$ and $n_i = n = 6$.

In the first case, the parameter vector $\beta(\xi_1) = (\eta \quad \tau_1 \quad \tau_2 \quad \tau_3 \quad \tau_4 \quad \tau_5)^T$ and the model matrix, $X(\xi_1)$, is coded in Matlab as

```
x1=[1 1 0 0 0 0; 1 0 1 0 0 0; 1 0 0 1 0 0; 1 0 0 0 1 0;1 00 0 0 1; 1 1 0 0 0 0; 1 0
1 0 0 0; 1 0 0 1 0 0; 1 0 0 0 1 0;1 0 0 0 0 1; 1 1 0 0 0 0; 1 0 1 0 0 0; 1 0 0 1 0
0; 1 0 0 0 1 0;1 0 0 0 0 1; 1 1 0 0 0 0; 1 0 1 0 0 0; 1 0 0 1 0 0; 1 0 0 0 1 0;1 0
0 0 0 1; 1 1 0 0 0 0; 1 0 1 0 0 0; 1 0 0 1 0 0; 1 0 0 0 1 0;1 0 0 0 0 1; 1 1 0 0 0
0; 1 0 1 0 0 0; 1 0 0 1 0 0; 1 0 0 0 1 0;1 0 0 0 0 1];
```

The model matrix, $X(\xi_2)$, for the case of the zero-sum constraint is obtained by setting $\tau_5 = -\sum_{i=1}^{4}\tau_i$ in the parameter vector $\beta(\xi_1)$ so that, using the first four rows of the product $X(\xi_1)\beta(\xi_1)$, we get

$$X(\xi_1)\beta(\xi_1) = \eta\begin{pmatrix}1\\1\\1\\1\\1\end{pmatrix} + \tau_1\begin{pmatrix}1\\0\\0\\0\\-1\end{pmatrix} + \tau_2\begin{pmatrix}0\\1\\0\\0\\-1\end{pmatrix} + \tau_3\begin{pmatrix}0\\0\\1\\0\\-1\end{pmatrix} + \tau_4\begin{pmatrix}0\\0\\0\\1\\-1\end{pmatrix}.$$

Thus $\beta(\xi_2) = (\eta \quad \tau_1 \quad \tau_2 \quad \tau_3 \quad \tau_4)^T$ and $X(\xi_2)$ is coded in Matlab as:

```
x2=[1 1 0 0 0; 1 0 1 0 0; 1 0 0 1 0; 1 0 0 0 1; 1 -1 -1 -1 -1; 1 1 0 0 0; 1 0 1 0 0; 1
0 0 1 0; 1 0 0 0 1; 1 -1 -1 -1 -1; 1 1 0 0 0; 1 0 1 0 0; 1 0 0 1 0; 1 0 0 0 1; 1 -1 -1
-1 -1; 1 1 0 0 0; 1 0 1 0 0; 1 0 0 1 0; 1 0 0 0 1; 1 -1 -1 -1 -1; 1 1 0 0 0; 1 0 1 0 0;
1 0 0 1 0; 1 0 0 0 1; 1 -1 -1 -1 -1; 1 1 0 0 0; 1 0 1 0 0; 1 0 0 1 0; 1 0 0 0 1; 1 -1 -
1  -1 -1];
```

To obtain the model matrix, $X(\xi_3)$, of the baseline constraint, we drop the column of τ_1 in matrix $X(\xi_1)$ so that $\beta(\xi_3) = (\eta \quad \tau_2 \quad \tau_3 \quad \tau_4 \quad \tau_5)^T$. The resulting model matrix $X(\xi_3)$ is coded in Matlab as

```
x3=[1 0 0 0 0; 1 1 0 0 0; 1 0 1 0 0; 1 0 0 1 0;1 0 0 0 1; 1 0 0 0 0; 1 1 0 0 0; 1 0 1 0
0; 1 0 0 1 0;1 0 0 0 1; 1 0 0 0 0; 1 1 0 0 0; 1 0 1 0 0; 1 0 0 1 0;1 0 0 0 1; 1 0 0 0
0; 1 1 0 0 0; 1 0 1 0 0; 1 0 0 1 0;1 0   0 0 1; 1 0 0 0 0; 1 1 0 0 0; 1 0 1 0 0; 1 0 0
1 0;1 0 0 0 1; 1 0 0 0 0; 1 1 0 0 0; 1 0 1 0 0; 1 0 0 1 0;1 0 0 0 1];
```

The rows and columns of the model matrices are obtained using the following: a1=size(x1)

a2=size(x2)

a3=size(x3)

The general measure of impression $\Psi\{M(\xi)\}$ for the D-optimality criterion is $\Psi\{M(\xi)\}=\lim_{k\to 0}\left(\frac{1}{p}\sum_{i=1}^{p}\lambda_i^{-k}\right)^{\frac{1}{k}}$, which simplifies as

$\Psi\{M(\xi)\}=\exp\left\{\lim_{k\to 0}\left(\frac{1}{p}\sum_{i=1}^{p}\lambda_i^{-k}\ln\frac{1}{\lambda_i}\right)\right\}=\exp\left\{\frac{1}{p}\ln\prod_{i=1}^{p}\frac{1}{\lambda_i}\right\}=\left(\prod_{i=1}^{p}\frac{1}{\lambda_i}\right)^{\frac{1}{p}}$ (see [1]), where $\lambda_1,\ldots,\lambda_p$, are eigenvalues of $\left(X^T(\xi_\nu)X(\xi_\nu)\right)$, $\nu=1,2,3$. We compute the information matrix $\left(X^T(\xi_\nu)X(\xi_\nu)\right)$ for $\nu=1,2,3$, using a software, Matlab.

3.0 Results

We obtain the following results:

$$M(\xi_1)=\left(X^T(\xi_1)X(\xi_1)\right)=\begin{bmatrix}30&6&6&6&6&6\\6&6&0&0&0&0\\6&0&6&0&0&0\\6&0&0&6&0&0\\6&0&0&0&6&0\\6&0&0&0&0&6\end{bmatrix},\ \det(M(\xi_1))=0,$$

$$M(\xi_2)=\left(X^T(\xi_2)X(\xi_2)\right)=\begin{bmatrix}30&0&0&0&0\\0&12&6&6&6\\0&6&12&6&6\\0&6&6&12&6\\0&6&6&6&12\end{bmatrix},\ \det(M(\xi_2))=194400,\ \text{and}$$

$$M(\xi_3)=\left(X^T(\xi_3)X(\xi_3)\right)=\begin{bmatrix}30&6&6&6&6\\6&6&0&0&0\\6&0&6&0&0\\6&0&0&6&0\\6&0&0&0&6\end{bmatrix},\ \det(M(\xi_3))=7776.$$

where $M(\xi_1)$, $M(\xi_2)$, and $M(\xi_3)$ are the information matrices for the designs ξ_1, ξ_2, and ξ_3, respectively. The eigenvalues, $\lambda_1,\ldots,\lambda_p$, are obtained from the characteristic equation of matrix $\left(X^T(\xi_\nu)X(\xi_\nu)\right)$ by solving $\left|\left(X^T(\xi_\nu)X(\xi_\nu)\right)-\lambda \mathrm{I}_{(k+1)\times(k+1)}\right|=0$, where $\mathrm{I}_{(k+1)\times(k+1)}$ is a $(k+1)\times(k+1)$ identity matrix. Using the Matlab package, the eigenvalues for each design are obtained from the codes:
[v1,d1]=eig(inv(D1)),[v2,d2]=eig(inv(D2)), [v3,d3]=eig(inv(D3)), and we obtain the reciprocal of the eigenvalues for each design using the following codes:

Reciprocal of *eig*(ξ_1): b1=1/d1(1,1), b2=1/d1(2,2), b3=1/d1(3,3), b4=1/d1(4,4), b5=1/d1(5,5), b6=1/d1(6,6),

Reciprocal of *eig*(ξ_2):c1=1/d2(1,1), c2=1/d2(2,2), c3=1/d2(3,3), c4=1/d2(4,4), c5=1/d2(5,5),

Reciprocal of *eig*(ξ_3): e1=1/d3(1,1), e2=1/d3(2,2), e3=1/d3(3,3), e4=1/d3(4,4), e5=1/d3(5,5),

The criterion $\Psi\{M(\xi)\} = \left(\prod_{i=1}^{p} \frac{1}{\lambda_i} \right)^{\frac{1}{p}}$, for each design, is implemented in Matlab as

criterion1=(b1*b2*b3*b4*b5*b6)^(1/a1(1,2)), criterion2=(c1*c2*c3*c4*c5)^(1/a2(1,2)),
criterion3=(e1*e2*e3*e4*e5)^(1/a3(1,2)),

wherein the most preferred CRD model is obtained using
pref=min([criterion1,criterion2,criterion3]),

The outputs from the Matlab package are as follows:
criterion1 = Inf; criterion2 = 11.4219; criterion3 = 6.0000.

4.0 Discussion

The information matrix of design ξ_2 has the largest determinant value of 194400 when compared to other designs. Based on D-optimality general measure of impression, D-optimum CRD is the design with the minimum measure (of 6) on its information matrix when compared to other designs. Hence, we choose the design ξ_3 with the baseline constraint as the D-optimum CRD for an experiment with five treatments and six replicates.

References

[1] Atkinson, A. C. and Donev, A. N. (1992). *Optimum Experimental Designs*. Oxford Science Publications, Oxford.
[2] Biedermann, S. and Dette, H. (2005). Numerical construction of parameter maximin D-optimal designs for binary response models. Retrieved from http://eprints.soton.ac.uk/41830/1/numconst.pdf
[3] Boer, E. P. J. and Hendrix, E. M. T. (2000). Global optimization problems in optimaldesign of experiments in regression models. *Journal of Global Optimization*, Vol.18: 385-398.
[4] Debusho, L. K. and Haines, L. M. (2008). V- and D-optimal population designs for the simple linear regression model with a random intercept term. *Journal of Statistical Planning and Inference*, Vol. 138: 1116-1130.
[5] Fedorov, V. V. (1972). *Theory of Optimal Experiments*. Academic Press, New York.
[6] Joshi, D. D. (1987). *Linear Estimation and Design of Experiments*. Wiley Eastern Limited, New Delhi.
[7] Liski, E. P., Mandal, N. K., Shah, K. R., and Sinha, B. K. (2002). *Topics in Optimal Design*. Springer, New York.
[8] Pazman, A. (2007). Criteria for optimal design of small-sample experiments with correlated observations. *Kybernetika*, Vol. 43, No. 4: 453-462.
[9] Wu, C. F. J. and Hamada, M. (2000). *Experiments: Planning, Analysis, and Parameter Design Optimization*. John Wiley & Sons, Inc., New York.