

A-Optimal Split Plot Design for Estimating Variance Components

Nuga O.A^{1,*}, Amahia G.N², Fakorede A³

¹Department of Physical Sciences, Bells University of Technology, Ota, Nigeria

²Department of Statistics, University of Ibadan, Ibadan, Nigeria

³Department of Computer Science, Bells University of Technology, Ota, Nigeria

*Corresponding author: oluwolenuga@yahoo.com

Received June 16, 2014; Revised July 07, 2014; Accepted August 07, 2014

Abstract We study A-Optimal Split Plot designs for the maximum likelihood estimators of variance components. The work used the general linear model with one whole plot factor and one sub-plot factor and assumed that both factor effects are random variables. Candidates designs with the same number and sizes of whole plot were assigned to the level of the whole plot factor in such a way that formed a balanced one way design. Using the variances of the maximum likelihood estimator of the variance components, candidates designs were compared for A- optimality. The work introduced a method of classifying the five variance components to make comparison and presentation meaningful. The resulting optimal designs depend on the true proportional value of the variance components.

Keywords: local optimality, A-Optimality, whole plot, sub plot, variance component

Cite This Article: Nuga O.A, Amahia G.N, and Fakorede A, "A-Optimal Split Plot Design for Estimating Variance Components." *American Journal of Applied Mathematics and Statistics*, vol. 2, no. 4 (2014): 235-238. doi: 10.12691/ajams-2-4-11.

1. Introduction

Optimal designs for variance components model have been discussed fairly in experiment that are ran in a completely random order. For the one way model Hammersly (1949), Crump (1954), Anderson & Crump (1967) were some of the earliest authors. Hammersly (1949) showed that for a fixed N, the variance $Var(\sigma_\alpha^2)$ is minimized by allocating an equal number n, of observation to each class where $n = \frac{N\rho + N + 1}{N\rho + 2}$, since

this formula may not yield an integer value, it was suggested that the closest integer value for n be chosen. Crump (1954) and Anderson & Crump (1967) showed that for fixed K and N, $Var(\sigma_\alpha^2)$ is minimized when

$n_i = n = \frac{N}{a}$ for all i. The optimal value for a in this case

is given as $a_1 = \frac{N(N\rho + 2)}{N(\rho + 1) + 1} = \frac{N}{n}$

Other authors are Kussmaul & Anderson (1967), Thompson and Anderson (1975), Herrendofer (1979), Murkerjue & Huda (1988), Giovagnoli & Sebastiani (1989), Norell (2006). Norell (2006) studied design effect for the one way random model of the maximum likelihood estimators.

The construction of optimal design for the two way crossed models seems to have been considered first by Gaylor (1960). He considered the problem of optimal

designs to estimate variance components using the fitting constant method of estimation of variance components for the unbalanced data. Bush (1962) and Bush and Anderson (1963), HIrotsu(1966), Mostafa (1967) are some of the other contributors to the designing experiment using the two way random model

Some pioneering articles that address the problem of estimating variance components in a nested classification are Bainbridge (1965) Prairie (1962), Prairie and Anderson (1962), Bainbridge (1965), they proposed designs that systematically spread the information in the experiment more equally among the variance components. Goldsmith and Gaylor (1970) carried out extensive investigation on optimal designs for estimating variance components in a completely random nested classification. Delgado (1999) defined a class of unbalanced design for estimating variance components in the three stage nested classification using the ANOVA method of estimation.

Loeza-Serrano. S and A. Donev (2012) constructed D-optimal design for variance components estimation in a three stage crossed and nested classification.

For experiments that include both crossed and nested factor in the same model, no assumption of a complete random model has been made. Work that design experiment for variance components estimation are based on the linear mixed effect model. Beverly (1981), Ankenman, Liu, Karr, and Picka (2001) and Aviles and Pinheiro (2001) are authors that have published work. However experiments that complete randomization of order of runs is not feasible or might be too expensive to use is performed using split plot models.

1.1. Split Plot Design

Split plot designs originally developed by Fisher (1925) for use in agricultural experiments are basically the modified form of randomized block designs. These designs are used in situations where complete randomization of runs within block is not possible. These designs are used widely in industrial experiments, experiments where one set of factors may require a large amount of experimental materials(Whole Plot factors) while another set of factors might be applied to smaller experimental materials (Sub Plot factors). Another situation that leads to the use of split plot designs in industrial experiments is when there exist one or more factors called hard to change factors that are expensive or time consuming to change level settings (WP factors) and the other factors (SP factors) whose level settings are easier to change are called Easy to change factors.

In general split plot design can be used for any experimental situation that involves two different types of experimental unit (large and small) randomly assigned independently at the two different level

The optimal design for split-plot experiments has received attention by Goos and Vandebroek (2001b, 2003a, 2004). The work was based on a more complex design structure that used the first and second order polynomial model to represent the response, but in general the work used a linear mixed effect model which assumed fixed effects for the settings of the whole plot factors and sub-plot factors and two variance components associated with the whole plot errors and sub-plot errors.

The study constructed A- optimal designs for maximum likelihood estimators in a split plot experiment with one whole plot (WP factors) and one sub plot (SP factor) with the assumption of random effect for both factors.

2. Model and Variance Structure

The model equation for the split plot design with one WP factor (Factor A) and one SP factor (Factor B) can be written as

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma_{ik} + e_{ijk} \tag{1}$$

$i = 1, 2, \dots, a, j = 1, \dots, b, k = 1, \dots, r_i$

r_i is the number of whole plot at i_{th} level of factor A, but $r_1 = r_2, \dots = r_a = r$ since equal number of whole plot are allocated to whole plot factor A. There are ar whole plot of equal sizes available.

y_{ijk} is the response at the k_{th} replicate of the i_{th} level of factor A and the j_{th} level of factor B, μ is the general mean, α_i is the effect of the i_{th} level of factor A, β_j is the effect of the j_{th} level of factor B. $(\alpha\beta)_{ij}$ is the interaction effect of the i_{th} level of factor A and the j_{th} level of factor B. γ_{ik} is the error term of the k_{th} replicate of the i_{th} level of factor A (WP error term), e_{ijk} is the error term corresponding to individual y_{ijk} (SP error term). the random variables

$\alpha_i, \beta_j, (\alpha\beta)_{ij}, \gamma_{ik}, e_{ijk}$ and are assumed to be normally distributed with zero mean variance $\sigma_\alpha^2, \sigma_\beta^2, \sigma_{\alpha\beta}^2, \sigma_\gamma^2, \sigma_e^2$, respectively. In matrix form the model can be written as

$$Y = \mu 1 + Z_1\gamma_1 + Z_2\gamma_2 + Z_3\gamma_3 + Z_4\gamma_4 + Z_5\gamma_5 \tag{2}$$

is a vector of abk observations, μ is the overall mean, z_i is an indicator matrix associated with the i_{th} variance component, γ_i is a vector of normally distributed random effects associated with the i_{th} variance component such that $w_i \approx N(0, \sigma_i^2 I)$. The variance matrix of observations can be written as

$$Var(Y) = V = \sum_{i=0}^5 \sigma_i^2 Z_i Z_i' \tag{3}$$

2.1. Large Sample Variance of the MLE

Maximum likelihood estimates of variance components cannot be obtain explicitly except for some balanced data, but their large sample asymptotic dispersion matrix can be derived. It is known that the large sample asymptotic dispersion matrix of the maximum likelihood estimators for any model is the inverse of the information matrix. This matrix is the negative of the expected value of the second order partial derivatives (Hessian Matrix) with respect to the parameters of the log-likelihood function.

For our data vector $y \approx N(X\beta, V)$, the likelihood function is

$$L = L(\beta, V / y) = \frac{e^{-\frac{1}{2}(y-X\beta)'V^{-1}(y-X\beta)}}{(2\pi)^{\frac{1}{2}N} \left| V \right|^{\frac{1}{2}}} \tag{4}$$

Upon taking the log of the likelihood

$$l = \text{Log}L = -\frac{1}{2}N \log 2\pi - \frac{1}{2} \log |V| - \frac{1}{2}(y-X\beta)'V^{-1}(y-X\beta)$$

The matrix is given as

$$\begin{bmatrix} \frac{\partial^2 l}{\partial \beta \partial \beta'} & \frac{\partial^2 l}{\partial \beta \partial \sigma^2} \\ \frac{\partial^2 l}{\partial \beta \partial \sigma^2} & \frac{\partial^2 l}{\partial \sigma^2 \partial \sigma^2} \end{bmatrix} \tag{5}$$

which gives the following

$$In \begin{bmatrix} \beta \\ \sigma^2 \end{bmatrix} = \begin{bmatrix} X'V^{-1}X & 0 \\ 0 & \frac{1}{2}tr(V^{-1}Z_i Z_i' V^{-1}Z_j Z_j') \end{bmatrix} i, j = 1, \dots, 5 \tag{6}$$

the model for this work involve only the variance components and therefore the information matrix become

$$In[\sigma^2] = \frac{1}{2} \left[\sum_m tr(V^{-1}Z_i Z_i' V^{-1}Z_j Z_j') \right] i, j = 1, \dots, 5.$$

Using Searle (1970) this reduces to

$$\frac{1}{2} [{}_m \text{tr}(V^{-1}Z_i Z_i' V^{-1} Z_j Z_j')] = \frac{1}{2} [{}_m \text{sesq}(Z_i' V^{-1} Z_j)] , j = 1, \dots, 5.$$

The inverse of v is also obtained using the results of Henderson and Searle (1979),

$$V^{-1} = \theta_0^{-1} (I_a \otimes C_b \otimes C_r \otimes C_n) + \theta_1^{-1} (C_a \otimes C_b \otimes \bar{J}_r \otimes \bar{J}_n) + \theta_2^{-1} (I_a \otimes \bar{J}_b \otimes C_r \otimes \bar{J}_n) + \theta_3^{-1} (C_a \otimes \bar{J}_b \otimes \bar{J}_r \otimes \bar{J}_n) + \theta_4^{-1} (\bar{J}_a \otimes C_b \otimes \bar{J}_r \otimes \bar{J}_n) + \theta_5^{-1} (\bar{J}_a \otimes \bar{J}_b \otimes \bar{J}_r \otimes \bar{J}_n) \tag{7}$$

Where

$$\theta_0 = \sigma_e^2, \theta_1 = (\sigma_e^2 + bn\sigma_\gamma^2), \theta_2 = (\sigma_e^2 + rn\sigma_{\alpha\beta}^2), \theta_3 = (\sigma_e^2 + rn\sigma_{\alpha\beta}^2 + arn\sigma_\beta^2), \theta_4 = (\sigma_e^2 + bn\sigma_\gamma^2 + rn\sigma_{\alpha\beta}^2 + brn\sigma_\alpha^2), \theta_5 = (\sigma_e^2 + bn\sigma_\gamma^2 + rn\sigma_{\alpha\beta}^2 + brn\sigma_\alpha^2 + arn\sigma_\beta^2)$$

a computational information matrix can thereafter be obtainedSearle et al (2006) pg 247.

$$\begin{bmatrix} \sigma_e^2 \\ \sigma_\gamma^2 \\ \sigma_{\alpha\beta}^2 \\ \sigma_\beta^2 \\ \sigma_\alpha^2 \end{bmatrix} = \begin{bmatrix} t_{ee} & t_{\gamma\gamma}/bn & t_{\alpha\beta}/rn & t_{\beta\beta}/arn & t_{\alpha\alpha}/arn \\ & t_{\gamma\gamma} & t_{\alpha\alpha}r & abrn^2/\theta_5 & t_{\alpha\alpha}/r \\ & & t_{(\alpha\beta)}^2 & t_{\beta\beta}/\alpha & t_{\alpha\alpha}/b \\ & symmetric & & t_{\beta\beta\alpha} & ab(rn)^2/\theta_5 \\ & & & & t_{\alpha\alpha} \end{bmatrix} \tag{8}$$

The diagonal element are

$$t_{ee} = \frac{v_e}{\theta_0^2} + \frac{v_{rr}}{\theta_1^2} + \frac{t_{\alpha\beta}}{(rn)^2}, t_{\alpha\alpha} = b^2 (rn)^2 \left(\frac{v_\alpha}{\theta_4^2} + \frac{1}{\theta_5^2} \right), t_{\beta\beta} = a^2 (rn)^2 \left(\frac{v_\beta}{\theta_3^2} + \frac{1}{\theta_5^2} \right), t_{\alpha\beta} = (rn)^2 \left(\frac{v_{\alpha\beta}}{\theta_2^2} + \frac{v_\beta}{\theta_3^2} + \frac{v_\alpha}{\theta_4^2} + \frac{1}{\theta_5^2} \right), t_{rr} = (bn)^2 \left(\frac{v_r}{\theta_1^2} + \frac{v_\alpha}{\theta_4^2} + \frac{1}{\theta_5^2} \right), \tag{9}$$

The V's are the degree of freedom corresponding to individual sources of variation.

The A-Optimal design require the minimization of the trace of the variance-covariance matrix, but by studying the information matrix, its required to maximize the trace.

3. Design Optimality and Generation

As stated above we seek the design that maximizes the trace of the optimality criterion information matrix of the five variance components. Optimal design in a linear random effects model depends on the relative size of the true values of the variance components; we borrow an idea from optimization on theory of nonlinear models and

use the local optimality. The five variance components were classified into two sets, the first set consist of the main effects and interaction variance components (MIVC)

which consist of σ_α^2 , σ_β^2 , and $\sigma_{\alpha\beta}^2$. The second set include the whole plot and sub plot error variance components (WSEVC) σ_γ^2 and σ_e^2 .

The work will initially assign proportional value of variance components to the two sets in such a way that the sum equals one, and thereafter distribute proportional value to each set based on initial allocation. As an example

$$\text{MIVC} = \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_{\alpha\beta}^2 = 0.5 \quad \text{WSEVC} = \sigma_\gamma^2 + \sigma_e^2 = 0.5, \text{ sub assigning MIVC and WSEVC, we have } \sigma_\alpha^2 = 0.05, \sigma_\beta^2 = 0.40, \sigma_{\alpha\beta}^2 = 0.05 \text{ and } \text{WSEVC} = \sigma_\gamma^2 = 0.37, \sigma_e^2 = 0.13$$

The optimality of the design depends on the proportion of the true value of the variance components and not the total variance components. Proportional value of the variance components will be assigned to each of the five variance components. We employed this approach to make comparison easy and establish a pattern for replicable comparison

A MATLAB code was written in the context of the information matrix of section (2.1) in such a way that enumerated design for a particular number and sizes of whole plot can be compared for A-Optimality based on any configuration of the true values of the variance components.

3.1. Design Generation

We will assigned $ar = R$ whole plots of equal size randomly to the level of the whole plot factor such that equal number of whole plots is assigned to individual levels of whole plot factor. Each level of the sub plot factor is applied once within each whole plot and one observation is measured within each sub plot. The resulting design structure is balanced.

Table 1. Design Space for some R

Number of whole plot	Possible Designs	Number of whole plot	Possible Designs
6	[a=3, r=2] [a=2, r=3]	15	[a=3, r=5] [a=5, r=3]
8	[a=4, r=2] [a=2, r=4]	16	[a=2, r=8] [a=8, r=2] [a=4, r=4]
10	[a=5, r=2] [a=2, r=5]	18	[a=2, r=9] [a=9, r=2] [a=3, r=6] [a=6, r=3]
12	[a=6, r=2] [a=2, r=6] [a=3, r=4] [a=4, r=3]	20	[a=2, r=10] [a=10, r=2] [a=4, r=5] [a=5, r=4]
14	[a=7, r=2] [a=2, r=7]		

The number of possible designs (design space) equals the total number of ways to partition $ar = \sum_i^a r_i$ subject to $r_i = r_2 = \dots = r_a$ and $2 \leq a < ar$. The assignment of

$ar = R$ whole plots to the levels of whole plot factor formed a balanced one way design.

3.1.2 An Algorithm

1. List the possible design for a fixed number and size of whole plot (Generate the design space).
2. Specify the available information about individual variance components. i.e. proportional value available to MIVC and WSEVC. Such that $MIVC+WSEVC=1$.
3. Redistribute the proportional above to the variance components within each set.
4. Calculate the criterion of optimality(A-optimal) for all design in the design space using the MATLAB programme code and identify the A-optimal design.
5. By making no changes to MIVC and based on the A-optimal design above, one can obtain the regions where some other designs in the design space is optimal. i.e. When $a \leq r$ based on the optimal design identified in (4) decreasing the proportional value of σ_γ^2 by 0.01 and increasing the proportional value of σ_e^2 by 0.01 in sequence until another design in the design space is optimal. When $a \geq r$ based on the optimal design obtain in (4) increasing the proportional value of σ_γ^2 by 0.01 and decreasing the proportional value of σ_e^2 by 0.01 in sequence until another design in the design space is optimal.

Example: Consider the situation where we have 6 whole plot of size 2, following the steps in the algorithm the list of designs in the design space is {a=2, b=2, r=3} and {a=3, b=2, r=3}

MIVC=0.5 and WSEVC=0.5 by (2) and applying (3) above individual variance components as

$$\left(\begin{array}{l} \sigma_\alpha^2 = 0.05, \sigma_\beta^2 = 0.40, \sigma_{\alpha\beta}^2 = 0.05, \\ \sigma_\gamma^2 = 0.26, \sigma_e^2 = 0.24 \end{array} \right)$$

Using the MATLAB code the D-optimal is {a=3, b=2, r=2}. For this design we increase the proportion value of the whole plot error variance components from 0.26 and reduce the proportional value of the sub plot error variance components from 0.24 in sequence by 0.01, there is a change in the D-optimal design at. In general for a fixed MIVC, the range of proportional value for which the two designs are optimal is given below

Table 2. Ranges of Optimality for R=6

Designs	Whole plot	Sub plot
{a=2, b=2, r=3}.	$0.38 \leq \sigma_\gamma^2 \leq 0.49$	$0.01 \leq \sigma_e^2 \leq 0.12$
{a=3, b=2, r=2}.	$0.01 \leq \sigma_\gamma^2 \leq 0.37$	$0.13 \leq \sigma_e^2 \leq 0.49$

4. Conclusion

We provided a procedure for constructing A- optimal design for estimating variance components in a split plot model that has one whole plot factor and one sub plot factor. The main problem studied is how to assign a given number of whole plots to the level of the whole plot factor in such a way that a balanced one way design is formed.

By following the algorithm in section, one can construct Local A- Optimal designs for other selected number and sizes of whole plot.

The extension of the work to situations where there is an unbalanced one way structure in the assignment of whole plot to the whole plot factor is a subject of current research by the authors.

References

- [1] Anderson R.L and Crump P.P (1967), Comparisons of designs and estimation procedures for estimating parameters in two stage nested process, *Technometrics*, 9, 499-516.
- [2] Ankenman, B. E., Liu, H., Karr, A. F. and Picka, J. D. (2001). A class of experimental designs for estimating a response surface and variance components. *Technometrics* 44, 45-54.
- [3] Ankermann B. E., Avilles, A.I. , Pinheiro J.C., (2003), Optimal Design for Mixed Effects Models With Two Random Nested Factors, *Statistica Sinica*, 13, 385-401.
- [4] Bainbridge T.R (1965), Staggered nested designs for estimating Variance components, *Industr. Quality Control*, 22-1, 12-20.
- [5] Crump S.L (1951) The present Status of variance components Analysis, *Biometrics*, 7, 1-16.
- [6] Delgado, J., Iyer, H., 1999. Search for optimal designs in a three stage nested random model. *Statistics and Computing* 9, 187-193.
- [7] Giovagnoli & Sebastiani (1989), Experimental Design for mean and variance estimation in variance component models, *Computational Statistics*. Data Analysis. 8, 21-28.
- [8] Goldsmith, C. H. and Gaylor, D. W. (1970). Three stage nested designs for estimating variance components. *Technometrics* 12, 487-498.
- [9] Goos, P., Vandebroek, M., 2003, D-optimal Split-Plot Designs with Given Numbers and Sizes of Whole Plots, *Technometrics*, 45, 235-245
- [10] Goos, P., Vandebroek, M., 2001, Optimal Split-Plot Designs, *Journal of Quality Technology*.
- [11] Graybill, F.A (1976) Theory and Application of Linear Models, Duxbury, North Scituate, MA.
- [12] Hammersley J.M (1949), The Unbiased estimate and standard error of interclass Varince. *Metron*, 15, 189-205.
- [13] Khuri A.I., (2000), Designs for Variance Components Estimation, *International Statistica Review* 68, 311-322
- [14] Loeza S, Donev A.N (2014). Construction of Experimental Design to estimate Variance Components. *Computational Statistics and Data Analysis* 71, 1168-1177.
- [15] Norell L. (2006) Optimal Designs for Maximum Likelihood Estimators in One way Random Model, U.U.D.M report 2006: 24, Department of Mathematics, Upsalla University.
- [16] Sahai H, Ojeda M., 2004 Analysis of Variance for Random Models. Vol. 1. Theory, Applications and Data Analysis, Birhauser Boston.
- [17] Sahai H, Ojeda M., 2004 Analysis of Variance for Random Models. Vol. 2. Theory, Applications and Data Analysis, Birhauser Boston.
- [18] Searle, S. R., Casella, G. and McCulloch, C. E. (1992). *Variance Components*. Wiley, NewYork.
- [19] Searle S.R, Harold V. Anderson (1979) Dispersion matrices for Variance Components Models, *Journal of American Statistical Association*. Vol. 74, No 366,465-470.