

Minimum Aberration and Some Optimality Criteria for Competing 2^{p-k} Resolution III Factorial Designs

Babatunde Adeniyi Agbaje¹ and Reuben Adeyemi Ipinoyomi²

Department of Statistics University of Ilorin, Nigeria.

Abstract: Searching for optimal designs among competing resolution III factorial designs needs some optimality criteria for experimenter who may prefer one criterion over another. This research work considers Variance, A-, D-, E-, and Rc-, in connection with the judgment of estimation capacity, estimation quality and minimum aberration. The method of construction for competing designs is also examined.

Key words: Construction, finite projective geometry, symmetric designs, resolutions, defining relations.

1. Email: realtunde34@yahoo.com

2. Email: raipinyomi@gmail.com



1. INTRODUCTION

Finite projective geometry method is important for generating maximum number of designs (Mukerjee and Wu, 1999) when an experimenter is searching for optimal designs. Finite number of lines and points is one of the characteristics of finite projective geometry, where any two lines meet at a point. The points are the treatment combinations and the lines show connectedness. The geometric system under finite projective geometry approach needs careful selection of independent effects which will determine the points in the system. However, exhaustive search is needed on the independent effects to be confounded having a particular resolution in mind.

Many attributes of experimental units called factors, affect the variable of interest in the experiment and the aim of an experimenter is to screen out less important ones (Jaynes 2016). Construction of designs to suit a particular number of factors are very important to solve certain problems (Arabie and Hubert 1992, Li et al., 2016) and for the study of scientific models across different fields (Flajolet and Sedgewick 2009). Having alternative designs to solve a particular problem requires finding important properties of their design matrices to determine the

optimal design (Dasgupta 2010). The following design matrix properties known as optimality criteria, **Variance**, **A-**, **D-**, **E-**, and **Rc-**, evaluating the judgment of **estimation capacity**, **estimation quality** and **minimum aberration** are used.

Estimation capacity deals with a design's ability to handle a number of main effects and levels of interaction in a model (Ching-Shui Cheng and Rahul Mukerjee 1998). Estimation quality deals with a design's ability, handling the quality of alias structure. Minimum aberration is a method for considering a design with fewer minimum word length effects in the defining relation as the best in fractional factorial designs (Ching-Shui Cheng and Rahul Mukerjee, 1998).

A fractional factorial design starts from proper identification of some structures in full factorial design (Fontana and Sampo 2013), in which generating function also known as defining contrast could be of help (Flajolet and Sedgewick 2009, Jaynes 2013). Generating function can help to partition the full factorial design into natural segments capable of solving a problem, with a reduced cost (Jaynes 2013). This function is important in the construction of fractional factorial designs. Symmetric fractional factorial design procedure which permits uncorrelated main effects estimates is considered here with an assumption that no interaction is consequential. Part of a full factorial design is selected using the concepts of modulus and generating functions. Defining contrast is a function from a full factorial design of an experiment to a set of integers defined mathematically as

$$Dc : d_f \rightarrow [0,1]$$

and

$$d_f = \begin{cases} d_0 & \text{if } Dc = 0 \pmod{2} \\ d_1 & \text{if } Dc = 1 \pmod{2} \end{cases}$$

Where Dc is a defining contrast, d_f is a full factorial design, d_0 is a fractional factorial design when $Dc = 0 \pmod{2}$, and d_1 is a fractional factorial design when $Dc = 1 \pmod{2}$ such that $d_0 \cap d_1 = \emptyset$.

One of the consequences of using fractional factorial designs in an experiment is the aliasing of effects (Jaynes 2013). Suppose k independent effects are selected

for the experiment which allows estimation of lower order interactions, design d_0 would have alias structure like

$$I_0 + D_{c01} + D_{c02} + \dots + D_{c0k}$$

equivalently written as

$$I_0 + \sum_{i=1}^k D_{c0i}$$

and design d_1 would have alias structure like

$$I_1 + D_{c11} + D_{c12} + \dots + D_{c1k}$$

equivalently written as

$$I_1 + \sum_{i=1}^k D_{c1i}$$

On the concept of limit

$$\lim_{k \rightarrow \infty} (I_0 + \sum_{i=1}^k D_{c0i}) \rightarrow \infty$$

Meaning that factor of interest is very likely to be significant when it is actually significant and

$$\lim_{k \rightarrow \infty} (I_1 - \sum_{i=1}^k D_{c1i}) \rightarrow 0$$

Meaning that factor of interest is very likely to be insignificant when it is actually significant.

Isomorphic designs are easily identified (Flajolet and Sedgewick 2009) due to the fact that they have countless properties in common like word length pattern, estimation capacity, estimation quality, degrees of freedom and so on (Cheng and Mukerjee 1998).

Every $d \in D$ and $d_1 \in D$ designs are said to be isomorphic, if the following conditions are satisfied (Cheng and Mukerjee 1998).

$$E(d_1) = E(d) \quad \forall d \in D$$

$$R(\beta, d_1) = R(\beta, d) \quad \forall \beta \in \Theta, d \in D$$

A design $d_0 \in D$ is said to have the highest distance between the levels of factors in D iff

$$R(\beta, d_0) \leq R(\beta, d) \quad \forall \beta \in \Theta, d \in D$$

This research work focus on using finite projective geometry in the construction of resolution III factorial designs to have competing designs and, likewise provide a procedure for comparison using the concept of distance and singular value decomposition.

2.

METHODOLOGY AND MATERIALS

2.1. FACTORIAL DESIGN

A model for symmetric fractional factorial experiments can be defined generally in matrix notation as

$$y = X\beta + \varepsilon \quad (1)$$

Where X is the design matrix, β is the vector containing overall mean and main effects, y is a vector of measured observations, and ε is a vector of residuals such that $E(\varepsilon\varepsilon') = \sigma^2 I$.

2.2 Definition

2.2.1 Definition : *D-Optimality Criterion*

A design $d^* \in D$ is said to be *D*-optimal in D if and only if

$$\det(C_{d^*}) \geq \det(C_d) \text{ for any } d \in \mathcal{D}.$$

2.2.2 Definition : A-Optimality Criterion

A design $d^* \in \mathcal{D}$ is said to be *A*-optimal in \mathcal{D} if and only if

$$\text{trace}(C_{d^*}) \geq \text{trace}(C_d) \text{ for any } d \in \mathcal{D}.$$

2.2.3 Definition : E-Optimality Criterion

A design $d^* \in \mathcal{D}$ is said to be *E*-optimal in \mathcal{D} if and only if

$$\min_i \{EV(C_{d^*})\} \geq \min_i \{EV(C_d)\} \text{ for any } d \in \mathcal{D}.$$

2.2.4 Definition : Rc-Optimality Criterion

A design $d^* \in \mathcal{D}$ is said to be *Rc*-optimal in \mathcal{D} if and only if

$$\text{recc}(C_{d^*}) \geq \text{recc}(C_d) \text{ for any } d \in \mathcal{D}.$$

where *EV* eigenvalue, C_d is information matrix and *recc* is reciprocal condition number (Mitchell and Bayn, 1978).

2.3 Distance

Distance is a means of determining correlation where $||.||$ stands for Euclidean matrix norm. Reliable effect estimates of interest are attained by increasing the distance between low and high level of the factors in an experiment (Montgomery 2001).

2.4 Singular Value Decomposition (SVD)

Linear technique known as Singular Value Decomposition (SVD) is helpful to derive optimal parameter values of any designs (Bourlard and Kamp 1988). SVD is an extremely complex mathematical method for matrix decomposition.

$$S = UDV'$$

Where, U is the orthonormal eigenvectors of SS' , V is the orthonormal eigenvectors of $S'S$, D is a diagonal matrix containing the singular values. The rank of a design matrix would determine the number of non-zero singular values.

2.5 Mathematical Expressions

Let X be a design matrix, where $||X||$ is called the measure of dissimilarity. Mathematically,

$$||X|| = 1 - \text{cor}(X) \text{ where } r \text{ is a correlation matrix}$$

Therefore, $||X||$ is also known as the measure of distance

$$d_{svd} = \text{svd}(||X||) = U_d \mathbb{D}_d V_d' \text{ where } x_{r_i} \in \mathbb{D}_d \text{ and}$$

$$\mathbb{D}_d = \{x_{r_1}, x_{r_2}, \dots, x_{r_p}\} \text{ (singular values)}$$

$$\text{Since } \text{rank}(X) = p$$

$$x_{r_1} \geq x_{r_2} \geq \dots \geq x_{r_p} > 0$$

2.6 Optimality Criteria

$\text{Variance}(\mathbb{D}_d) = [\sum_{i=1}^p (x_{r_i} - \bar{x}_r)] / (p - 1)$ (variation in the singular values).

$\text{efficiency}(\mathbb{D}_d) = \text{Average}(\mathbb{D}_d)$ (Average in the singular values could be geometric mean (D), arithmetic mean (A) or minimum value of the singular values (E)).

3. RESULTS AND DISCUSSION

We use R codes for the construction and estimation of the optimality values for symmetric resolution III factorial designs.

Table 1: List of Resolution III Designs for $6 \leq p \leq 7$, $n = 16$

SN	P,K	Generators	singular values (Min, Max)	Variance	D-, A-, E-, Rc- Criteria	Estimation Quality	Estimation Capacity	Word Length Pattern
1	6,2	ABC BDEF ACDEF	0.8284271, 26.5213	39.47702	[1] 1.744362 [1] 1.306727 [1] 0.8284271 [1] 0.0102389	3, 3	6,9	3,4,5
2	7,3	ABC CDE EFG ABDE ABDFG ABCEFG CDFG	0.8284271, 28.7009	46.0043	[1] 1.915367 [1] 1.430545 [1] 0.8284271 [1] 0.0102401	5, 4	7, 8,0	3,3,3,4, 4,5,6

Column 1: shows the serial number,

Column 2: shows the number of factors and the independent interaction effects.

Column 3: shows the generators of the design.

Column 4: shows the minimum and maximum of singular values.

Column 5: shows the variance of singular values (or roots) of the designs constructed

Column 6: shows the efficiency criteria of singular values (or roots) of the designs constructed.

Column 7: shows the estimation quality

Column 8: shows the estimation capacity

Column 9: shows the word length pattern

Estimation quality shows the number of times Main effects and two-order interaction effects, two two-order interaction effects, and two and three order interaction effects occur together in the alias structure.

Estimation capacity shows the number of Main effects, two-order interaction effects and three order interaction effects that a design allowed to be estimated respectively.

4. CONCLUSION

Construction of resolution III factorial designs and the selection of optimal designs are considered here. It is fast sometimes to find the optimal design by mere looking at the word length patterns in the defining relations, due to this, the concept of minimum aberration become very important to some experimenters. When an experimenter is interested on possible effects to be estimated from a design, the concept of estimation capacity is important. When an experimenter is concerned about how effects are aliased, the concept of estimation quality is the better option. Some experimenters may prefer some computations, then **Variance, A-, D-, E-, and Rc-** criteria are good in the selection of optimal designs. All these optimality criteria work effectively well in the selection of good designs.

A design that has large number of factors and independent interaction effects, tends to have large number of designs possible to generate which provide an opportunity to a researcher to select better design among many.

REFERENCES

- Bourlard and Kamp (1988). Auto-Association by Multilayer Perceptions and Singular Value Decomposition. *Biol. Cybern.* 59, 291-294
- Ching-Shui Cheng and Rahul Mukerjee (1998). Regular Fractional Factorial SDesigns with Minimum Aberration and Maximum Estimation Capacity. *The Annals of Statistics*, Vol. 26, No. 6, 2289-2300
- Hao Li, Zhengyun Zhang, Zunqiang Zhou, Xianting Ding, Gwangwen Zhou (2016). Optimal Combination of Gemcitabine Sorafenib, and S-1 shows increased efficacy in treating Cholangiocarcinoma in vitro and in vivo. *Anti-Cancer Drugs* 27,7,600
- Jessica Jaynes, Xianting Ding, Hongquan Xu, Weng Kee Wong and Chih-Ming Ho (2013). An Application of Fractional Factorial Designs to Study Drug Combinations. *Quality and Reliability Engineering International* Vol. 32, 307-318
- Jessica Jaynes, Yitong Zhao, Hongquan Xu, and Chih-Ming Ho (2016). Use of Orthogonal Array Composite Designs to Study Lipid Accumulation in a Cell-Free System. *Quality and Reliability Engineering International* Vol. 32, 1965-1974
- Montgomery (2001). *Design and Analysis of Experiments*. 222 Rosewood Drive, Danvers, MA 01923, (508) 750-4470

Mukerjee and Wu (1999). Blocking in Regular Fractional Factorials: A Projective Geometric Approach. The Annals of Statistics, Vol. 27, No. 4, pp. 1256-1271

Nairanjana Dasgupta, Mike Jacroux, Rita SahaRay (2010). Partially Replicated Fractional Factorial Designs. Metrika Vol. 71, pp. 295-311

Phipps Arabie and Lawrence J. Hubert (1992). Combinatorial Data Analysis. Annu. Rev. Psychol. 43: 169-203

Philippe Flajolet and Robert Sedgewick (2009). Analytics Combinatorics. Cambridge University Press ISBN 978-0-521-89806-5

Roberto Fontana and Sabrina Sampo (2013). Minimum-Size Mixed-Level Orthogonal Fractional Factorial Designs Generation: A SAS Based Algorithm. Journal of Statistical Software Vol. 53

T. J. Mitchell and C. K. Bayne (1978). D-Optimal fractions of Three-Level Factorial Designs. Technometrics, Vol. 20, No. 4, Part 1, pp. 369-380

IJSER